TABLE OF CONTENTS

SYMPOSIUM: WHEN THE STATE SPEAKS, WHAT SHOULD IT SAY?

Legitimacy and Autonomy: Values of the Speaking State  
Frank I. Michelman  985

Freedom of Expression and the Golden Mean  
Steven G. Calabresi  1005

Democratic Rhetoric: How Should the State Speak?  
Josiah Ober  1015

You’re All Individuals: Brettschneider on Free Speech  
Andrew Koppelman  1023

Liberty, Equality, and State Responsibilities: Review of Corey Brettschneider’s When the State Speaks, What Should it Say?  
Robin West  1031

The Liberal Tightrope: Brettschneider on Free Speech  
Sarah Song  1047

Democratic Persuasion and Freedom of Speech: A Response to Four Critics and Two Allies  
Corey Brettschneider  1059

ARTICLES

Discovery and Darkness: The Information Deficit in Criminal Disputes  
Ion Meyn  1091

Cruel and Invisible Punishment: Redeeming the Counter-Majoritarian Eighth Amendment  
Aliza Cover  1141
“Good Fences Make Good Neighbors”: An Environmental Justice Framework to Protect Prohibition beyond Reservation Borders

NOTES

The Securities Act of 1933: A Jurisdictional Puzzle


Following Suit with the Second Circuit: Defining Gambling in the Illegal Gambling Business Act

Prevent Defense: Will the Return of the Multiyear Scholarship Only Prevent the NCAA’s Success in Antitrust Litigation?

The JOBS Act and Crowdfunding: How Narrowing the Secondary Market Handicaps Fraud Plaintiffs

Sean J. Wright 1197

Stephen O'Connor 1233

John H. Runne 1265

Kaitlyn Dunphy 1295

Vincent J. DiFortе 1333

Sherief Morsy 1373
INTRODUCTION

A. Aims of This Essay

When, if ever—on behalf of what sorts of causes, using what sorts of instruments and powers—ought the state engage in political debates, or be required, invited, or permitted by law to do so? In When the State Speaks, What Should It Say?,¹ Corey Brettschneider’s message is that the good state will sometimes find itself obliged, prudentially and morally, to assume an active role as partisan in highly charged political controversies. Brettschneider urges, accordingly, that we will do best to read—and, as necessary, to change—our constitutional law to keep or bring it in line with that view.

These propositions are contentious. In support of them, Brettschneider urges that riding on their acceptance is no lesser a stake than the legitimacy of the state. Surprisingly, though, our author nowhere really stops to spell out what “legitimacy” is—to say what it means to judge a political practice as legitimate or not, or why we should care about such judgments. My aim in this commentary is partly supplementary: to fill this gap in Brettschneider’s exposition. It is also partly critical: to see how (or whether) the gap can be filled in a way that will actually support the philosophical and constitutional-legal

---

¹ Robert Walmsley University Professor, Emeritus, Harvard University. I thank Richard Fallon for his helpful comments.

¹ COREY BRETTSCHNEIDER, WHEN THE STATE SPEAKS, WHAT SHOULD IT SAY?: HOW DEMOCRACIES CAN PROTECT EXPRESSION AND PROMOTE EQUALITY (2012).
stances of *State Speaks*. I think it can be done, but perhaps only by ascribing to Brettschneider a somewhat novel (but I do not say necessarily mistaken) conflation of a liberal-theoretic notion of normative political legitimacy with a personal-ethical value of autonomy.

B. A State Duty of Political Partisanship?

Provocative as Brettschneider’s claims surely are meant to be, they nevertheless start out from a set of premises that will likely strike most readers as both innocuous and right-minded. Somewhat loosely restated, and pending more precise analysis to follow below, these initial premises run as follows.

State political regimes, including democratic ones, are designed to impose a measure of coercive rule on citizens. The legitimacy (in the sense of the moral justification) of any state’s practice of coercive rule is closely tied to the visibility within that practice of a commitment to conform its operations to certain principles of regard for the freedom and equality of citizens.\(^2\) Actors in civil society—persons and organizations—nevertheless can, and they sometimes do, emit messages of contradiction and rejection of these legitimacy-sustaining principles. Such messages and the ideas they carry are what the book calls “hateful.”\(^3\) Persons and groups sometimes send around hateful ideas by the sentences they utter, and sometimes by the policies and practices they adopt and follow. Of course there will always be controversy over which words and practices do and do not cross the line of contradiction of the legitimacy-sustaining values of free and equal citizenship (partly because there will always be controversy over exactly how to define those values).\(^4\) But wherever you or I would draw the line of contradiction, we might be tempted, at least, by this key thought of Professor Brettschneider: where the state stands silent in the face of the exhibition of ideas that flagrantly cross the line of contradiction, that silence can possibly cloud the visibility and certainty of the state’s own commitment to the contradicted values, with attendant risk of a resultant loss of societal attachment to those values.\(^5\)

All of that, as far as it goes, will doubtless make good enough sense to most readers. Where many will want to enter a

\(^2\) See, e.g., *id.* at 49.

\(^3\) See *id.* at 1 (classifying as “hateful” those views or viewpoints that express “an idea or ideology that opposes free and equal citizenship”).

\(^4\) See *id.* at 47, 90.

\(^5\) See, e.g., *id.* at 39.
strong cautionary reservation is at the point of the book’s further claim: that the state stands under some kind of necessitation to depart from a posture of silence in the face of hateful speech.\(^6\) That does not follow, the doubters will say, for two interconnected reasons: primarily and decisively, because the “silence” in question is in line with a principle of state neutrality on the field of political debate, which is itself a basic (or even a legitimacy-sustaining) commitment of liberal democracy;\(^7\) secondarily and supportively, because attentive observers will accordingly construe the state’s silence as an act of fidelity to that principle and therefore not as the state’s condonation of hateful views.\(^8\)

And so we draw near to what is controversial in Professor Brettschneider’s book. In \textit{State Speaks}, Brettschneider asserts the state’s obligation—“obligation” is his repeated word of choice\(^9\)—to respond to hateful speech by its own engagement in a type of expressive activity that he labels “democratic persuasion.”\(^10\) Democratic persuasion comprises the state’s use of a congeries of communicative means to denounce and oppose the propagation of “hateful” ideas that plainly and sharply contradict a somewhat thinly defined, “political” ideal of free and equal citizenship.\(^11\) By Brettschneider’s argument, the state is profoundly duty-bound to defend that political ideal against attack and attrition, for the particular reason that a continued public commitment to it provides the necessary foundation for the state’s legitimacy: “If . . . hateful doctrines were left to prevail, they could subvert the basic principles of a legitimate democratic state.”\(^12\)

\(^6\) See, e.g., \textit{id.} at 6 ("[T]he state has the obligation to use its expressive capacities to defend the values of free and equal citizenship . . . .").

\(^7\) See \textit{id.} at 9 (describing, while rejecting, “neutralism” as the political doctrine that “the state should not promote or express any particular set of values”).

\(^8\) \textit{But see id.} at 43-44 (asserting that “[a] state that fails to answer” hateful views would “risk being seen as . . . complicit in” those views); \textit{id.} at 84 (same).

\(^9\) See, e.g., \textit{id.} at 6, 47, 111, 122. Brettschneider also sometimes (and apparently equivalently) speaks in this connection of a “-duty” of the state. \textit{See, e.g., id.} at 18, 114, 119.

\(^10\) See \textit{id.} at 25 ("[T]he state rightly engages in democratic persuasion when it exercises its expressive capacity to promote the values of free and equal citizenship.").

\(^11\) See \textit{id.} at 14 (initially defining this strictly political ideal, as distinct from “comprehensive” conceptions of metaphysical or social equality); \textit{id.} at 42-49 (recounting various expressive means available to the state for defending the ideal against attack).

\(^12\) \textit{Id.} at 7. The message recurs in the book like a drum-beat. \textit{See id.} at 4 ("These democratic . . . values of free and equal citizenship should be . . . promoted by the state because they ground the legitimacy of government and justify protecting rights . . . ."); \textit{id.} at 18 (ascribing to any “legitimate democracy” a duty to promulgate values that justify rights); \textit{id.} at 34 (calling “central to legitimacy” the idea that “citizens should all be treated as free and equal”); \textit{id.} at 39 (explaining when considerations of “democratic legitimacy” require efforts by the state to influence public
approved means of defense encompassed by democratic persuasion stop short of direct coercive muzzling of hateful speech, but they do include viewpoint-discriminatory grants and refusals of state financial support that many might construe as punitive or regulatory, in purpose or effect.\textsuperscript{13}

This positive duty to defend might not be strictly a \textit{legal} one. While Brettschneider affirms that “the ideal of democratic persuasion . . . provides a guide to identify when state speech is appropriate,”\textsuperscript{14} I have found nothing in \textit{State Speaks} to say that a court should ever treat a state’s failure to speak as a violation of constitutional law.\textsuperscript{15} Rather, when Brettschneider refers to a state’s obligation, it seems he has in mind a higher calling of moral or ethical responsibility.\textsuperscript{16} Be that as it may, Brettschneider plainly does mean for recognition of this calling to carry over into our practice of judicially enforced constitutional law, at least to this extent: he means for it to serve as a decisive reason for \textit{not} reading the Constitution in ways that would \textit{prevent} the state’s due engagement in democratic persuasion—for why, in other words, “it should be constitutionally permissible for the state to speak in favor of values of free and equal citizenship . . . .”\textsuperscript{17}

\section*{C. A Question of Legitimacy?}

Readers will ponder and doubtless will differ over how far, if at all, our author’s claims and proposals depart from
broad-sense liberal political morality, political wisdom, and sound constitutional law. Those will not be my questions here.\textsuperscript{18} I shall rather be looking upstream in Brettschneider’s argumentation, specifically at his reliance on a political–theoretic notion of legitimacy to help him move his proposals safely past predictable resistance from (some) liberal ideologists and American constitutional lawyers.

Brettschneider writes, after all, as a dedicated disturber of a widely prevailing, prescriptive imagery of the American state as neutral umpire in a free contest of ideas. He classes his work as critical in tenor, fraught with calls for changes (more politely “reinterpretations”) in the extant jurisprudence of the Supreme Court.\textsuperscript{19} He explains with care how his interventions will likely be found to deviate from American constitutional law as currently construed by the Court.\textsuperscript{20} He faults the Court’s work in general for excessive submission to a doctrine of state neutrality.\textsuperscript{21} He expects his views to be contentious among liberal mainstreamers, with whose outlook he does himself, as a liberal, feel a strong tug of sympathy.\textsuperscript{22} He anticipates objection both from those who worry that viewpoint-discriminatory deployment of the state’s vocal powers (and of course even more so its powers of the purse) is already past the limit of coercive control over political opinions that is tolerable in a democracy,\textsuperscript{23} and from those who

\textsuperscript{18} For what it is worth, I would generally associate myself with the views of Professor Calabresi in this Symposium, see Steven Calabresi, \textit{Freedom of Expression and the Golden Mean}, 79 Brook. L. Rev 1005 (2014).

\textsuperscript{19} See Brettschneider, supra note 1, at 144.

\textsuperscript{20} See id. at 98-99 (faulting the Supreme Court’s application of the Bill of Rights in Wisconsin v. Yoder, 406 U.S. 205 (1972)); see also id. at 115-25 (faulting the Court’s neutralist groundings for decisions in regard to the state’s management of “limited public forums,” its impositions of conditions on recipients of state funds, and its selective refusals of display-space on state real estate).

\textsuperscript{21} See id. at 47 (finding the Court’s “jurisprudence” to be “often couched in excessively value-neutral terms”).

\textsuperscript{22} Brettschneider defends a strict constitutional rule against viewpoint-discriminatory restrictions on speech. He thus takes sides with civil-libertarian liberals, against objections that such a rule effectively bars the state from its most straightforward means of combating the kind of “hateful” speech that he, himself, sees as posing a threat to values essential to legitimacy. His position is that, at least as long as a regime is committed to the state’s due engagement in democratic persuasion, its incorporation of a strict rule against coercive restrictions on viewpoints counts distinctly in favor of its legitimacy. See Brettschneider, supra note 1, at 80-81, 105; infra Part II.A.4.

\textsuperscript{23} See Brettschneider, supra note 1, at 7 (addressing those “who are concerned about excessive state power”); at 12 (same); at 94 (addressing the view that “state action—even when limited to expression—falls into the category of coercion per se”).
start with a strong intuition that, coercion entirely aside, strict state neutrality is essential in a just political regime.24

Neither group will be initially disposed to draw much comfort from the suggestion that the state is morally licensed to speak only in defense of “the right values.”25 And how, then, will our author hope to convert the liberal doubters? I have already told you a key part of the answer: “Legitimacy!” A due regard for that sovereign political-moral value should lead reluctant liberals toward acceptance of a departure from “neutralism” in deference to a higher necessity of the state’s due engagement in democratic persuasion.26

I. THE ARGUMENT FROM LEGITIMACY

A. A Standard Liberal Discourse of Legitimacy

Whatever Brettschneider means by “legitimacy,” it must be a value of relatively urgent concern to liberals. But what exactly is this value?

Common political discourse applies the terms “legitimate” and “legitimacy” to any of a number of different objects.27 These can range from particular acts or behaviors of specific institutional components of a legal system (say, a decision of the Supreme Court), to entire political orders or regimes, as perhaps represented by a country’s constitution or body of constitutional laws.28 The discourse furthermore uses these terms to target any of a number of different virtues or merits in or of those objects.29 Some measure of ambiguity, therefore, must attend upon Brettschneider’s invocations of “legitimacy” in support of his prescriptions for state engagement

24 See id. at 9 (addressing the “neutralist” doctrine that “the state should not promote . . . any particular set of values”); id. at 73 (same).
25 Id. at 123.
26 See authorities cited in note 12, supra.
27 See Richard H. Fallon, Jr., Legitimacy and the Constitution, 118 Harv. L. Rev. 1787, 1802 (2005). For a thorough and meticulous survey and analysis of these usages and their complex (often subtle) interconnections, including references to a wide sweep of the literature, see generally id.
29 See Fallon, supra note 27, at 1790-91 (differentiating among “legal,” “sociological,” and “moral” legitimacy); id. at 1792 (differentiating between “ideal” and “minimal” moral legitimacy).
in democratic persuasion, until we have narrowed down to one each (a) the class of targeted objects and (b) the desired virtue or merit in objects in that class, to which Brettschneider means those invocations to refer. To this end, we rely as closely as possible on Brettschneider’s texts, construing them in the light of the apparent needs of his argument.

1. The Object Class: Legitimacy of the State Political Order

We can deal quickly with object classes. Words and purposes conspire to show that Brettschneider’s concern is with attributions of merit to entire, constitution-bound, state-centered practices of social ordering maintained, in part, by coercive laws. His text consistently refers, in terms, to the question of the legitimacy of “the state” (or, equivalently, of the “government,” “society,” or “democracy” in question).\(^{30}\) His claims characteristically turn on an idea of the state’s special calling to uphold the terms and conditions on which its legitimacy is said to depend.\(^{31}\) Such usages are standard within the common discourse.\(^{32}\) And of course Brettschneider’s argument to reluctant liberals becomes orders of magnitude more urgent when taken to mean that the meritoriousness (in some sense) of an entire constitutional order (not just of one or another specific institutional act or arrangement within that order) depends on a state’s due engagement in democratic persuasion.

2. The “Moral” (Not Merely “Sociological”) Legitimacy of the State Regime

We come, then, to the question of the particular meritorious character—the particular targeted virtue of a political regime—at which a legitimacy judgment (in

\(^{30}\) See authorities cited in note 12, supra. Brettschneider does twice speak of the Supreme Court treating “certain laws” as legitimate (or not), see BrettSchneider, supra note 1, at 82, or striking down certain “laws” as illegitimate, see id. at 149. These applications of legitimacy judgments to individual laws are best understood as parasitic on their application to state level regimes or constitutions: an “illegitimate” law being simply a law that fails to conform either to a legitimate constitution or to principles that any legitimate constitution supposedly would contain. (Thus, in striking down certain laws as illegitimate, the Supreme Court does so “because the reasons and beliefs for such laws violate public principles that are central to the state’s own legitimacy.” Id. at 149.).

\(^{31}\) See authorities cited in note 12, supra.

\(^{32}\) See Fallon, supra note 27, at 1796 (“The leading theories of moral and political legitimacy have primarily addressed the legitimacy of constitutions or governmental regimes . . . .”).
Brettschneider’s usage) is aimed. A sufficiently complete response will take us through several steps. The first is that Brettschneider here is not just talking about legitimacy in what has been called a “sociological” sense of that term, referring to current facts of acceptance by the populace of the regime’s claim to merited political authority. Rather, he means legitimacy in its normative or regulative sense of a standard of “rightful rule”—a measure of the regime’s moral “worthiness to be recognized”—that we as external evaluators bring to the table.

As between a descriptive–sociological and a regulative–moral construction of legitimacy judgments, only the latter will fit comfortably with both our author’s words and his argumentative purposes. Brettschneider carefully differentiates “the crucial value of legitimacy” from the value of (mere) “stability.” “Legitimacy,” in his pages, comes regularly coupled with “justification.” In order for the state “to be legitimate and for the laws to be justifiable to all,” he writes, the state must be committed to the equal status of citizens. “Central to the legitimacy of the democratic state” is the idea that coercion of citizens by law requires sufficient “justifying reasons.”

The apparent point of these remarks is to supply reluctant liberals with a super-compelling reason to give ground on what they have been accustomed to regard as a main principle of democratic political rectitude: strict state neutrality on the field of political debate. Avoidance of a loss to the state’s moral-sense legitimacy undoubtedly supplies a reason of that kind. Avoidance of a substantial risk of loss of the state’s sociological legitimacy might also, of course, supply such a reason, but only if we are shown credible grounds for belief that such a risk is really pending. I believe few readers would think it plausible to

33 See id. at 1790, 1795 (defining a “sociological” usage for “legitimacy”).
34 Cory Brettschneider, Democratic Rights: The Substance of Self-Government 18 n.23 (2007).
35 Fallon, supra note 27, at 1796 n.25 (quoting Jürgen Habermas, Communication and the Evolution of Society 178 (Thomas McCarthy trans., Beacon Press 1979) (1976)).
36 See id. at 1797-98 (noting several kinds of evaluative theories that various evaluators employ).
37 See Brettschneider, supra note 1, at 107 (“If we focus solely on the value of stability and ignore the crucial value of legitimacy, we would have no way of distinguishing between a stable Hobbesian leviathan and a stable rights-protecting liberal society.”); id. at 30 & 180 n.4 (endorsing John Rawls’s distinction between a “modus vivendi” and “stability for the right reasons” and equating the latter with a regime’s legitimacy).
38 Id. at 14.
39 Id. at 49.
40 See infra Part I.A.5.
maintain, and Brettschneider at no point suggests any reason to believe, that the American state’s sociological legitimacy has been placed at risk by the state’s obvious failure, to date, to commit to a policy of engagement in democratic persuasion, up to, or anywhere near the level prescribed in State Speaks.\footnote{In further support of a regulative construction of Brettschneider’s use of “legitimacy” in BRETTSCHNEIDER, supra note 1, we may take note of his use of “legitimacy” in a prior book to mean the standard of “rightful rule” or of “the justifications of coercion” or of “the state’s use of force.” BRETTSCHNEIDER, supra note 34, at 8, 11, 18 n.23, 24, 54, 59.}

3. Legitimacy as the Test of Our Moral License for Support of the State Regime

Brettschneider’s “legitimacy,” we have so far rather comfortably discerned, calls for a regulative, not a merely descriptive, judgment aimed at entire state-level regimes of political rule. And so we come to a third point, which also, like those first two, fits easily into the common discourse: to wit, the legitimacy judgment is specifically concerned with the presence or absence of moral justification for the coercive aspects of legal ordering by and under the regime in question. In a prior book, Brettschneider wrote that a certain, normative conception of democracy provides “the best way to legitimize the state’s use of force.”\footnote{Id. at 11.} In State Speaks, he writes that “central to the legitimacy of the democratic state” is the idea that “[c]itizens should be coerced” only in accordance with certain kinds of reasons.\footnote{BRETTSCHNEIDER, supra note 1, at 49.}

Thus, Brettschneider endorses a standard line of liberal thought, which runs, in short, as follows:\footnote{See Michelman, supra note 28, at 345-47 (recapitulating this standard line of liberal thought); see also Fallon, supra note 27, at 1798 (describing in similar terms a class of “minimal” theories of moral legitimacy as applied to political regimes).} Effective legal ordering is a very great political value. In pursuit of this value, the state must be able to exert credible and effective demands on everyone for a general regularity of compliance with its duly issued laws, like or agree with them or not. Such a demand requires a degree of further justification before supposedly free and equal citizens.\footnote{See BRETTSCHNEIDER, supra note 1, at 54 (calling it “bedrock” in liberal political theory that “state action must be justified to individuals who are recognized as free and equal citizens”).}

To judge a state regime legitimate is to say that it contains a sufficient set of justice-serving or human-serving structures, assurances, and commitments to give everyone prevailing reasons for compliance—and thus to make morally supportable our own
collaboration with the regime’s demand for compliance—despite everyone’s awareness of unresolved uncertainties and disagreements about the justice or other merits of this, that, or the other of the state’s laws and legislative policies. “Legitimate” connotes compliance with a “threshold” standard, “above which legal regimes are [deemed] sufficiently just to deserve the support of those who are subject to them in the absence of better, realistically attainable alternatives.”

4. Legitimacy as a Categorical Judgment

Judgments regarding legitimacy are “categorical” if they allow only “yes” and “no” as answers. They are “scalar” if they allow for adverbially inflected answers such as “highly,” “weakly,” and so on. Judgments regarding “sociological” legitimacy are scalar. By contrast, a judgment regarding the moral-sense legitimacy of a state must be categorical, because it is simultaneously a judgment regarding our moral permission to call on others to go along with a regime that cannot credibly claim to be fully just in the eyes of many (if any) of its constituents. Permission cannot be more or less; either you have it or you don’t. The judgment regarding permission may be close and doubtful or plain and certain, but it has to be made, yes or no, one way or other. “Out” or “safe,” “goal” or “no goal,” there is nothing in between. Collaboration with the regime is either permitted—it is “adequately morally justified”—or it is not.

Is Brettschneider on board? Does he really mean that a state’s failure of active democratic engagement, when the occasion calls, results in a cancellation of our moral permission to support the existing state order’s demands for compliance in general with its laws? The question merits some discussion.

Consider Brettschneider’s account of the state’s obligation of respect for individual rights of expression, association, and conscience. In the name and service of legitimacy, Brettschneider calls for a strict constitutional rule against direct, viewpoint-discriminatory restrictions of speech by the state. His claim, however, is specifically not that state legitimacy is categorically dependent on the adoption of such a rule. Rather, he says that an otherwise well-formed regime can improve its legitimacy-score by

---

46 See Fallon, supra note 27, at 1798.
47 See id. at 1796 (commenting that legitimacy in the “sociological” sense is "a variable, not a constant").
48 Id. at 1799.
49 See, e.g., BREITTSCHNEIDER, supra note 1, at 12, 22, 73.
incorporating such a rule. No doubt the rule may cost something in reduced protection against the spread of hateful views, but, still, allowing in this way for recognition of individual rights can, in the right conditions, yield an “increase” or “gain” to democratic legitimacy “overall.”

Is Brettschneider then treating judgments of a state’s moral-sense legitimacy as scalar? I do not think so. We ought not to read him as meaning, nonsensically, that we could be “a little bit permitted” to collaborate as opposed (say) to “very permitted.” We should rather read him as saying, with perfect good sense, that restricting the state to combating hateful views by means that also respect individual freedom rights makes it likelier than otherwise that we will conclude a complex weighing of pros and cons by finding—categorically—in favor of permission.

5. Interim Summation of Brettschneider’s “Legitimacy”

According to our findings so far, Brettschneider’s “legitimacy” means an external moral standard, applicable to an entire state regime or to its constitution, by which we judge the following: first, the regime’s moral entitlement (or lack of it) to claim from its citizens or subjects a general disposition to comply with its duly enacted laws, just because they are its laws; and, second, our own moral permission (or lack of it) to give our continuing support to that regime. It should be clear how neat is the fit between this typical-liberal notion of legitimacy and Brettschneider’s argumentative needs. When our author invokes legitimacy in support of his prescriptions regarding the state’s energetic engagement in democratic persuasion, in defense of the ideal of free and equal citizenship, he is suggesting to reluctant liberals that a rejection of those prescriptions sets us well on the way toward forfeiture of our moral license to support the extant rule-of-law regime in our country. That surely would give anyone a super-compelling sort of reason to reconsider whatever resistance to those prescriptions he or she might otherwise be feeling. But of course the reason will hold for you or for me only insofar as we find persuasive the proposition that the state’s due engagement in democratic persuasion, in defense of a political ideal of free and equal citizenship, really is a requirement of state legitimacy. We now turn our attention toward that question of persuasiveness.

50 See id. at 16, 80-81, 105.
B. Toward a Problem for Brettschneider

1. Legitimacy: “Strong” or “Weak”?

We asked rhetorically, above, whether Brettschneider “really means” that a certain kind of failure of our state and its officials—to engage sufficiently in a vocal, pedagogical defense of a certain ideal of free and equal citizenship—can result in a cancellation of our moral permission to support the existing state order’s demands for compliance in general with its laws. The answer, so far, is that indeed Brettschneider does mean that. The question then becomes, for each of us, whether we believe it. Is that, to us, a plausible proposition?

Notice, now, that the answer may depend on how strong and demanding—or, oppositely, how weak and forgiving—is our conception of a standard for the legitimacy of a state regime. The lower the level of demandingness at which we set the standard, the less likely are we to conclude that any given feature of the state or its operations is a make-or-break requirement for its fulfillment. If (say) all that is required for the legitimacy of a state regime is its credible constitutional commitment to periodic elections of officials and respect for such libertarian fundamentals as freedom from arbitrary arrest and freedom from censorship of political expression—a “weak” standard—then Brettschneider’s claim that legitimacy depends on a state’s active engagement in a vocal defense of free and equal citizenship will ring hollow. But if, to the contrary, the standard of legitimacy incorporates every true element of political and social justice as you or I might understand them—a “strong” standard—the claim could have a good deal more traction. It follows, interestingly, that to the extent we might find Brettschneider himself endorsing the idea of a weak legitimacy standard, we will also find him, to that same extent, impeaching the credibility of his claim that legitimacy stands or falls with the adequacy of the state’s engagement in democratic persuasion in support of free and equal citizenship. And that extent, as we are about to find, is considerable.

2. The “Substance-Based Limit”: Defense Only of a Political Ideal, Thinly Defined

Brettschneider imposes what he calls a “substance-based limit” on the state’s moral calling, and corresponding constitutional license, to engage in democratic persuasion. The calling and license only cover responses to clear, unmistakable
attacks on a strictly political (as opposed to “comprehensive”) ideal of equal citizenship. For example, expressions of views opposing egalitarian economic policies, or denying that the races, or the sexes, or the sexual orientations, have been created equal, or that they are equally fit for social companionship, or that secular liberals have truer ideas than devout Christians about the path to the good life, would all easily fall on the “comprehensive” side of the line and would fail to activate the state’s limited license to respond. Such expressions are to be distinguished from attacks on the strictly political proposition of civic equality, equality before the state and its law. This means an equality of all persons in matters of rights, of citizenship, of public official and legal standing; it is an equality, therefore, of attribution to persons of entitlements to civic respect and, correspondingly, of their endowments with the “moral powers” that engage and make meaningful the enjoyment by persons of all those forms of public and legal recognition.

“Hateful” speech, in Brettschneider’s vocabulary, includes only speech that more-or-less flagrantly transgresses the thinly defined, political ideal of civic equality. Speech that is otherwise more comprehensively inegalitarian does not activate the state’s license to respond. Doubtless, there are hard cases. The line will not always be clear or easily drawn. That fact does not obscure Brettschneider’s ambition to restrict his claim of an active state duty to the protection of an ideal that is “public” as opposed to “comprehensive,” “political” as opposed to “metaphysical”—thin (as we may say) as opposed to thick. Why so? The Rawlsian vocabulary (not to mention the express appeals to the authority of John Rawls) point straight to the answer. It seems that Brettschneider, like Rawls and like “political” liberals more widely, is in search of basic terms of social cooperation that leave the maximum possible latitude for the moral autonomy of citizens, and so could be found acceptable by citizens holding widely differing, conflicting, even irreconcilable views about some of the deepest questions humans can face about the right and the good. From there, it would seem

51 See id. at 14, 47, 89-90.
52 See id. at 14, 18, 36.
53 See id. at 8, 31, 34-35, 88.
54 See id. at 47, 90.
55 Id. at 14, 30-31.
56 See id. at 30, 34, 35 & 178 n.10, 52 & 180 n.4, 53 & 180 n.5.
57 See, e.g., RAWLS, supra note 28, at xx (inquiring how “free and equal citizens profoundly divided by reasonable though incompatible religious, philosophical,
to follow that Brettschneider will join political liberals, as well, in their advocacy of a relatively weakened standard of state legitimacy. The next two sections will briefly explain why.

3. Legitimacy as a “Weakened” Standard

What, after all, are we doing with this normative-regulative notion of a state regime’s “legitimacy”? Why don’t we simply speak in terms of the regime’s compliance (or not) with the principles of justice for a state regime, and have done with it? A key to the answer, surely, is our knowledge that, in modern free societies, citizens will divide, gravely and intractably, not only over their conceptions of the right and the good, but over their conceptions of what justice truly requires of a state regime in exactly those circumstances: regarding, say, affirmative action, campaign-finance controls, socioeconomic rights, gay marriage, assisted suicide, vouchers for religiously affiliated schools, and on we go. Fifty years hence the list will be different but no shorter. The point of introducing a regulative concept of political legitimacy, to stand beside our various, competing conceptions of justice, is to open a path to morally justified collaboration, by citizens and officials, in demands for a prevailing regularity of compliance by everyone with constitution-conforming laws and policies, regardless of expected, persisting conflict regarding the compliance with justice of this policy or that one, or indeed of this or that feature in the constitution itself. Obviously, in order to do that work, the standard of legitimacy will have to be weaker, less demanding, more forgiving, than will be many, if not all, of the competing conceptions of full-fledged justice that various citizens and parties may support.

It is important, for our purposes, to note that the liberal-theoretic fallback from full and perfect justice to a standard of legitimacy is not simply a practical-minded compromise with human frailty. It is furthermore and distinctly, in the minds of many, a morally principled response to perceived facts of reasonable disagreement or “pluralism” in modern liberal

and moral doctrines” can find a way to “live together and all affirm the political conception of a constitutional regime”).

societies. Legitimacy, then, is the notion that aims to allow for a reasonable convergence by everyone on the regime’s minimum-baseline moral worthiness of support,\(^59\) while at the same time conveying due respect and regard for all parties to intractable disagreements about the justice and other merits of the regime and its various legislative stances.\(^60\) Brettschneider’s sympathy with views of this kind is plain on the face of his text.\(^61\) Such views, as we see, inevitably point toward a relatively lax and forgiving standard of legitimacy.

4. Legitimacy and Transparency

There remains yet a further word to say on this point, and one that goes to deepen somewhat our doubt of the plausibility of the claim that a state’s energetic engagement in democratic persuasion could be make-or-break for a standard of legitimacy that is sufficiently relaxed to do the work we want it for. In the political-liberal view with which Professor Brettschneider associates himself, a standard of legitimacy represents a kind of lower common denominator of core liberal principles of political right and wrong. The standard cannot be excessively demanding or detailed and still do the work we want it for—which is to provide a publically viable standard of justification for collaboration in the country’s practices of coercion by law, in the face of severe and protracted political disagreement, including disagreement about what justice ideally does or does not require. Nor, if it is to do that work, can the standard of legitimacy contain requirements that are excessively prone to reasonable interpretive disagreement at the point of application. This concern for transparency-in-application of the standard of legitimacy goes far, for example, to explain the hesitation of liberals to incorporate into their standard of legitimacy a positive duty of the state to exert itself toward fulfillment of acknowledged demands of economic-distributive justice, whether cast in terms of basic-needs satisfaction or of materially fair equality of opportunity.\(^62\)

\(^{59}\) See Fallon, supra note 27, at 1798-99.

\(^{60}\) See RAWLS, supra note 28, at 229 (arguing on those grounds that the standard of legitimacy must mainly be limited to the ”central ranges” of certain basic liberties on which all reasonable citizens can be expected to agree).

\(^{61}\) See BREITTSCHNEIDER, supra note 1, at 63, 81, 105.

\(^{62}\) See SANDRA LIEBENBERG, SOCIO-ECONOMIC RIGHTS 63-76 (2010) (describing reasons for resistance and suggesting responses). A leading exemplar is John Rawls, who both asserts that a principle of materially fair equality of opportunity is a strict requirement of justice for a regime and declines to treat such a principle as an “essential” component of a
A hesitation of that kind would have obvious application to the question of making a state’s commitment to a practice of Brettschneider-style democratic persuasion into a major test for state legitimacy. An attentive reading of State Speaks shows how complex, delicate, and potentially divisive would be a judgment about whether the Obama Administration, say, has been applying the right amount of anti-hateful persuasive torque to the elbows of civil society: enough but not too much, in response to true but not false occasions of need (as defined by Brettschneider’s “substance-based limit”), and just up to but not across the line that separates persuasion from coercion (as defined by Brettschneider’s “means-based limit”). It seems to me unlikely in the extreme that Brettschneider could mean that every American who sincerely answers “no” to that question has thereby thrown into grave danger his or her moral license for continued general loyalty to the regime of American law.

II. TOWARD A BETTER READING

But if Brettschneider does not mean that, then what does he mean by his invocations of legitimacy as a compelling reason for reluctant liberals to give ground? Accept with me that he does, indeed, use “legitimacy” to mean a standard (i) to be applied categorically (ii) to entire political regimes as (iii) a relatively relaxed, regulative test for the moral supportability of a regime that no one could reasonably, in the face of disagreement, presume to certify as free of serious defect from the standpoint of justice. And so he also means (if you accept my reading) that a regime that we would honestly judge to fail such a test is one that we would lack a moral permission to support. What Brettschneider nevertheless does not mean (or so I want now to suggest) is the implausible claim that a state or government risks forfeiture of its claim to our support by reason of its rejection, by word or deed, of a commitment to democratic persuasion à la Brettschneider.

---


63 See supra Part I.B.2.

64 See Brettschneider, supra note 1, at 87-88 (defining a “means-based limit” on proper democratic persuasion so as to rule out the use of threats and sanctions against exercise of “fundamental rights” of freedom of expression, association, or conscience); id. at 116 (proposing to distinguish “persuasion” by “refusal of a subsidy” from impermissible sanction or threat).
Brettschneider does not ever actually put his claim in those terms, as far as I can see. What he does plainly say is that a regime that allowed its signature operations to stray too far, for too long, from the committed pursuit of the conditions of free and equal citizenship would thereby forfeit its claim to legitimacy. It is that moral fact (supposing you accept it as such)—and not a claimed moral fact that a state’s failure of due engagement in democratic persuasion amounts, just in itself, to a forfeiture of legitimacy—that Brettschneider believes should strongly motivate a demand from us that our governments live up to a policy of due engagement. The point merits recapitulation: Brettschneider is best understood to argue not that a state’s failure of a due-engagement test is, just in itself, a fact that strips it of legitimacy, but rather that a regime’s more-or-less egregious failure in the pursuit of free and equal citizenship, thinly defined, would render it non-legitimate. It is the latter sort of moral fact that is supposed to give us a compelling reason to want our governments to be energetically engaged in democratic persuasion.

A. A Consequentialist Version

And what, then, would be that compelling reason? How would it run? Most obviously and directly, it would be consequentialist in form: a regime that fails of a sufficiently robust commitment to the state’s engagement in democratic persuasion thereby courts an excessive risk of attrition of both its own and the society’s commitment to free and equal citizenship, perhaps eventually to a point where the regime could no longer claim to pass a categorical test of minimal moral legitimacy. That would be a cogently formed, consequentialist argument proceeding from a moral-categorical conception of legitimacy. Brettschneider quite explicitly makes this argument. “If . . . hateful doctrines were left to prevail,” he writes, “they could subvert the basic principles of a legitimate democratic state.” The argument does not lack for force. It could well make good headway with a sizeable fraction of Brettschneider’s readership.

It may not, however, make much of a dent on the resistance of those principled, reluctant liberals who enter the debate braced by strong intuitions that no form of nudge from the

---

65 Cf., e.g., supra note 12 and accompanying text.

66 BRETT SCHNEIDER, supra note 1, at 75. Brettschneider devotes pages to explaining the mechanisms by which this might occur. See id. at 38-42 (drawing arguments to this effect from socio-political considerations of “congruence,” “stability,” “interconnection,” and “public trust”).
state can be construed as non-coercive, so that strict state neutrality becomes in itself (so to speak) a formal requirement for legitimacy.\footnote{See supra notes 23-24 and accompanying text.} It seems they would be likely to respond to the consequentialist argument just as many do when, for example, cherished rights of privacy come up against quite plausible assertions of the needs of national security. A legitimate state, they will say, is and only can be one that treats itself as very strongly bound to search out and prefer other ways to meet the risk, even if at some non-negligible cost in reduced expected effectiveness.

It seems, therefore, worth asking whether Brettschneider has anything more or different to say to our group of hard-core principled reluctant liberals. I believe that he has.

\textbf{B. The “Full Autonomy” of Reasonable Citizens}

Brettschneider writes, he says, in response to liberal theory’s “overemphasis on issues related to the justification of coercion.”\footnote{BRETTSCHNEIDER, supra note 1, at 13.} “The ideal of free and equal citizenship,” he remarks by way of partial explanation, “is not just relevant to justifying coercion; it is relevant to our own moral identities, to the way we order our various public and private commitments.”\footnote{Id. at 58.} It is not, please note, the state’s or the political collective’s identity, but our own (several) personal identities that are being thus bound up with a political ideal of free and equal citizenship. What does Brettschneider mean by this? Perhaps John Rawls, whom we have already seen serving as a guide to Brettschneider in some other respects, could tell us.\footnote{I have drawn the following condensation of John Rawls’s idea of a person’s full autonomy from Frank I. Michelman, \textit{The Subject of Liberalism}, 46 STAN. L. REV. 1807, 1829 (1994).}

Suppose we envisage citizens as possessed of “higher-order interests” in the development and exercise of two so-called “moral powers,” in virtue of which they are owed respect as free and equal individuals.\footnote{RAWLS, supra note 28, at 74, 81.} The powers comprise both a capacity for embrace of a public sense of justice and a capacity for the pursuit of a self-determined conception of the good.\footnote{See RAWLS, supra note 28, at 18-20, 74-75, 81-82. Brettschneider is on board. See BRETTSCHNEIDER, supra note 1, at 34-35.} On that understanding of human personal capability and interest, a convergence of a society’s members on a thin set of principles for the conduct of politics among citizens thus constituted could

\footnote{RAWLS, supra note 28, at 74, 81.}
be a very good great for every participant. Obviously, the conduct of politics in one’s society is bound to bear heavily on the values and satisfactions in a person’s life, and the existence of the thin consensus could put within everyone’s reach the fulfillment of a necessary condition of a satisfying life for persons endowed with higher-order interests in the exercise and development of both the moral powers. The consensus makes achievable by everyone the condition that Rawls calls “full autonomy,” in which a person realizes both the moral powers synchronously—grasping and acting upon a public conception of political decency, by and through the same acts of judgment by which she holds to her own self-responsibly determined conception of the goods to be pursued in life. The consensus thus further enables satisfaction of what Rawls calls the “conception-dependent desire” to realize in one’s person an ideal conception of liberal citizenship.

We have seen how Brettschneider adopts the Rawlsian idea of a thin or “political” conception of justice, putting that distinction to work in explaining what he calls a “substance-based limit” on state engagement in persuasion. Remember: it is not every “inegalitarian” view that the state has any proper business opposing, but only those that are openly and directly hostile to a thin, “political” ideal of equal citizenship. Why so? “Why not “abandon the limitation of promoting only thin values,” and instead endorse” as a public value “a full-fledged conception of equality in all aspects of life?” The answer lies in due respect for the moral and ethical autonomy of citizens. “For citizens,” writes Brettschneider,

to think freely about whether they endorse the ideal of equal citizenship, it is important that they be able to reflect about what they take to be a good life. Part of equal respect entails a respect for citizens to reflect on matters of the good and to make up their own minds freely . . . . The value of equal respect is [thus] central to the idea that we should limit ourselves to a [thin, political] conception of equal citizenship . . . .

Read in the light of the Rawlsian lesson on full autonomy, that would explain the connection, alleged by Brettschneider,
between political ideals such as free and equal citizenship and “our own moral identities, . . . the way in which we order our various public and private commitments.” Establishment of that connection also might make a good start at explaining to reluctant liberals, in non-consequentialist terms, how liberal principles might call for opposition, in a public voice, to corrosive, direct assaults on a thin political ideal of free and equal citizenship. Reduced to the briefest possible summation, the argument would be that the assaulted ideal is a crucial component of a minimal conception of political decency, on the stalwart public affirmation of which the society’s members depend for the possibility of the realization in their lives of a highly valued state of full autonomy. That is, I believe, the deepest argument planted by Corey Brettschneider in the pages of State Speaks.

---

80 Id. at 58.
Freedom of Expression and the Golden Mean

Steven G. Calabresi†

Corey Brettschneider’s splendid new book, *When the State Speaks, What Should it Say?*,\(^1\) is a refreshing and magnificent reinterpretation of the application of First Amendment principles to speech by the government and to hate speech more generally. Professor Brettschneider’s book addresses an extremely difficult and important problem: How should a liberal society approach the topic of hate speech? Professor Brettschneider posits two dystopias that we need to avoid.\(^2\) The first is the dystopia of the Invasive State, which is so eager to militantly protect democracy that it regularly invades people’s rights.\(^3\) The second is the dystopia of the Hateful Society, which is so tolerant that it will not even intervene to defend its core norm of tolerance.\(^4\) Professor Brettschneider describes the harms inherent to each before proposing a new solution designed to occupy the ideological middle ground that protects both expression and the rights of citizens to be free and equal members of society.

Both of the dystopias Professor Brettschneider describes have existed in major constitutional democracies during the last century. The United States, for example, was an Invasive State during the red scares of World War I and during the Senator Joe McCarthy period, which followed World War II and continued in some form throughout the 1950s. The U.S. Supreme Court ratified the dystopia of the Invasive State in cases like the

---

† Clayton J. and Henry R. Barber Professor of Law, Northwestern University; Visiting Professor of Political Science Brown University; Visiting Professor of Law, Yale University. I would like to thank my Dean, at Northwestern, and John Tomasi and the Political Theory Project at Brown University for creating the working environment that made it possible for me to get to know Corey Brettschneider and to write this essay.

\(^1\) COREY BRETTSCHNEIDER, *WHEN THE STATE SPEAKS, WHAT SHOULD IT SAY? HOW DEMOCRACIES CAN PROTECT EXPRESSION AND PROMOTE EQUALITY* (2012).

\(^2\) *Id.* at 10.

\(^3\) *Id.*

\(^4\) *Id.* at 10-11.
unanimous 1919 opinion in *Debs v. United States*,\(^5\) which upheld a 10-year jail sentence and lifetime disenfranchisement for Socialist Party leader Eugene V. Debs for making a seditious speech critical of U.S. involvement in World War I. The government’s specific complaint in the case was that Debs, in a speech, had sought to stir up public non-compliance with the military draft in violation of the Espionage Act of 1917. Debs was convicted and imprisoned, though his speech merely expressed a viewpoint and arguably was not intended to stir up imminent lawless action. While in prison, Debs ran for President of the United States as the Socialist Party candidate and received 919,799 votes (3.4% of the total popular vote), thus confirming the fact that his views had some degree of popular support.

Another landmark U.S. Supreme Court case ratifying the dystopia of the Invasive State is *Dennis v. United States*.\(^6\) In 1951, the Supreme Court upheld the conviction of Eugene Dennis, the General Secretary of the Communist Party USA, in a six to two decision.\(^7\) Dennis’s crime was that he had argued, in the abstract, for the overthrow of the United States government by force and violence, even though he had not incited any specific imminent unlawful action. A plurality of the Court upheld the conviction under the clear and present danger test. This ruling was overturned, *de facto*, by the Supreme Court in its 1969 decision in *Brandenburg v. Ohio*.\(^8\)

A final famous example of the Supreme Court endorsing the Invasive State is the Court’s 1959 decision in *Barenblatt v. United States*.\(^9\) In that case, the Supreme Court held that the House Un-American Activities Committee had the power to hold Lloyd Barenblatt, a college professor, in criminal contempt for refusing to answer questions about his membership in the Communist Party and for refusing to identify other Communists who happened to be college professors. Once again, Barenblatt, like Debs and Dennis, was incarcerated for expressing views on political matters while not attempting to incite imminent lawless action.

All three of these Supreme Court majority opinions—*Debs v. United States*, *Dennis v. United States*, and *Barenblatt v. United States*—are overwhelmingly viewed today as having been wrongly decided, and they form a paradigmatic instance

\(^5\) 249 U.S. 211 (1919).
\(^6\) 341 U.S. 494 (1951).
\(^7\) *Id.* at 495.
of the Invasive State, which Professor Brettschneider warns us against. The Supreme Court’s hyper-strong First Amendment case law today\textsuperscript{10} is, in effect, a response to the dystopia of the Invasive State.

The dystopia of the Hateful Society is epitomized by the German Weimar Republic, which held power in Germany from the end of World War I until the rise to power, in 1933, of Adolf Hitler and the Nazis.\textsuperscript{11} The Weimar Republic was rife with anti-Semitism and was characterized by the existence of powerful political parties and groups that were openly opposed to democracy, constitutionalism, and tolerance.\textsuperscript{12} The officer corps of the German army during this period was disdainful of democracy and yearned for a return of the autocratic and imperial German monarchy. The army refused to protect the government from coup attempts, and many officers blamed German democrats for the army’s loss in World War I.\textsuperscript{13} In addition, there were politically powerful parties in the German parliament that were committed to the overthrow of the democratic regime.\textsuperscript{14} The German Communist Party and Hitler’s Nazi Party loathed the Weimar democracy and would sometimes form so-called red-brown coalitions to vote down democratic measures or bring down the government in parliament.\textsuperscript{15}

Freedom of speech during the Weimar era led to a flourishing of the arts, but it also led to a flourishing of hate speech in an increasingly poisonous political climate. Professor Cindy Skach describes the climate in Weimar, Germany, well:

\begin{quote}
[T]here was a rather exuberant and creative society . . . . ‘When we think of Weimar, we think of modernity in art, literature, and thought; we think of the rebellions of sons against fathers, Dadaists against art, Berliners against beefy philistinism, libertines against old-fashioned moralists; we think of The Threepenny Opera, The Cabinet of Dr. Caligari, The Magic Mountain, the Bauhaus, Marlene Dietrich. And we think, above all, of the exiles who exported Weimar culture all over the world. Indeed, Weimar Germany is still often remembered for
\end{quote}

\begin{footnotes}
\item[10] See, e.g., Snyder v. Phelps, 131 S. Ct. 1207 (2011) (striking down, in an 8 to 1 vote, a tort judgment for intentional infliction of emotional distress as violating the First Amendment because the speech in question was about a public issue on a public sidewalk).
\item[13] See The BREAKDOWN OF DEMOCRATIC REGIMES, supra note 11.
\item[14] See id.
\item[15] See id.
\end{footnotes}
producing some of the most brilliant scholars, artists, and free thinkers in the framework of a golden age . . . .

And yet, not all political parties or all of this flourishing society was enthusiastic about democracy. Several groups remained loyal to the idea of restoring the monarchy, and this included a substantial part of the landed classes, the Lutheran Church, and the nation-building elites that remained a part of the state bureaucracy.¹⁶

In 1925 and 1926, Adolf Hitler published his manifesto Mein Kampf—“Struggle”—which was filled with anti-Semitism and warnings about “the Jewish menace.”¹⁷ The book had sold 240,000 copies by the time Hitler came to power in 1933, and its publication played an important role in the breakdown of the Weimar Republic. Weimar Germany thus typifies what Professor Brettschneider quite accurately calls the dystopia of the Hateful Society.

The extension of political freedoms and First Amendment-type rights to the Nazis and the Communists during the Weimar Republic is widely viewed as a monumental mistake and a catalyst for World War II and the Holocaust. Present day Germany’s doctrine of Militant Democracy that bans totalitarian political parties as well as hate speech is widely seen today as an effective and necessary response to what Professor Brettschneider calls the dystopia of the Hateful Society. Professor David Currie explains that “the concept of militant democracy (‘streitbare’ or ‘wehrhafte Demokratie’) [is seen as being necessary because of] the bitter experience of the Weimar Republic, in which antidemocratic forces took advantage of political freedoms to subvert the constitution itself.”¹⁸ Militant Democracy in Germany, today, is thus a direct response to what Professor Brettschneider calls the dystopia of the Hateful Society in Weimar, Germany before the rise to power of Adolf Hitler and the Nazis.

The core insight of Professor Brettschneider’s book is that we need to find a Golden Mean that lies somewhere between the two dystopias. Aristotle describes many of the virtues he writes about in The Nicomachean and Eudemian Ethics as existing as a Golden Mean between two vices. He thus defines courage as being a Golden Mean between the vices of recklessness and cowardice. Aristotle explains in The Nicomachean Ethics that:

---

¹⁶ SKACH, supra note 12, at 30-31.
[The virtues] are naturally destroyed through deficiency and excess, just as we see in the case of strength and health . . . : excessive as well as deficient gymnastic exercises destroy strength, and, similarly, both drink and food destroy health as they increase or decrease in quantity, whereas the proportionate amounts create, increase, and preserve health. So it is too with moderation, courage, and the other virtues: he who avoids and fears all things and endures nothing becomes a coward, and he who generally fears nothing but advances toward all things becomes reckless. Similarly, he who enjoys every pleasure and abstains from none becomes licentious; but he who avoids every pleasure, as the boorish do, is a sort of “insensible” person. Moderation and courage are indeed destroyed by excess and deficiency, but they are preserved by the mean.\(^{19}\)

Similarly, in *The Eudemian Ethics*, Aristotle describes courage as being the Golden Mean between recklessness and cowardice; modesty as being the Golden Mean between shamelessness and bashfulness; truthfulness as being the Golden Mean between boastfulness and dissimulation; friendliness as being the Golden Mean between flattery and curmudgeonliness; and wisdom as being the Golden Mean between unscrupulousness and unworldliness.\(^{20}\)

Similarly, Professor Brettschneider correctly intuits that the right approach to freedom of expression is one that falls in between the two dystopias of the Invasive State and the Hateful Society. Constitutional democracies must strive to both protect freedom of expression, and actively affirm the free and equal citizenship of all of their citizens. This is the Golden Mean between the Invasive State and the Hateful Society.

I agree with this sentiment entirely. There is a Golden Mean with respect to freedom of expression that we should strive to attain. I agree with Professor Brettschneider that—given our history during the Red Scares and the McCarthy era—the United States should not criminally punish hate speech because there is more of a danger that the government will persecute individuals unjustly than there is a danger that hate groups will take over our complex federal republic with all its checks and balances. There may be societies, like West Germany after World War II, where it is necessary to outlaw hate speech and anti-regime political parties, but the United States is not such a society. I thus emphatically agree with *Brandenburg v. Ohio* (raising the bar on prosecutions for incitement of lawless action);\(^{21}\)


At the same time, however, I also emphatically agree with Professor Brettschneider that we should not expect that, when the government itself speaks, its speech must always be viewpoint neutral. As Professor Brettschneider's book points out, we celebrate great men and women and great events in our history for value-laden reasons and that is precisely as it should be. We should celebrate Martin Luther King's birthday, and not Robert E. Lee's, because it is vital that our government actively affirm the ideal of free and equal citizenship for all its citizens. This is a way of affirming that ideal. Similarly, the president and other high government officials should praise democracy and civil rights, and should not give equal time in their remarks to hateful or opposing viewpoints.

I would add that public colleges, universities, and secondary schools could not even function if they did not choose to praise some viewpoints and criticize others. The praising of some things and the disapproving of others is basically at the core of what education itself is all about. The Corporation for Public Broadcasting, PBS, National Public Radio, and the National Endowment for the Arts are among the many examples of governmental institutions that do not censor their speech to make it viewpoint neutral. Obviously, these entities are all of great value, and they should be retained.

Professor Brettschneider's book is a breath of fresh air when it comes to First Amendment scholarship because he recognizes an obvious truth which has not been adequately recognized until now—that when the government speaks it ought to speak in support of democracy, and free and equal citizenship. At some level, we all probably know that, but Professor Brettschneider brings the point out into the open and explains where it comes from. It is true that the government could in theory use its bully pulpit to attack free and equal citizenship or individual liberty, but I think Professor Brettschneider is right in his intuition that this is not a concern, at least in the U.S. context. I therefore think that Professor Brettschneider comes close to striking the right balance in avoiding the dystopias of the Invasive State or the Hateful Society in the U.S. today.

---

I do, however, disagree with Professor Brettschneider on the ease with which he would revoke 501(c)(3) tax exempt status from groups like the Boy Scouts that he rightly thinks are engaged in speech that denigrates free and equal citizenship. My disagreement with Professor Brettschneider here is more practical than normative, but attaining a Golden Mean between two dystopias often requires attention to practical concerns. I do not think that groups like the Boy Scouts that denigrate free and equal citizenship deserve a 501(c)(3) tax subsidy, but I do fear that once government officials get in the business of evaluating which 501(c)(3)s ought to lose their tax exemption on the basis of political alignment, a very rabid partisanship will set in which will be destructive of the goals of free and equal citizenship that Professor Brettschneider and I favor.

An effort to revoke the tax exempt status of the Boy Scouts, in my opinion, would rapidly lead to a counter effort by social conservatives to revoke the tax exempt status of groups that support, for example, gay rights. It would politicize the Internal Revenue Service, which, as we can see from current events in the government today under President Obama, is not something we should want. IRS officials who know they have more power to revoke 501(c)(3) status might be more likely to audit the returns of political opponents—the Eugene V. Debs of the current day—and to excuse tax code violations by political supporters of the regime in power. A politicization of the IRS would be bad for civil liberties and bad for the government because it would undermine support for paying taxes. This seems to be, in my opinion, prudentially, a dangerous road to go down.

Many political conservatives would leap at the chance to eliminate 501(c)(3) status for liberal universities and foundations if they thought this behavior was tolerable. We should not go down that road. America’s huge not-for-profit corporate sector is one of the many things that makes the United States so much stronger than Europe and Japan. Section 501(c)(3) organizations are what Edmund Burke called “mediating institutions.”

27 For a discussion of the way in which political parties function as mediating institutions, see Steven G. Calabresi, Political Parties as Mediating Institutions, 61 U. CHI. L. REV. 1479 (1994).
people. Mediating institutions between the government and the people include: the family; churches, synagogues and mosques; charitable associations; civic associations such as the League of Women Voters; labor unions, family owned and other small businesses, local government units such as New England town hall meetings, and political parties. They are invaluable assets, and should be left alone.

Professor Brettschneider may at some level recognize that, because he further argues in his book that the Boy Scouts should lose their tax exempt status but that the Catholic Church should not, even though both organizations oppose free and equal citizenship with respect to sexual orientation. I find it hard to discern a principle here given that the Boy Scouts now admit gay scouts whereas the Catholic Church would presumably excommunicate a gay church member who refused to pledge not to engage in same-sex relationships. (Perhaps Pope Francis’ recent statements on not judging gay people indicate a change in what has been the Catholic Church’s position until now.) To be clear, I think it is a violation of the rights of free and equal citizenship for the Boy Scouts to ban gay scoutmasters, and I would vote for a federal law that banned sexual orientation discrimination in employment so long as it exempted religious organizations. I am just not persuaded by the book that there is a significant difference between the position of the Boy Scouts and the position of the Catholic Church, at least prior to Pope Francis.

Professor Brettschneider argues in the book that his position is supported by the Supreme Court’s opinion in Bob Jones University v. United States. In that case, the Supreme Court upheld the authority of the IRS to strip Bob Jones University of its 501(c)(3) tax exempt status because of its racist policy of not allowing interracial dating for what the University claimed were religious reasons. I agree that the federal government was within its rights under current law and that it acted correctly in that case in revoking the 501(c)(3) tax exempt status of Bob Jones University. There has never been a religion in the United States under which interracial dating was forbidden as a matter of widespread religious belief. Bob Jones’ claim to a church’s tax exempt status was therefore

28 EDMUND BURKE, REFLECTIONS ON THE REVOLUTION IN FRANCE 315 (Conor Cruise O’Brien ed., 1968). Burke described mediating institutions as being societies’ small “platoon[s].” Id. at 135.
29 For a fuller discussion of the concept of mediating institutions, see Calabresi, supra note 27, at 1479-1533.
correctly denied because the interracial dating ban could not have been a matter of religious principle. I would not conclude from this, as Professor Brettschneider does, that the Boy Scouts’ tax exemption should be withdrawn because of their refusal to recognize gay rights. I think the Boy Scouts can be and should be publicly criticized for this, but I would ostracize the Scouts rather than withdraw their tax status.

I agree with Professor Brettschneider that free and equal citizenship is the proper goal of the liberal state, but I would go further. I would ban laws or executive actions that deprive people of life, liberty, or property on the basis of religion as well as on the basis of race and gender. Most people are born into their parents’ religion, a reality that may be characterized as an immutable characteristic. Secular and Christian people of Jewish descent were stunned to find that Hitler thought them to be Jewish by blood and accordingly sent them to concentration camps. Muslims today are born into their faith and may face charges of apostasy if they try to convert.

I thus disagree with the Supreme Court’s decision in *Locke v. Davey*, which held that it was constitutional for Washington State to give scholarships for students doing graduate school work, so long as they were not studying theology or religion. This seems to me to be a blatant attempt by the State of Washington to discriminate on the basis of religion. It is objectionable for the same reason it would be objectionable if the State of Washington disallowed funding graduate work at historically African American schools or at an all-women’s school. The government deprives individuals of their free and equal citizenship when it discriminates on the basis of religion. Thus, Justices Scalia and Thomas were right to dissent in *Locke v. Davey*.

Similarly, the so-called Blaine Amendments in state constitutions violate the equal protection guarantee of the Fourteenth Amendment. Blaine Amendments are so named after Senator James G. Blaine, a nineteenth-century politician, who supported state and federal constitutional amendments that forbade any government funding from ever going to any religious school, charity, or organizations, even if similar secular schools, charities, and organizations did receive government funds. Many

---


states passed Blaine Amendments, but the proposed Blaine Amendment to the federal constitution was not adopted.

The equal protection guarantee would also be violated if state schools were run to tolerate a hostile environment with respect to race, sex, or religion. State-run schools cannot be suffused with racism or sexism because the existence and state funding of such schools would violate the free and equal citizenship which is at the core of the Fourteenth Amendment. Similarly, state-run schools cannot be suffused with hostility to religion or militant secularism. The federal and state governments cannot discriminate on the basis of religion, and Title VII of the Civil Rights Act of 1964 forbids private employers from discriminating on the basis of religion as well. It follows that just as Title VII does not allow employers to maintain a workplace that is hostile on account of race or sex, so too they cannot maintain a workplace that is hostile on the basis of religion.

Finally, I would note that the concept of free and equal citizenship that Professor Brettschneider defends does not by definition apply to longtime resident aliens, whether legal or illegal. This would seem to be a flaw. Surely government speech must show respect for the fundamental human rights even of non-citizens. It is for this reason that the Fourteenth Amendment’s due process and equal protection clauses apply to all persons and not merely to all citizens as does the privileges and immunities clause.\(^{33}\) I doubt Professor Brettschneider disagrees with this, but he may in the future want to elaborate on the extent to which his ideas are based on fundamental human rights as well as the rights of citizens.

All in all, however, Professor Brettschneider’s book is a major step forward. He is absolutely right to seek a Golden Mean between the two dystopias of the Invasive State and of the Hateful Society. I would locate that Mean in a slightly different place than Professor Brettschneider would, but it seems that fundamentally we are in agreement.

---

\(^{33}\) The amendment reads:

No state shall make or enforce any law which shall abridge the privileges or immunities of citizens of the United States; Nor shall any state deprive any person of life, liberty, or property, without due process of law; Nor deny to any person within its jurisdiction the equal protection of the laws.

U.S. CONST. amend. XIV, § 1.
Democratic Rhetoric: How Should the State Speak?

Josiah Ober†

INTRODUCTION

I agree with most of the substantive arguments in Corey Brettschneider’s book When the State Speaks, What Should It Say: I think that the middle ground, between a highly activist coercive and invasive state (the Intrusive State) and a fully neutral state that allows hateful speech to go unanswered (the Hateful Society), is the right “solution space” for liberal democracy. I applaud Brettschneider’s emphasis on free and equal citizenship as the core value that a liberal democracy must defend. Indeed, I would go further, and say that any democracy that deserves the name, whether liberal or not, must defend that ground.

My comments will not focus on the theoretical foundations of the book’s argument, but on possible extensions of it. At the core of my response is the thought that a lot more can be done by a state (by its representatives) in the way of expressive articulation of what Brettschneider calls “reasons for rights.”¹ So I will be focusing on an issue that does come up in the book, but does not seem to get its full due: once we have decided the substance of what the state should say (public justifications for free and equal citizenship, and public refutations of those who would deny the relevant sorts of political freedom and equality to citizens) and once we have agreed on the means (persuasion, not coercion) by which the state will say it, how ought the state to speak?

My primary concern is that Brettschneider leaves to one side what I see as the centrally important issue of democratic rhetoric: once we have decided what the state should say, we

† Stanford University. Constantine Mitsotakis Professor in the School of Humanities and Sciences, Professor of Political Science and Classics. Author of books and articles on Greek history, classical political thought and practice, and democratic theory.

¹ COREY BRETTSCHNEIDER, WHEN THE STATE SPEAKS, WHAT SHOULD IT SAY?: HOW DEMOCRACIES CAN PROTECT EXPRESSION AND PROMOTE EQUALITY 4-5, 12-14, 20-21 (2012).
need to ask whether, when the state speaks, anyone will be persuaded. The answer is surely, “only if the state speaks persuasively.” That raises the question of how and why speech persuades—which is, of course, what rhetoric is all about. While contemporary analytic political theory (in stark contrast to classical political theory) tends to avoid the topic of rhetoric, perhaps because of its association with misleading and pernicious speech, studying democratic speech without attending to democratic rhetoric is rather like studying an automobile without attending to its engine—rhetoric is the engine of democratic politics; if we ignore it we will never understand what makes the thing go.  

Brettschneider offers two main answers to the question of how the state ought to speak. First, public officials ought to speak out against attacks on equality (notably against racism, sexism, and gay-bashing) and they ought to honor historical movements that pushed in the direction of greater freedom and equality for citizens (e.g. by national holidays and monuments and school curricula). Second, the state ought to be more selective about its subsidies to private organizations, including religious organizations; organizations that reject or denigrate free and equal citizenship ought to be denied public money and tax breaks.

It is the first of these two categories that most interests me—and that seems to get somewhat short shrift in the book. One of Brettschneider’s important points is that there is a qualitative difference between the duty of citizens to speak out in defense of democratic rights—the point was strongly made by John Stuart Mill and brilliantly exemplified by Martin Luther King, Jr.—and the duty of agents of the state to do so. And yet, the book does not seem to offer a very detailed road map that might guide us in thinking about what that special duty of representatives of the state would mean in democratic practice.

The positive examples that Brettschneider offers of the proper use of non-fiscal democratic persuasion by agents of the state are overwhelmingly drawn from U.S. Supreme Court opinions and jurisprudence. Yet, as Brettschneider rightly and repeatedly points out, Supreme Court opinions are not generally models of persuasive public speech. Almost no one outside of the tribe of political theorists and constitutional lawyers reads

---

them, and it seems unlikely that many citizens who do read them alter their viewpoints on rights as a result. The other examples I found in the book were the establishment of Martin Luther King Day; history standards that mandate teaching about the civil rights, women’s rights, and gay rights movements; Bill Clinton’s apology for the Tuskegee experiments; Michael Bloomberg’s chastisement of anti-Muslim speech related to the Cordoba Center; and Barack Obama’s lukewarm support for Iranian democracy. I will stipulate, without much fear of objection, that none of these belongs on any plausible list of the top ten moments in the history of pro-democratic, rights-justifying political speech.

I. TOP TEN PRO-DEMOCRATIC SPEECHES

What would Brettschneider say about examples of public speech that do have a good claim to belong on a top ten list? Although there are many candidates for inclusion on the top ten list, here (in chronological order, and limited to Athens, England, and the USA) are ten candidates that come to mind:


2. Pericles, 431 B.C.E. Funeral Oration. “Its administration favors the many instead of the few; this is why it is called a democracy.”

3. Demosthenes, 346 B.C.E. Speech 21 Against Meidias. “[T]he laws are strong through you and you through the laws.”

4. Edmund Burke, 1774. Speech to the Electors of Bristol. “Parliament is a deliberative Assembly of one Nation, with one Interest.”

---


5 Demosthenes, Speech 21: Against Meidias § 224 (A. T. Murray trans., Harvard Univ. Press 1939) (c. 346 B.C.E.), available at http://www.perseus.tufts.edu/hopper/text?doc=Perseus%3Atext%3A1999.01.0074%3Aspeech%3D21%3Asection%3D24. Demosthenes was not stricto sensu speaking an agent of the state, but as prosecutor, Demosthenes was serving a public function.
5. Abraham Lincoln, 1863. Gettysburg Address. “[O]f the people, by the people, and for the people.”

6. Dwight Eisenhower, 1957. Address on the events at Little Rock. “Mob rule cannot be allowed to override the decisions of our courts.”


10. Barack Obama, 2013. Speech on race in America: “Trayvon Martin could have been me.”

These speeches are not all focused specifically on free and equal citizenship per se, but each seems to me, one way or another, to invoke rights, and reasons for rights (or in the Greek examples, quasi-rights), and democracy. So my question for Brettschneider is this: which speeches pass the test of appropriate democratic persuasion, which fail the test, and why?

Brettschneider’s core claim for democratic persuasion is that it must articulate the reasons for rights. I believe the speeches I have listed do that, one way or another, but they use very different kinds of rhetoric than analytic political theorists are used to employing in professional scholarship. In brief, we write in a deliberately-measured, not to say arid, prose that is largely stripped of appeal to sensibilities other than reason.

---

itself. The rhetoric of the speeches listed above is, I believe, grounded in appeals to reason, but in each case it goes well beyond the bare appeal to reason, appealing to an array of positive, non-hateful emotions and to an array of civic virtues. They use techniques of narrative, invention, style, and delivery that are familiar from classical rhetoric (as canonized by, for example, the Latin writers Cicero and Quintilian). Brettschneider alludes briefly to legitimate uses of emotional appeals, noting that “the requirement that democratic persuasion include explicit reasons does not mean that it must avoid emotion or rhetorical persuasiveness.”13 But that tip of a hat to rhetoric is bracketed by much more extensive discussion of worries about manipulation, subliminal messages, propaganda, and so on. It appears that Brettschneider is nervous about letting the rhetorical camel’s nose into the tent of public reason—a nervousness that he shares, I dare say, with many members of the political theory tribe.

There is reason enough for that nervousness—after all, rhetoric can be and sometimes is divorced from reason-giving in ways that can devolve into overt manipulation and propaganda. The fear of bad rhetoric begins with the founding works of western political theory—most notably of course with Plato (especially in his dialogues Gorgias, Phaedrus, and Republic), who drew a bright line between philosophy as the realm of truth and reason, and rhetoric as the realm of opinion and deception.

II. A NEW THEORY OF RESPONSIBLE PUBLIC RHETORIC

Once we have set our foot on the road, as Brettschneider urges us, of using persuasion rather than coercion to achieve important public ends—to resist devolution into the Hateful Society without bringing in the heavy artillery of the Intrusive State—we need to take a fuller inventory of the tools in the toolbox of persuasion. No doubt denying subsidies is one powerful tool. Expressing reasons for rights in the austere technical languages of analytic liberal political philosophy and Supreme Court jurisprudence may also be tools of sorts. But surely we cannot afford to overlook the potential tools of rhetorical persuasion. Nor, if we are to take up the theoretical challenge of democratic persuasion, can we afford to simply leave rhetoric to disciplines other than political theory or to political practitioners.

There is nothing intrinsically odd about saying that political theory should take rhetoric seriously, given that

13 BRETT SCHNEIDER, supra note 1, at 88-89.
rhetoric was once a major part of political theory. Plato not only criticized the misuse of rhetoric, he understood it deeply and his dialogues are masterpieces of persuasive rhetoric—how many of us got started in political theory in part because we fell in love with the Socrates of the *Apology*, *Crito*, and *Republic*? Aristotle’s great trilogy on political theory concludes with the *Art of Rhetoric*. And of course Cicero was a political philosopher, theorist, and practitioner of political rhetoric; ditto Machiavelli and Hobbes.

Classical rhetorical theory is, however, inadequate to the task of fully explaining the sorts of democratic persuasion Brettschneider envisions. We need a new theory of responsible public rhetoric that is (1) suited to the sort of democratic persuasion that Brettschneider advocates, (2) capable of accommodating the emotional force of great public speeches that have actually had a meaningful impact on citizens in terms of transforming attitudes, and (3) delivered in a way that motivates citizens to change their behavior so as to more reliably act in defense of free and equal citizenship. If we do not develop the analytic tools to understand how and why and when rhetoric moves people, we will never really understand the toolbox of democratic persuasion—much less make it available for use in the real world in which the state speaks to citizens. A new theory of public rhetoric requires expanding the domain of contemporary analytic political theory. Surely, that is exactly what Brettschneider is urging us to do as we mark out the terrain between the Hateful Society and the Intrusive State and seek to learn how that terrain can best be defended.

What does Brettschneider think about the proposition that accepting the core argument of his book demands a substantial expansion of the field of analytic liberal theory in the direction of a new theory of rhetoric? That expansion does not require that we reframe Brettschneider’s arguments in terms of emotions or virtues, but it does require that we develop a robust theoretical framework capable of determining when appeals to emotions and virtues are, and are not, compatible with offering citizens reasons for rights.

III. POSSIBLE POLICY CHANGES

Brettschneider’s chapters on the use of selective subsidy could readily be employed to draft the legislation necessary to

---

14 *Id.* at 20 (rejecting arguments based on virtues).
implement the policy changes demanded by his normative arguments. By contrast, there seems to be no very obvious legislative agenda implied by his discussion of the use of public speech by representatives of the government for democratic persuasion—other than perhaps tweaking standards for teaching American history. By way of conclusion, let me offer a few sample policy changes and pose the question of whether (assuming, counterfactually, that they could be passed) Brettschneider would regard these as promoting or violating his conception of legitimatedemocratic persuasion.

1. Require the president (and state governors) to give a “State of the Citizenship” address each year (or, if we wanted to mimic the Athenians, every year in which citizen-soldiers died in combat). The purpose of the speech would be to give reasons for rights in light of challenges to and advances in rights at home and abroad in the last year.

2. Require that each U.S. Supreme Court justice issue a public justification (or concur in another justice’s justification), expressed briefly and in non-technical language, for each of his or her votes on cases affecting free and equal citizenship. This would be separate from lengthy and often technical legal opinions.

3. Require that elected officials (or maybe all government officials) (1) take a public oath of personal commitment to, and belief in, the fundamental rights of free and equal citizenship (i.e. they not only swear to uphold the laws guaranteeing rights, but affirm their belief in those rights) and (2) make a public statement of their own reasons for their commitment to those rights (which could appeal variously to political or comprehensive conceptions). Those whose public behavior or votes blatantly violated their oath would be appropriately sanctioned (by, for example, public exposure, fines, and loss of office).

4. Require an oath similar to number three above as part of the naturalization process for new U.S. citizens.

5. Seek to ensure (through mandating coverage on all TV stations, emailing every citizen, or similar means) that important expressions of reasons for rights (e.g. numbers one and two above) are privileged, and thus
have a better chance to cut through the chaotic and noisy barrage of information/entertainment to which all contemporary citizens are constantly exposed.

My guess is that Brettschneider would not want to endorse all (or maybe not any) of these policy changes, but if I am right in that, I would like to hear why not. In my view, each of the sample changes seems, on the face of it, to be in line with the core claims of the book (as, for example, demanding commitment oaths of all citizens would not be). In any event, I think that sketching a range of policy changes, and then determining which ones fit and fail to fit Brettschneider’s theory of democratic persuasion—along with the test case of the “top ten list” above—would help to put some flesh on the skeleton of a theory presented in Brettschneider’s stimulating book.
You’re All Individuals: Brettschneider on Free Speech

Andrew Koppelman†

INTRODUCTION

Brian: No, no. Please, please please listen. I’ve got one or two things to say.

The Crowd: Tell us! Tell us both of them!

Brian: Look, you’ve got it all wrong. You don’t need to follow me. You don’t need to follow anybody! You’ve got to think for yourselves! You’re all individuals!

The Crowd (speaking in impressive unison): Yes! We’re all individuals!

Brian: You’re all different!

The Crowd: Yes! We are all different!

Man in crowd: I’m not . . .

Man in crowd: Shhh!

Brian: You’ve all got to work it out for yourselves.

The Crowd: Yes! We’ve got to work it out for ourselves!

Brian: Exactly!

The Crowd: Tell us more!!

Liberalism, Corey Brettschneider tells us, is uneasily poised between two dystopias, which he calls the Invasive State and the Hateful Society. The Invasive State constantly and intrusively monitors and punishes ordinary citizens for saying and doing anything, even in private, that is antagonistic to the values of free and equal citizenship. The Hateful Society is

† John Paul Stevens Professor of Law and Professor of Political Science, Northwestern University. Thanks to Corey Brettschneider and Martin Redish for comments on an earlier draft, and to Marcia Lehr for research assistance.

1 MONTY PYTHON’S LIFE OF BRIAN (Sony Pictures 1979).
characterized by a discriminatory culture that generates political inequality. “It has seemed that we are faced with a dilemma” between these two choices.

The solution—the solution, Brettschneider argues—is that the liberal state must allow hateful speech, but must also publicly condemn it. “Although citizens should retain rights to disagree with antidiscrimination laws, the state has the obligation to use its expressive capacities to defend the values of free and equal citizenship against criticism from hateful or discriminatory groups and individuals.” This expression by the state evidently is both necessary and sufficient to avoid both dystopias: “In this way the state can protect the right to express all viewpoints and, at the same time, it can defend the values of freedom and equality against discriminatory and racist challenges.”

It isn’t that simple.

I. ONE SIZE FITS ALL

Brettschneider is offering an ideal theory: in a well-functioning society, the state will tell citizens to reject hateful ideas. The citizens will think what the state tells them to think. Although “there should be a wide role for citizens in defending the core liberal democratic values,” the state has little discretion about whether or when to exercise its persuasive capacities.

This ideal is not contingent on how effective government speech is likely to be—something that certainly varies from one society to another. Nor is it contingent on how prevalent such ideas happen to be in any particular society. That varies too. Prejudice against the Irish used to be pervasive in the United States. Now it is nearly nonexistent. But if, anywhere in this vast land, there is a single lonely crank who casts aspersions on the liberty and equality of the Irish, government has a duty to hunt him down and denounce him.

Brettschneider thinks the state’s exercise of its moral authority will “effectively counter the spread of hateful

---

3 Id. at 6.
4 Id. at 3-4.
5 Id. at 13.
6 Id. at 119.
viewpoints,”7 and that this authority should “counter the concerns”8 of those who worry about the stability of democracy. Does he really believe that this alone will guarantee that minority rights will be respected in democratic politics?9 He never explicitly says that, but he endlessly repeats variations on the sentences I just quoted, and he presents this and nothing else as the solution to the dilemma he describes.

Brettschneider’s most innovative proposals call for the state to expressly discriminate against hateful views in the granting of subsidies and nonprofit status, and accordingly to abandon neutrality in limited public fora. If it does not do this, he claims, it could be perceived as endorsing the hateful views that it is protecting.10 But linguistic meaning always depends on the background assumptions shared by speaker and audience. When the Westboro Baptist Church is granted the same tax-exempt status as every other church, does anyone really think that the state is endorsing its hateful antigay and anti-Catholic views?11

Sometimes state pedagogy is helpful to the cause of liberal democracy. But this is a non-ideal state of affairs. In his earlier, much better book, Democratic Rights, Brettschneider defended judicial review, but argued that it is decidedly a second-best solution: respect for individual rights ideally emerges from an unfettered electoral process rather than from being imposed from above.12 That respect also ideally emerges from civil society, rather than being dictated by the state.13 The state is not a reliable source of moral wisdom. Good liberal citizens know that.

Brettschneider’s idea that the government, in particular, has a duty to speak implies that the state is uniquely authoritative.

---

7 Id. at 110.
8 Id. at 17. He acknowledges that in some “emergency situations” his solution may not be adequate. Id. at 18.
9 In South Africa, the government has been a world leader in gay rights, but violence against gay people is pervasive, and police ignore it. “Corrective rape,” the practice of raping gay men and lesbians in an effort to change their sexual orientation, is not only common, but is frequently abetted by the victims’ families. Clare Carter, The Brutality of “Corrective Rape,” N.Y. TIMES, July 27, 2013, http://www.nytimes.com/interactive/2013/07/26/opinion/26corrective-rape.html.
10 BREITTSCHNEIDER, supra note 2 at 16, 84, 113, 138-39.
11 Id. at 133-34.
13 At one point, Brettschneider concedes this: “Ideally, reflective revision might be prompted by the persuasion and reasoning of one’s fellow citizens in civil society.” BREITTSCHNEIDER, supra note 2 at 66. But this ideal sits uneasily beside the incessantly repeated claim that the state has an “obligation to speak. In arguing for a distinctive state role, he also observes that “only a representative of the people as a whole . . . can speak on behalf of the polity in condemning and apologizing for past injustices.” Id. at 41. That is certainly true, but it is a minor part of the state speech he has in mind.
That premise is a familiar part of the legitimizing mythology of the modern state, which imagines itself as a kind of secular God—a uniquely stabilizing and valid source of obligations.\textsuperscript{14} That mythology is a fact of life, and sometimes liberals will need to deploy it rhetorically. But they should never do so without disgust.

Brettschneider is oblivious to these objections. The counterarguments he sees fit to address are often silly. Does anyone really believe that what happens in civil society ought to be immune from criticism?\textsuperscript{15} That the state has an “obligation to guarantee the equal success of all viewpoints”?\textsuperscript{16} That any criticism of specific objectionable tenets of a particular hateful religious group will pressure citizens to abandon religion itself?\textsuperscript{17} He critiques these claims at length, but his citations offer no evidence that they have any proponents.

The question of just how the proper operation of a liberal society—a complex system with a complex range of possible pathologies—is to be promoted is one that presents different specific problems, and so requires different strategies from one time and place to another. Sometimes the Monty Python joke I quoted at the beginning really is the appropriate path. Post-World War II Japan under U.S. General MacArthur looked a lot like that: liberal prescriptions issued by a military dictator.\textsuperscript{18} The problem Brettschneider raises calls for a rich variety of strategies. What would you think of someone who claimed that whatever goes wrong with your car can be fixed by replacing the transmission fluid?

II. \textbf{Freely They Stood Who Stood}

What does not vary is the fact that there is no way to be \textit{certain} that citizens will not abandon the values that (Brettschneider has nicely shown in \textit{Democratic Rights}) are integral to democracy. “If citizens are protected in their ability to say and believe whatever they wish by a set of liberal rights, there is no guarantee that they will engage in reflective revision [of hateful beliefs], as shown by the spread of hate groups in the United States and Europe.”\textsuperscript{19} Yes. There is no guarantee. Ever.

\textsuperscript{14} See \textsc{William T. Cavanaugh}, \textit{The Myth of Religious Violence: Secular Ideology and the Roots of Modern Conflict} 128-29 (2009).
\textsuperscript{15} \textsc{Brettschneider}, \textit{supra} note 2 at 27, 52, 69, 110, 166, 168.
\textsuperscript{16} \textit{Id.} at 80; see also \it{id}. at 79, 110, 112. This one is not only silly but incoherent.
\textsuperscript{17} \textit{Id.} at 157, 159-60.
\textsuperscript{18} \textsc{William Manchester}, \textit{American Caesar: Douglas MacArthur 1880–1964} 460-70 (1978).
\textsuperscript{19} \textsc{Brettschneider}, \textit{supra} note 2, at 43.
The problem cannot be solved by state speech or by anything else. People with free will can make bad choices.

This has been a persistent anxiety in free speech theory, and it is one that we will always have to live with.\(^{20}\) It is already present in John Milton’s germinal defense of free speech. Milton’s principal reason for opposing the licensing of printing is religious: free will means the freedom to choose evil. Salvation is to be achieved only by struggle against temptation. “Assuredly we bring not innocence into the world, we bring impurity much rather: that which purifies us is trial, and trial is by what is contrary.”\(^{21}\) It follows that “all opinions, yea errors, known, read, and collated, are of main service and assistance toward the speedy attainment of what is truest.”\(^{22}\)

Like Brettschneider, Milton was sanguine about the consequences of unregulated speech:

And though all the winds of doctrine were let loose to play upon the earth, so Truth be in the field, we do injuriously by licensing and

\(^{20}\) It is not, however, a “paradox,” as Larry Alexander alleges:

If it outlaws illiberalism, its credentials as a liberal state appear to be undermined. If it permits illiberalism, it licenses Robert Frost’s derogatory quip that liberalism can’t take its own side in an argument. Either way, liberalism appears self-contradictory and incoherent. It must either betray its principles or betray itself (and thereby betray its principles). Liberalism both appears to be possible—we’ve seen it done—and impossible (it can’t be done).

Larry Alexander, *Free Speech and ‘Democratic Persuasion’: A Response to Brettschneider* (June 11, 2013), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2277849. I can take my own side in an argument without needing to hit anyone who disagrees with me. The incoherence is present only if one presumes that human will finds intolerable the existence of other beings with free will. Jean-Paul Sartre argued this, see Andrew Koppelman, *Sex Equality and/or the Family: From Bloom vs. Okin to Rousseau vs. Hegel*, 4 YALE J.L. & HUMAN. 399, 427-28 (1992), but he wasn’t a liberal.


\(^{22}\) *Id.* at 727. The importance of a free choice between good and evil is likewise emphasized in JOHN MILTON, PARADISE LOST, bk. 3, l. 102-10 (1674), *reprinted in COMPLETE POEMS AND MAJOR PROSE* 173, 260 (Merritt Y. Hughes ed., 1957). The speaker here is God the Father, explaining why it was right to allow the rebel angels and, later, Adam to transgress:

Freely they stood who stood, and fell who fell.
Not free, what proof could they have giv’n sincere
Of true allegiance, constant Faith or Love,
Where only what they needs must do, appear’d,
Not what they would? what praise could they receive?
What pleasure I from such obedience paid,
When Will and Reason (Reason also is choice)
Useless and vain, of freedom both despoil’d,
Made passive both, had serv’d necessity,
Not mee.
prohibiting to misdoubt her strength. Let her and Falsehood grapple; who ever knew Truth put to the worse, in a free and open encounter.\(^{23}\)

But Milton’s confidence also depended on his religious views. Vincent Blasi observes:

Milton’s sweeping generalization about the strength of truth was not offered in the spirit of empirical demonstration, nor even that of didactic history. Milton here was simply affirming, once again, his faith in divine providence.”\(^{24}\) This faith “can have no secular analogue.\(^{25}\)

Milton did not rely on the state as a source of moral (which for him meant religious) authority. He doubted the state’s reliability,\(^{26}\) but he also thought that even correct religious doctrine would not bring about salvation if it was the consequence of blind conformity rather than active engagement with religious questions. “A man may be a heretic in the truth; and if he believe things only because his pastor says so, or the Assembly so determines, without knowing other reason, though his belief be true, yet the very truth he holds becomes his heresy.”\(^{27}\) Milton’s ideal citizens are, in this respect, better liberals than Brettschneider’s.

Alexander Meiklejohn, on whom Brettschneider relies, exalts free will for Miltonic reasons. In modern free speech theory, there has been a persistent problem about whether those who reject democracy are entitled to free speech.\(^{28}\) Meiklejohn offered the best response: “A government is maintained by the free consent of its citizens only so long as the choice whether or not it shall be maintained is recognized as an open choice, which the people may debate and decide, with conflicting advocacies, whenever they may choose.”\(^{29}\) Meiklejohn’s argument is essentially

\(^{23}\) MILTON, supra note 21, at 746.


\(^{25}\) Id.

\(^{26}\) See generally JOHN MILTON, CONSIDERATIONS TOUCHING THE LIKELIEST MEANS TO REMOVE HIRELINGS OUT OF THE CHURCH (1659), reprinted in COMPLETE POEMS AND MAJOR PROSE 856, 878-79 (Merrit Y. Hughes ed., 1957).

\(^{27}\) MILTON, supra note 21, at 739. Milton’s caricature of deference to religious authority anticipates Monty Python. Id. at 739-40.


\(^{29}\) Alexander Meiklejohn, What Does the First Amendment Mean?, 20 U. CHI. L. REV. 461, 468 (1953). Another, similarly Miltonic formulation: “If men are not free to ask and to answer the question, ’Shall the present form of our government be maintained or changed?’; if, when that question is asked, the two sides of the issue are not equally open for consideration, for advocacy, and for adoption, then it is impossible to speak of our government as established by the free choice of a self-governing people.” ALEXANDER MEIKLEJOHN, POLITICAL FREEDOM: THE CONSTITUTIONAL POWERS OF THE PEOPLE 123 (1960).
the same as Milton’s: in order for the choice of good to be authentic, there must be a real option to choose evil. Meiklejohn, however, is curiously evasive about the dangers. He does not have Milton’s assurance of God making it all right in the end.\footnote{Other early proponents of free speech were more explicit about the hazards of the undertaking. Justice Holmes wrote that free speech “is an experiment, as all of life is an experiment.” Abrams v. United States, 250 U.S. 616, 630 (1919) (Holmes, J., dissenting). Justice Brandeis thought that maintaining a free society required courage. Whitney v. California, 274 U.S. 357, 375, 377 (1927) (Brandeis, J., concurring); see Vincent Blasi, The First Amendment and the Ideal of Civic Courage: The Brandeis Opinion in Whitney v. California, 29 WM. & MARY L. REV. 653 (1988).}

If autonomy is the only thing that matters, of course, then this is not a problem. T.M. Scanlon argued that respect for citizens’ autonomy entails that speech cannot be prohibited simply because it results in listeners having false beliefs, or in listeners coming to believe that they ought to perform certain harmful actions, however bad the consequences of such beliefs might be.\footnote{T.M. Scanlon, Jr., Freedom of Expression and Categories of Expression, 40 U. PITT. L. REV. 519, 533 (1979).} He later recanted, precisely because the principle was too cost-insensitive. It may be that there should be some restriction on the costs that government should take as a justification for restrictions on speech, he wrote, but such an argument “must itself be based on a full consideration of all the relevant costs.”\footnote{Thomas Scanlon, A Theory of Freedom of Expression, 1 PHIL. & PUB. AFF. 204, 213-14 (1972).} Scanlon invoked Milton’s concern about authentic choice, but he retreated from his absolutism, because he lacked Milton’s assurance that authentic choice is more important than anything else.\footnote{Frederick Schauer’s critique of Scanlon misses his Miltonian dimension: Schauer thinks that if the state can deem some conduct impermissible, it logically follows that it can censor speech advocating that conduct. FREDERICK SCHAUER, FREE SPEECH: A PHILOSOPHICAL ENQUIRY 70-71 (1982). By this logic, God should not have permitted the serpent to communicate with Eve.} You can still support free speech on the grounds Scanlon articulated—I do—but you have to be open to the possibility that those grounds are overridden if the consequences are bad enough.

Brettschneider is in this tradition. Toleration of illiberal speech is demanded by respect for the independent judgment of free and equal citizens: “[I]t is essential to the legitimacy of value democracy that they could choose to embrace inegalitarian principles and policies.”\footnote{BRETTSCHNEIDER, supra note 2, at 77.} But the whole project demands that they not do so. It is inescapably consequentialist.

Milton accepted the dangerous consequences of his doctrine, because he believed there is an ultimate Guarantor to
assure us that the dangers are worth it. Brettschneider, on the
other hand, has no assurance that one day Christ will return
and make everything okay. Happy talk about state speech is a
poor substitute.

There is no way to be sure that the freedom that is argued
for will not be used to bring about the consequences that are
feared. Somehow the citizens must be persuaded to use their
freedom well. Perhaps government can do the persuading. But it
is better if the ideas of freedom and equality are so deeply
ingrained in the culture that the government has no work to do.

Any consequentialist case for a given liberty is likely to
rest on some combination of optimism (about what will happen
absent regulation) and distrust (of the government actors who
would regulate the liberty in question).\(^\text{35}\) The warrant for either
of these will vary from one society to another. Too much
optimism and you have the Hateful Society. Too little distrust
and you have the Invasive State.

In the United States, as it happens, there is plenty of
evidence that supports optimism about a regime with robust free
speech protection.\(^\text{36}\) Overt racism has been nearly eliminated
during the period when the First Amendment was construed to
give strong protection to racist speech, and it is not at all clear that
the state’s tutelary function, rather than vigorous discussion within
civil society, was responsible.\(^\text{37}\) The growth of religious tolerance,
too, is largely the product of private initiative rather than state
speech.\(^\text{38}\) Maybe state speech can help, but assessing when and
whether this is so requires a more nuanced understanding of the
cultural basis of liberalism than Brettschneider offers.

\(^{35}\) Andrew Koppelman with Tobias Barrington Wolff, A Right to
Discriminate? How the Case of Boy Scouts of America v. James Dale Warped

\(^{36}\) See Andrew Koppelman, Waldron, Responsibility-Rights, and Hate Speech,

\(^{37}\) The most important state intervention was probably the Civil Rights Act of
1964, but it did considerably more than lecture citizens about equality.

\(^{38}\) See Kevin M. Schultz, Tri-Faith America: How Catholics and Jews
Held Postwar America to Its Protestant Promise (2011).
Liberty, Equality, and State Responsibilities

REVIEW OF COREY BRETTSCHNEIDER’S WHEN THE STATE SPEAKS, WHAT SHOULD IT SAY?

Robin West†

Corey Brettschneider’s When the State Speaks confronts a core dilemma for liberalism and indeed for liberal states: how (and whether) liberal states should respond to the existence of hateful speech and practices, and the groups that sponsor and promote them. Brettschneider advocates for an approach that checks the damage the hateful speech does to underlying liberal principles of free and equal citizenship, while at the same time respecting the rights of the speakers and groups that engage in it. He rejects what he considers to be the two polar responses that pervade state responses, both in the U.S. and elsewhere, to this dilemma, and that virtually exhaust the scholarly treatment of the issue, at least in the U.S., the U.K., and Canada.

On the one hand, Brettschneider rejects the civil libertarian (or “neutralist”) claim, popular in the United States, that private speech is just that—private—and therefore of no concern or relevance to public values, public deliberation, or public law. He likewise rejects the “militant democrats” (his phrase), some feminists, and most of the European liberal democracies, who argue that private hateful speech has very harmful and fully intended consequences and should be banned or censored in some meaningful way to stop its noxious spread.¹

These two poles, Brettschneider argues, veer toward one or the other of two dystopian visions of the relation of the state to its citizens. The “militant egalitarian’s” view, which urges greater criminalization of hate speech, risks what he calls the “Invasive State,” meaning a state overly involved in our private

† Professor of Law, Georgetown University Law Center. Part of this review has been published as a part of Liberal Responsibilities, 49 TULSA L. REV. 393 (2013).
¹ COREY BRETTSCHNEIDER, WHEN THE STATE SPEAKS, WHAT SHOULD IT SAY?: HOW DEMOCRACIES CAN PROTECT EXPRESSION AND PROMOTE EQUALITY 1-3, 5 (2012).
lives—the traditional bogeyman of civil libertarians everywhere.\(^2\)

The liberal or neutralist view, on the other hand, according to which the state is and should be fundamentally unconcerned with the content of private speech, no matter how hateful or indeed how consequentially harmful, risks what he calls the “Hateful Society.”\(^3\) In this dystopia, all rights and liberties are vigorously protected, but hate runs like an open sewer, undercutting the reasons we have rights in the first place. Under such a regime, some groups of citizens—women, racial minorities, gay and lesbian citizens—are frequently and even routinely subjected to hateful practices and utterances.\(^4\)

The question Brettschneider raises and tries to answer in the book is how we can steer a middle course between these undesirable societies. On the one hand is the Hateful Society, in which rights are protected but hate runs rampant, resultantly feeding a lesser regard for the equality of citizens. On the other is the Invasive State, in which hate is checked but the state is a far-too-intrusive censor into our private lives, risking the protection of all speech, and thus compromising the joy and value those deeply human activities and attributes bring.

Brettschneider’s provocative suggestion is to introduce a third possibility, fully captured by his introduction and descriptive account of the “Persuasive State.” The Persuasive State, unlike the Invasive State, refrains from coercion, and thus avoids the invasive state’s pitfalls. However, it also recognizes the relevance to public values of privately held and promulgated hateful beliefs, including those promulgated within the family and within religious traditions.\(^5\) It can thereby at least attend to, though it is uncertain whether it can fully counter, the dangers of the Hateful Society.

Brettschneider argues that the state should seek to persuade citizen holders of hateful beliefs to transform, modify or drop their hateful beliefs, to whatever extent those beliefs conflict with public democratic values, notably, values of free and equal citizenship. Thus, in effect, the Persuasive State counters hateful speech with argument—argument that hateful beliefs undercut the very values of free and equal citizenship that undergird the rights enjoyed by the holders of those beliefs.

\(^2\) Id. at 10.
\(^3\) Id.
\(^4\) Id. at 10-11.
\(^5\) Id. at 12-19.
themselves. Perhaps those with hateful views will be persuaded, and will drop the views. But even if not, other citizens will hear the dialogue, the result being that the state will have been respectful of the equal rights of all and will not have been complicit in the spreading of beliefs that fundamentally undercut liberal democracy. Indeed, the state has a moral responsibility to engage in this speech; it owes it to the victims of the hate speech it protects, and to the greater mandates of constitutionalism and liberalism. It is essential to the stability, fraternity, and equality of the community that the hate speech it protects seeks to undermine.

I am largely sympathetic to this project. I think it is entirely right for us to recognize the relevance of private hateful beliefs to public values such as equality and freedom, equal respect, and due regard. I also agree that it is entirely right for us to shift our focus somewhat from concerns about the state’s overuse of its coercive role to the possible good a state can do when it acts in its persuasive capacity. And I think it is right and extremely important that we move our attention from debates over individuals’ rights to engage in hate speech to the state’s responsibilities to respond to it and what the state should say in response.

This is a refreshing change of focus. Brettschneider shows us how to think of the state as a fully moral actor in this ongoing liberal project, leaving behind the two roles our traditional debates on this issue have articulated for it: roles as the generator of unconstitutional laws that inhibit speech, or the sometimes overly zealous protector of individual rights to engage in it. The state is more complex than this simplistic dichotomy and can multi-task with the best of us. One thing the state can and should do, Brettschneider wisely points out, is publicly make the case for the liberty and equality for which rights exist. This is particularly important where the speech those rights protect undermines the values free speech is designed to serve. That insight alone, and that prescriptive suggestion for change, is a contribution to First Amendment doctrine and theory, as well as to our understanding of equality.

However, there are problems, largely of coherence, that make it hard to understand the full import of Brettschneider’s proposal. I will focus on three such problems, as well as some more substantive objections, providing suggestions for addressing these issues.

First, it is truly difficult to understand who and what Brettschneider is talking about when he talks about state speech, or what he could possibly envision by suggesting that the state
take up and use the microphone pictured on the front jacket cover of this book. The state does, after all, already “speak” and attempt to persuade constantly; it is never quiet. The state speaks when it passes laws; when it justifies them in judicial decisions; when it promulgates administrative regulations and adjudicates those regulations; when it imposes sanctions in civil cases; when it educates children in public schools; and when it imprisons, fines, and executes people. Almost all of that speech, furthermore, is “persuasive.” Persuasive state speech is as present as air. Of course the state could and should use its rhetorical powers to promote liberal values of equality and freedom, equal respect, due regard, and human dignity. It already does this, but there is no reason it should not be urged to do so both more, and more reflectively and effectively. There is also no reason that the state should not do so in the specific context of hate speech and pornography. This is what I take Brettschneider to be urging, and I support the effort.

It would be a much clearer proposal, though, if he specified what sort of “speechifying” he has in mind. In the last chapter, he suggests doctrinal changes to our First Amendment case law that would give executive branch officials greater latitude to deny tax exemptions and deductions (as well as affirmative grants) to even purportedly religious groups that engage in hate speech. This is the first example Brettschneider provides of the sort of persuasive state speech he is urging, finally giving some context to an otherwise vague suggestion: “persuasion” apparently includes the act of denying tax exemptions and withholding economic tax dollar support, and the “Persuasive State” apparently includes the state that so withholds.

But Brettschneider seems to have in mind more than this doctrinal change that comes late in the book and late in the argument. It would have been helpful if he had simply provided more examples of the sorts of state speech he envisions. For example, a thorough treatment of issues surrounding public education (which he does give some cursory attention)—from struggles over the content of curricula and school textbooks, to the merits and perils of unregulated homeschooling—might have been a good place to start in putting a bit more meat on the bones of what is otherwise a somewhat empty, although aspirational, exhortation: the state should engage in more liberal moralizing.

---

6 Id. at 161-63.
7 Id. at 96-104.
Second, there seems to be something unrealistic, and a bit naïve-sounding, about the nature of the hate speech Brettschneider seeks to both protect (through traditional First Amendment doctrine) and counter (through the persuasive efforts of the liberal state). Brettschneider writes as though that speech is largely ideological, spawned in political groups such as the KKK, private organizations such as the Jaycees, or in families that have and transmit to children overtly racist or sexist world views. The counter to such bad ideas, Brettschneider sensibly maintains, are good ideas. What the state should do, then, about these groups with their wrongheaded and inegalitarian ideologies, is not censor them, but persuade their members that they should abandon their beliefs, at least to whatever degree they conflict with a thin and liberal conception of equal and free citizenship. That is, families that teach their daughters to accept a lesser and subordinate role than men, groups such as the Jaycees that teach their members to accept discriminatory membership policies, the Boy Scouts with their homophobic refusal to acknowledge the full humanity of gay men and boys, and groups such as the KKK that spread genocidal messages of hatred and contempt, should not be censored. Rather, their members should be exhorted not to harbor their false beliefs. The members should in turn listen, and then change their beliefs accordingly. (It sounds a bit like Romney’s “self-deportation” solution to the immigration problem.) The state should, in effect, reason and argue, not punish. If it does, it will win the argument, and that’s that.

There are two problems with this part of Brettschneider’s prescriptive claim. First, at least sometimes, hate speech is motivated not by worked out noxious ideologies, such as a doctrinaire belief in white supremacy or male superiority, or any other set of beliefs, but rather, by literal hate, pure and simple. Its targeted audience is not society in general, reachable through marches or rallies in the public square with signs and cross burnings, but rather a particular teenage girl, or a gay boy, or an insecure and vulnerable tween, targeted through harassing speech and images on Facebook. That harassing and harmful speech spawns not just feelings of inferiority, but self-inflicted cuttings, eating disorders, and suicides.

---

8 Id. at 8, 72.
9 Id. at 54.
10 Id. at 130-32.
11 Id. at 8.
What prompts it? Of course, sometimes ideology plays a role: the harasser thinks, wrongly, that women or gay boys or black children are of lesser worth. If he or she could simply understand the wrongness of that belief the behavior would stop. But, at least sometimes and perhaps more often than not, that speech is the product of twisted psyches rather than noxious beliefs—not the sort of mindset to be changed through persuasive state-sponsored speech. Brettschneider seems to have forgotten, or just underplayed, the role of the “hate” in the hate speech he rightly deplores. Hate speech is not solely the product of wrong beliefs; it is also the product, in part, of hate.\(^{12}\)

Second, and relatedly, it is difficult to buy into Brettschneider’s insistence on the moral responsibility of people who hold these noxious views to change them so that they accord with the minimal decency required by liberal principles, and that the state should be engaged in the project of exhorting them to do so. Is it really morally incumbent upon everyone who harbors these false beliefs to listen to the state’s arguments against them and change their beliefs accordingly? Consider the families that keep their children home so as to avoid the liberalizing content of a public school education, and teach their daughters submissive and lesser roles, or the boy who torments his gay peers, or the white housewife who despises her black neighbors and lets them know it. Presumably, these individuals dislike the egalitarian, assimilationist, and integrated liberal state and its exhortative, persuading ways, at least as much or more as the loosely liberal feminist and egalitarian ideas of gender and race equality the liberal state promulgates.\(^{13}\) Like teenagers who are spurred on to double their consumption of cigarettes and alcohol by public health campaigns, surely illiberalists will be spurred to greater contempt for any state that tries to persuade them out of their entrenched habits of thought, feeling, and child raising—habits that have at their core, not periphery, fear of and contempt for a state that preaches liberal values.

It is difficult to comprehend what is meant by the claim that the liberal state should exhort people who hate the liberal state precisely because of its tendency toward exhortation of views they despise. It seems imprudent to advocate that the state’s solution to citizens stating, for example, “I hate the state

\(^{12}\) For a very different description, see DANIELLE CITRON, HATE 3.0: THE RISE OF DISCRIMINATORY CYBER HARASSMENT AND HOW TO STOP IT (forthcoming).

\(^{13}\) For an account from a judicial decision adjudicating a conflict between such a family and a public school, see Mozert v. Hawkins Cnty. Bd. of Educ., 827 F.2d 1058 (6th Cir. 1987).
because it beats up on me,” should be to continue bullying these citizens until they change their beliefs that a bullying state that beats up on them is something they hate. The very attempt at persuasion may very well underscore precisely what extreme anti-liberals find so hateful about the liberal state; it is what spurs the fringes among them to arm themselves in defense against black helicopters. Persuasive exhortatory politically correct talk by the liberal state, in other words, will backfire—just ask any teenager. Thus, Brettschneider’s proposal would have been stronger had he specified more carefully the kind of hate speech he is targeting, and the kinds of responses that are envisioned, when he speaks of acts of state persuasion.

I also identified some substantive problems with the proposal itself: that the state’s response to hate speech, contra the censorship urged by militant democrats and contra the blind eye turned by civil libertarians, should be to engage in various acts of persuasion. There are three separate problems that I see. The first is a general worry about the way Brettschneider has characterized the Persuasive State as an alternative to his two dystopias: the Invasive State on the one hand and the Hateful Society on the other. Brettschneider’s Persuasive State is a well-intentioned, completely functional, and thoroughly liberal state, not beset by public choice woes, or administrative capture, or excessive corporate and private power, or know-nothing voters, politicians and their political parties, or noxious factionalism, or sagging economies, or even external threats. It’s a pretty utopian state. Of course, it is also a hypothetical and idealized construct, meant to rhetorically counter the two dystopias he portrays, and which are implied by the poles of debate he is seeking to interrupt. It is not intended to describe any actual state, liberal or otherwise. Nevertheless, there is an odd blindness in Brettschneider’s project to the distinctive dystopic possibilities to which his own utopian alternative might give rise.

Bluntly, it seems to me that in addition to the two dystopias that drive Brettschneider’s project—the Invasive State and the Hateful Society—he might have added a third, implied by his own proffered alternative: the Hypocritical State. Just as the Invasive State is the dystopia masked by the egalitarians’ political aspirations, and the Hateful Society, the dystopia implied by the civil libertarians’, so the Hypocritical State, I suggest, is the dystopia possibly masked and legitimated by Brettschneider’s proffered alternative. Sometimes, when actual liberal states speak of the values of equality, diversity, and liberty, they are doing so at the very moment they are pursuing
profoundly inegalitarian, stultifying and oppressive ends. All that egalitarian persuasion might be toward the end of distracting people from the inegalitarian acts it is undertaking. In this dystopian vision of Brettschneider’s proffered alternative, “state persuasion” is designed, roughly, to drive the listener crazy, by insisting the state is doing the opposite of what it is doing at that very moment. Sometimes, the Persuasive State’s egalitarian and freedom-respecting rhetoric is cover for actions that are viciously inegalitarian and disrespectful—even hateful. Sometimes, to use the critical language of the 1970s and 1980s, the Persuasive State is the “legitimating state”—its equality and freedom-promoting rhetoric is a cover for deeply illiberal impulses. In my view Brettschneider should have worried about that possibility, at least a little.

We need look no further in our own contemporary society than the very high-minded liberal, liberty enhancing, and thoroughly moral justifications that various state actors—Supreme Court Justices, ninth grade civics teachers, state and federal prosecutors, pro bono lawyers from prestigious private law firms, judicial opinions by the bucket-load, and academic lawyers in the field—proffer, when discussing norms of our criminal law and procedure. Juxtapose all that equality and liberty-promoting rhetoric about presumptions of innocence and the dignity-protecting rights of the worst criminal defendants, with the reality of our grotesquely dysfunctional criminal justice system, promulgated by those same state actors, that incarcerates more citizens per capita than some of the most ghastly, illiberal, and totalitarian dictatorships on the planet. The same “state” that speaks of the dignity, equality, and respect owed to criminal defendants, when justifying criminal law, imposes penalties for the possession of crack cocaine at one hundred times the harshness as possession of powder cocaine, and life sentences for trivial and victimless as well as unproven crimes, executes prisoners in the face of proffered and

unexamined evidence of innocence and does so in a way that disproportionately kills African American citizens, and targets citizens abroad for execution on hidden evidence of their alleged complicity with terror presented in secret tribunals.

The Hypocritical State or the Legitimating State (or the lamp-posting state) legitimates all of this mayhem and random violence it inflicts with high-flying language justifying its sanctions in terms completely congenial to liberal rights. It employs the rhetorical mechanisms of persuasion, in other words, so as to coercively impose its mandates in ways that express its utter contempt for the very moral and liberal values it self-righteously, loudly, and repeatedly extols.

Surely the Hypocritical State is a danger worth attending. But Brettschneider does not at all address this possibility. Rather, the Persuasive State of Brettschneider’s imagination utters and deeply believes liberal and democratic values, particularly the value of equal and free citizenship. There is no risk of hypocrisy anywhere in sight. The extraordinary gap between the liberal persuasive state and the liberal incarceral and executing state may be the largest in the criminal justice field, but it by no means resides solely there. It may, in fact, also reside in the distance between the rhetoric of equality that it deploys, and that Brettschneider wants it to deploy more loudly, and the wide berth it gives to the hate speech it tolerates. In other words, the Persuasive State of Brettschneider’s imagining might just be protesting a bit too loudly. It seems to me that particularly those of us attracted to the idea of the Persuasive State, and its potential for good, need to worry about this legitimating function of the Persuasive State, and to try to find ways to counter its influence.

The second substantive problem regards the myriad purposes served in Brettschneider’s book by the line drawn between the Persuasive and the Coercive State, or the state when it is acting in its “expressive” mode, and when it is acting in its coercive mode. The distinction is a vital one in the book, not just theoretically but practically. It lies at the heart of the most helpful doctrinal reform Brettschneider advocates: that First Amendment law, and particularly the unconstitutional conditions doctrine, be redrawn so as to unequivocally permit the state to refuse to fund groups that sponsor illiberal beliefs—even where those groups do so under cover of religious dogma—through withholding either direct grants or tax exemptions and deductions.

The liberal state cannot and should not censor the expression of offensive beliefs. The liberal state should not slide into the Coercive State and, for the most part, cannot by virtue
of the First Amendment. The Persuasive State, though, can and should respond to the content of the ideas it is required to protect, and one way it can respond is by refusing to assist those groups in the marketplace of ideas. This important suggestion, though, obviously depends quite heavily on the distinction, if it can be maintained, between the “Coercive State” and the “Persuasive State.”

According to Brettschneider, “persuasion” includes, notably, the act of withholding tax exempt status, and refusals to extend grants. However, as he notes, the Supreme Court for the most part disagrees, as do many commentators. A bit more development of the meaning of “persuasion,” “coercion,” or both, might have helped bridge the gap between them.

More generally, though, the distinction Brettschneider draws between the coercive and persuasive states is a little too black and white, even aside from its application to the problem of tax exemptions for religious or other groups that promulgate hate. The state does many things, at least some of which are not easily categorized as clear examples of coercion or persuasion. Tax deductions and exemptions are just one such close case. There are other borderline cases as well. The state also regulates and coordinates across vast areas of social life, all of which can be understood as either “coercion” or “persuasion” under sufficiently broad definitions of either term.

Here is one example of a potentially ambiguous case, which is completely undiscussed in Brettschneider’s book, but which is of direct relevance to his general thesis. The “state” administers a system of private law, including an array of tort remedies, so that citizens can pursue, through actions for monetary damages, some measure of corrective justice when they have been wronged without turning to the punitive “coercive” arm of the state. The state obviously facilitates these private actions. It is also in some sense responsible for the doctrine under which those actions proceed. It isn’t clear, though, in Brettschneider’s treatment, where this private-rights-and-private-remedies-providing function falls. Does the state’s provision of tort remedies for private wrongs constitute “persuasion” or “coercion”? It is not clear. It is also not clear where Brettschneider might think it falls, particularly given that he considers the construction of incentives and disincentives created through tax structures or the withholding of largesse from

---

15 Brettschneider, supra note 1, at 13, 109.
16 Id. at 128-40.
17 Id. at 137-40.
the public purse, examples of “persuasion” despite their arguably categorically coercive consequences.

It is a question, though, that is of obvious relevance to the issue that absorbs the book: the regulation, or recognition of hate speech, and the state’s response to it. The distinction between the criminal law—which is clearly coercive—on the one hand, and regulatory or tort regimes—in which the state provides the forum and determines the remedy but private parties bring the actions—has been of vital importance to lawyers seeking to provide remedies for victims of hate speech without encroaching on First Amendment rights. Brettschneider’s overall project is somewhat aligned with that of these lawyers, at least in its goal; he too wants to figure out a way to respond to the anti-equalitarian content of hate speech without offending First Amendment guarantees. Thus, he might have looked a little more closely at some of those campaigns.

For example, at least parts of the failed anti-pornography ordinances of the 1980s pursued by Catherine MacKinnon and Andrea Dworkin were specifically designed so as not to employ the punitive arm of the state. What was envisioned in some parts of those ordinances were civil actions for civil remedies, contemplating the imposition of monetary damages to be paid to the victims of pornography by purveyors; not punitive actions involving the criminal law, the state, and jail time. Do those actions, and the state’s facilitation of them, constitute exemplars of “coercion” or “persuasion”? It seemed to the backers of these ordinances that the state’s role in these actions was what Brettschneider calls persuasive, not coercive. I suspect that Brettschneider would disagree. But if so, this needs a defense. Is the state’s role in facilitating private actions for the harms done by hate speech or pornography, so injurious to the rights of purveyors to justify a ban on even such civil actions? This goes quite a bit further than what seems to be argued in the text, which is that criminal sanctions on hate speech would do such injury.

Recognizing the violations of rights arguably occasioned by criminalizing various forms of hate speech by no means implies that a state which permits civil actions for those harms, followed by the imposition of monetary damages, also violates those rights. Rather, it appears to require only that the speakers internalize the harms occasioned by that speech, when the

---

speech violates the civil rights of others, with the phrase “civil rights” there understood in its original and ordinary way: rights to dignity, to bodily integrity, and to be free of the fear of assault in both society and cyberspace. The state provides the forum—the courthouse, court personnel, and the cause of action—such that those civil rights can be vindicated in the form of private actions. It seems to me that the state is performing a persuasive, and not a coercive function, when it acts as such. But if there is a claim to be made to the contrary, it isn’t spelled out in this book.

Another way to put this concern is that it isn’t clear where civil actions, as opposed to criminal sanctions, for hate speech and pornography would fall on Brettschneider’s schematic divide between “coercive” and “persuasive” state functions. The state acts through common and constitutional tort law so as to inhibit or deter a wide range of speech acts, involving an equally wide range of types of harms. Thus, the state defines as “tortious” various types of libel, slander, fraud, blackmail, perjury, assault, stalking, harassment, group defamation and intentional infliction of emotional distress. (Some of these, of course, are also crimes.) Likewise, it defines as actionable various breaches of contract, copyright, patent, and trademark rights, virtually all of which also involve nothing or almost nothing but speech and speech acts. All of these torts and contract or copyright violations are occasioned by speech acts, and they all give rise to civil sanctions. The speech in all of these cases, like hate speech, does things, and (arguably) by so doing, causes harm. And the state’s response likewise does something: the state defines torts, and court rules, so as to facilitate a damage award, a punitive damages award, or a restraining order, in response to the harm this speech causes. For the most part, when it does so, it is facilitating a private remedy designed to require the actor to rectify the wrongs done by his actions, and to internalize those costs.

The same would be true of a (hypothetical) civil action for harms occasioned by hate speech or pornography. Yet, only with respect to such harmful speech, are liberals, and possibly Brettschneider (although it is not entirely clear), inclined to close the courthouse door, and preclude not only criminal responses by the state, but also civil responses by harmed citizens. But why? Why is it only with respect to hate speech that the state must not permit civil recourse? Why is it only the harms occasioned by pornography or racist speech that inspire worries that any state that allows civil remedies is thereby acting coercively, rather than persuasively? What is so peculiar about hate speech, and the
harms it occasions, that gives rise to the impulse to shield them with protection against a civil state response?

More generally, the failure to define coercion, and its reach, has additional normative consequences in Brettschneider’s argument. To reverse Robert Cover’s insight from decades ago that all state speech takes place on a field of violence, all coercive state actions, from executions to the impositions of regulatory fines, civil sanctions, parking tickets or taxes, also take place within a field of persuasive words. There is often coercion behind state attempts to be persuasive. Think of the interactions between the pregnant woman seeking an abortion and the abortion provider required to educate the woman with respect to various attributes of the fetal life inside her. But there is also often, perhaps very often and maybe even always, persuasion behind the state’s attempts to coerce.

The efficacy of persuasive action requires an audience and the state sometimes attempts to coerce attendance. Think again of public education. Is it always wrong to require attendance? If not, this suggests a continuum, rather than a bright line, between acts of the state that are primarily persuasive but accompanied by some measure of coercion—even if just coerced attendance—and acts of coercion accompanied by some measure of persuasion. Sometimes, the persuasion that accompanies coercive state action is just obnoxious and grating and we would be better off without it—just tell me what it will cost me, don’t lecture me, when you’re imposing a traffic fine. Sometimes, the coercion that might accompany persuasive action, however, is justified. For example, think of the attachment of the salary that accompanies the imposition of a civil or regulatory fine, or as Brettschneider argues in some detail, the withdrawal of a tax exemption from religious views that are hateful. The tax exemption, however, is not an anomaly; it is, rather, simply one example of a pervasive dynamic: persuasive state speech accompanied by some measure of coercion. That dynamic, I think, requires more general treatment.

A final concern goes to the narrowness of Brettschneider’s thesis. Why are we focused so exclusively on the role of the persuasive state in promoting race and sex equality, rather than also on values pertaining to a fair or just distribution of resources? Shouldn’t the persuasive state pursue, through persuasion, these liberal values as well? The paradox Brettschneider discovers of the state protecting hate speech for liberal reasons, even though the

---

speech it protects undermines liberalism, is not limited to the context of speech and speech rights. Much of the dynamic Brettschneider describes—of the back and forth between the dystopic images of an overly invasive state on the one hand or a hateful society on the other—is also true in the realm of economic liberty and justice.

In part for liberal-sounding reasons regarding the dangers of concentrated state power and the values of individualism and entrepreneurship, we tend to let market-generated outcomes lie, even when that results in massive inequalities in wealth, and hence concentrations of economic power, which can squelch true individualism and liberty in the private sector. Perhaps one way out of this box as well as in the speech context would be to look at the role of the persuasive state. A state that can explain its commitment to markets might go some way toward undoing the damage unregulated markets occasion; or on the other side, a state that can explain its commitment to regulating them might also bear a responsibility to explain how that regulation in fact furthers the liberties that it seemingly constrains. A shift in focus away from the pitfalls of the state’s coercive role in taxing and redistributing wealth, or in regulating private markets, and toward the potential of the state’s persuasive role in these contexts, might yield benefits comparable to the shift in focus that Brettschneider advocates, and largely accomplishes, in the context of hate speech. Were we to try this, however, we’d face some of the problems outlined above.

In addition to the invasive state and hateful society dystopias, we would also have to contend with the inegalitarian state, and the illiberal state, as well as the hypocritical or lamp-posting state. We would have to contend with the various state acts that impact distributive justice concerns, but where the action is neither cleanly coercive nor cleanly persuasive, but some combination of both. Here, the weakness of the “coercive versus persuasive” understanding of the state would come into sharp relief: the main way that the state affects distributive justice is through its regulatory, taxing and spending authority, all of which are viewed by some as fundamentally coercive, and by others as the paradigmatic exercise in persuasive statecraft. Similarly, its laws of inheritance, contract, and property rights, all of which aim to give full sway to private choice, are viewed by some—critical theorists most notably, but more generally, the political left—as so coercive as to be theft, and by others—primarily libertarians and liberals but some communitarians as well—as the heart of civil society.
In all of these areas the state’s acts of coercion are so intertwined with acts of persuasion that it will be hard to disentangle them, suggesting, I believe, the limits of the distinction. It might be better in this context, as well as in the context of hate speech, to look in a more granular way at the specific acts the state has taken without categorizing them as coercive or persuasive, and simply ask whether the state is saying the right things when it speaks, as the wonderful title of Brettschneider’s book suggests.
The Liberal Tightrope: Brettschneider on Free Speech

Sarah Song†

INTRODUCTION

Corey Brettschneider’s book takes up one of the most important questions in moral, political, and legal philosophy—what are the grounds and limits of toleration—by focusing on the problem of hate speech. I admire his aspiration of crafting a middle position that defends robust free speech protections while also subjecting those espousing hateful viewpoints to “democratic persuasion.”

Brettschneider begins with a critique of two dominant responses offered by political and legal theorists to the problem of hate speech. On the one hand, you have the “neutralists” who defend strong rights of free speech, including protections for “hateful viewpoints,” which he defines as “hostile to the core ideals of liberal democracy.”¹ On the other, you have the “prohibitionists” who endorse coercive bans on hate speech.² The first approach is reflected in the “United States Supreme Court’s current free speech jurisprudence,” whereas the second is “found in most liberal democracies outside the United States.”³ Brettschneider argues that both approaches are problematic. Neutralists fail to take seriously the ways in which hateful viewpoints undermine the ideal of free and equal citizenship. Prohibitionists disregard the autonomy of citizens, their capacity to choose, revise, and pursue their own conceptions of the good.

Brettschneider proposes a middle path between these two positions, based on a distinction “between a state’s coercive power, or its ability to place legal limits on hate speech, and its expressive

† Professor of Law and Associate Professor of Political Science, U.C. Berkeley.
¹ COREY BRETTSCHNEIDER, WHEN THE STATE SPEAKS, WHAT SHOULD IT SAY? HOW DEMOCRACIES CAN PROTECT EXPRESSION AND PROMOTE EQUALITY 1 (2012) (internal quotation marks omitted).
² Id.
³ Id.
power, or its ability to influence beliefs and behavior by ‘speaking’ to hate groups and the larger society.”4 The state should refrain from using its coercive capacity against groups espousing hateful viewpoints in the name of respecting freedom of expression while simultaneously criticizing such groups in its expressive capacity in the name of respecting equal citizenship. Such a middle path provides a more nuanced approach but it is also like walking a tightrope. Teeter too far in one direction and you endanger liberty; lean too far the other way and you sacrifice equality.

Another way of characterizing the aim of Brettschneider’s book is in terms of two liberal dystopias to be avoided. For neutralists, the *summum malum* is the Invasive State where government officials constantly monitor and punish citizens for speaking or acting on beliefs contrary to the ideal of free and equal citizenship. By contrast, prohibitionists fear something different: the Hateful Society in which robust liberal rights protections permit a thousand hateful viewpoints to bloom. Brettschneider argues we should guard against both dystopias. He agrees with the neutralists that the state should not coercively infringe upon free speech and privacy rights, but he also sides with the prohibitionists in recommending that the state actively engage in “democratic persuasion,” which involves criticizing and even condemning hateful viewpoints. When the state “speaks,” it should not only remind citizens of the content of rights; it should also present the reasons for rights. The reasons for the liberal rights of freedom of speech, association, and religion are none other than the values of freedom and equality at the core of liberal democracy. Any moral or religious viewpoints incompatible with these core values are properly subject to “democratic persuasion” even while being shielded from coercion.

My focus in this essay is to develop and amplify concerns that prohibitionists, on one side, and neutralists, on the other, will continue to have about Brettschneider’s “third alternative.” To build a middle position aimed at addressing the concerns of two opposing sides inevitably means building a framework around a tension. Each side has reason to think Brettschneider’s position betrays its core concerns to appease the other side. Are they right?

---

4 *Id.* at 3.
I. THE PROHIBITIONIST OBJECTION: DEMOCRATIC PERSUASION DOESN’T GO FAR ENOUGH

The traditional liberal approach to drawing the limits of state regulation has been based on a spatial metaphor between “the public” and “the private.” On this view, certain “spaces” of life—the family, the market, churches, and voluntary associations—are deemed to be private spaces and thereby free of state regulation. Building on the work of feminist theorists, Brettschneider argues that no space is immune from state regulation. The state can override privacy rights wherever those rights are exercised in ways that violate the rights of others. This point is pithily captured by the feminist slogan, “the personal is political.”

In lieu of the spatial metaphor, Brettschneider offers what he calls the “principle of public relevance” to determine the limits of the state regulation.5 “What makes certain beliefs and practices publicly relevant is that they conflict with our public status as free and equal citizens.”6 In cases of conflict, personal beliefs and actions “ideally should be changed to make them consistent with the ideal of free and equal citizenship.”7 To implement the principle of public relevance, we need to answer a prior question: what is required to treat others as free and equal citizens?

In fleshing out the ideal of free and equal citizenship, Brettschneider seems most concerned to show how minimalist that ideal is. Following John Rawls, the leading liberal philosopher of the twentieth century, he characterizes his principle of public relevance as “political, not metaphysical.” It “is not ‘deep’ in terms of either its justification or the types of beliefs to which it applies.”8 To demonstrate why the ideal is not deep, Brettschneider elaborates the kind of freedom and equality that is required:

According to value democracy, citizens must be treated as having equal status in that the rights of all citizens must be equally respected. These rights include freedom of expression, association, and religion, as well as rights of political participation and the rule of law. Citizens are to be regarded as “free” in the sense of being able to possess and exercise the rights of citizenship.9

The “political, not metaphysical” move appeals to many contemporary liberals for the familiar reasons Rawls gave.

5 Id. at 26.
6 Id.
7 Id.
8 Id. at 30.
9 Id. at 31.
Political liberals aim to develop political principles that could be acceptable to citizens who hold a wide variety of comprehensive moral and religious doctrines. Democratic societies are characterized by the fact of reasonable pluralism—the plurality of comprehensive doctrines that people hold is reasonable because when people are left to judge for themselves on fundamental moral and religious questions, they inevitably come to different conclusions due to what Rawls calls the “burdens of judgment.”

Rawls’s *Political Liberalism* offers a political conception of justice that may command widespread agreement in the form of an overlapping consensus in spite of the fact of reasonable pluralism. A political conception of justice is “political” in virtue of being presented independently of any comprehensive moral doctrines and because it is worked out from fundamental ideas seen as implicit in the public political culture of a democratic society. Brettschneider’s value democracy, like Rawls's political liberalism, requires all citizens to accept the core values of free and equal citizenship but not any comprehensive moral or religious doctrines. Rawls himself maintained that among the core values of liberal democracy “are the freedom and equality of women, the equality of children as future citizens, [and] the freedom of religion.”

Here is the objection I want to raise with “political, not metaphysical” approaches on behalf of “prohibitionists.” I draw here on the feminist political theorist Susan Okin’s critique of Rawls. Okin emphasized that political liberals’ “toleration of a wide range” of personal beliefs and practices tends to come at the expense of the equality of women.

Many political liberals, including Brettschneider, respond that public toleration of a plurality of moral and religious doctrines is conditional on those doctrines being compatible with the ideal of free and equal citizenship. That is a crucial limit of liberal toleration. In fleshing out the ideal of free and equal citizenship, Brettschneider says it requires recognizing citizens as “politically autonomous and equal.” He sums it up this way:

> In short, the “freedom” in free and equal citizenship requires respect for both the political and personal autonomy that citizens need to

---

11 *Id.* at 12-13.
develop opinions about politics and the good life. The “equal” in free and equal citizenship reflects a concern to ensure that these capacities are respected equally regardless of one’s race, ethnicity, or gender.\textsuperscript{14}

Brettschneider points to slavery, segregation, and race- and sex-based exclusion from the vote as clear violations of the ideal of free and equal citizenship. But what about less clear-cut cases that involve doctrines supporting inequalities based on race, sex, and sexual orientation that are perpetuated within families and houses of worship?

Take, for example, the case of the Catholic Church’s prohibition of female priests. Insofar as the Church does not directly oppose the equal status of women in society at large, Brettschneider says it should not be subject to democratic persuasion. If it were to advocate policies that dissuade women from serving in public life, however, the case for cutting off state-funded tax privileges becomes stronger. In Brettschneider’s view, the state’s subsidy power is a tool of persuasion, not coercion, and the state should withhold subsidies from groups that violate the ideal of free and equal citizenship. In the case of the Catholic Church, Brettschneider argues that the Church should be immune from democratic persuasion because whether the Church’s policies are actively opposed to the ideal of free and equal citizenship is “ambiguous” at best. Democratic persuasion should be restricted “to clear instances in which the ideal of free and equal citizenship is opposed.”\textsuperscript{15} He imposes this “substance-based limit” on democratic persuasion to appease the worries of neutralist liberals, but in doing so, does he betray his other commitment to combating the spread of inequality-breeding doctrines?

I think Brettschneider is too quick in dismissing the state’s use of democratic persuasion in this case. Why should the ambiguity of a case not trigger democratic persuasion instead of ruling it out? A fuller appreciation of the feminist slogan “the personal is political” requires recognizing that the influence runs in both directions. The slogan stands not only for the proposition that public regulation of private beliefs and practices is sometimes justified, but also that private beliefs and practices strongly shape political life. Okin emphasizes the latter when she says, “it is exceedingly difficult to see how one could both hold and practice (in one’s personal, familial, and associational life) the belief that women or blacks, say, are naturally inferior, without its seriously affecting one’s capacity to relate

\textsuperscript{14} Brettschneider, supra note 1, at 35.

\textsuperscript{15} Id. at 135.
(politically) to such people as citizens ‘free and equal’ with oneself.” 16 We need not agree with Okin’s characterization of the world’s major religions as “rife with sexism” in order to share her view that privately held beliefs strongly shape women’s access to citizenship rights. 17 Ensuring the worth of the equal rights of citizenship for women requires supportive social mores. This private–public dependence led Okin to a conclusion contrary to Brettschneider’s: that the state should withdraw tax-exempt status from the Catholic Church or any religious association “as long as they discriminate against women” in all its most important hiring decisions and in the distribution of institutional power. 18 The question, then, for Brettschneider’s value democracy is why the Church should be immune from democratic persuasion when its position on gender-based leadership roles is “ambiguous” at best.

We can amplify this feminist objection by considering the curious asymmetry among many political liberals in their treatment of race, as opposed to sex, discrimination in familial and associational life. In addition to discussing the Catholic Church’s ban on women from the priesthood, Brettschneider discusses Bob Jones University’s ban on interracial dating. 19 He says “Bob Jones University is a prime illustration of when a religious organization should not be given the tax privileges of non-profit status, because of its opposition to the ideal of free and equal citizenship.” 20 Why does Bob Jones University’s policy trigger democratic persuasion while the Catholic Church’s policy does not?

In spelling out further inquiry that might be pursued in these cases, Brettschneider says we should examine the Catholic Church’s current attitudes toward public officeholders who are women. Yet, he suggests a rosier picture than seems warranted when he says the Church “often celebrates Catholic women who attain high political office,” without providing any evidence. 21 By contrast, in the case of Bob Jones University, Brettschneider says “[a] more conclusive inquiry should ask whether [the university] could have some principled reason for

16 Susan Moller Okin, Political Liberalism, Justice, and Gender, 105 ETHICS 23, 29 n.16 (1994).
17 Susan Moller Okin, Justice and Gender: An Unfinished Debate, 72 FORDHAM L. REV. 1537, 1556 (2004); see also SUSAN MOLLER OKIN, IS MULTICULTURALISM BAD FOR WOMEN? 12-17 (Joshua Cohen et al. eds., 1999).
20 BRETTSCHNEIDER, supra note 1, at 137.
21 Id. at 135.
believing its prohibition on interracial dating was compatible with an endorsement of the legality of interracial marriage,” and he registers his doubt that any such justification is available. Why not apply a similar inquiry to the Catholic Church? Such an inquiry would involve addressing questions such as: What are the principled reasons for banning women from the priesthood, and what do they imply for Catholic women’s attainment of leadership positions both inside and outside the Church? Are those reasons really compatible with respecting women as free and equal citizens? Without hearing more, we are left to ponder why Brettschneider’s theory, like that of many political liberals, appears tougher on race-based, as opposed to sex-based, discrimination.

II. THE NEUTRALIST OBJECTION: DEMOCRATIC “PERSUASION” IS COERCION

Perhaps the most innovative aspect of Brettschneider’s theory of value democracy is the use of the state’s taxing and spending power as a key tool for responding to hateful viewpoints. The innovative claim is not so much that the state should use its subsidy power to respond to the problem of hate speech, but rather that we ought to view the state’s use of subsidies as an instance of democratic persuasion, not coercion. This claim depends, of course, on your conception of coercion and persuasion.

Brettschneider adopts the philosopher Robert Nozick’s influential account of coercion, which he interprets “as the state threatening to impose a sanction or punishment on an individual or group of individuals with the aim of prohibiting a particular action, expression, or holding of a belief.” I wish he had dug a bit deeper into Nozick’s account because there are some problems that it raises for his argument. On Nozick’s account,

\[ P \text{ coerces } Q \text{ if and only if:} \]

1. \( P \) aims to keep \( Q \) from choosing to perform action \( A \);
2. \( P \) communicates a claim to \( Q \);
3. \( P \)’s claim indicates that if \( Q \) performs \( A \), then \( P \) will bring about some consequence that would make \( Q \)’s \( A \)-ing less desirable to \( Q \) than \( Q \)’s not \( A \)-ing;

\[ Id. \text{ at 162.} \]
\[ Id. \text{ at 88.} \]
4. $P$'s claim is credible to $Q$;
5. $Q$ does not do $A$;
6. Part of $Q$'s reason for not doing $A$ is to lessen the likelihood that $P$ will bring about the consequence announced in (3).

On Nozick’s definition, the law criminalizing murder is a clear instance of coercion. An individual may choose not to commit murder because the state has made clear that it will impose harsh criminal penalties against him if he does so. The state also engages in coercion if it aims to prevent a group from associating and speaking out by threatening imprisonment or a fine as a consequence. As Brettschneider explains, “State coercion is employed in an attempt to deny the ability to make a choice. By contrast, financial inducements like pure persuasion, clearly attempt to convince citizens to make a particular choice, but they do not deny the citizen the right to reject it.” So when the state decides to withdraw state funding from a group espousing hateful viewpoints, the state is engaging in persuasion, not coercion.

Upon closer examination, however, democratic “persuasion” by threat of withholding state funding does appear to be an instance of coercion on Nozick’s definition. The state aims to keep hateful groups from choosing to perform a particular action (holding and espousing the hateful viewpoint). The state communicates this claim to the group, indicating that if the group continues to hold and espouse the hateful viewpoint, the state will bring about the consequence of withdrawing state funding from the group. To resist this conclusion, Brettschneider needs to stipulate a narrower category of “consequences” that the state can bring about in order to make it less desirable for hateful groups to hold hateful viewpoints. He attempts something like this when he interprets Nozick’s view in terms of imposing “a sanction or punishment.” Similarly, when he says there are “means-based” limits on democratic persuasion, he suggests the state cannot use coercion to prohibit expression, by which he seems to mean direct infringement on the basic rights of free speech, association, and conscience. But Brettschneider’s argument is at odds with Nozick’s view of

---


25 BRETTTSCHNEIDER, supra note 1, at 112.

26 Id. at 87.
coercion which notably diverged from prior accounts by excluding direct use of force and associating coercion only with proposals or conditional threats and by requiring that a judgment about coercion refer to facts about the coercee’s psychology. If we inquire into the psychology of an association faced with the state’s threat of withdrawing state subsidy, it seems reasonable to think it would view the state’s threat as an instance of coercion.

Brettschneider might respond to this objection by saying he adopts the spirit, not the letter, of Nozick’s account. So let me raise two additional arguments to cast doubt on his claim that the state’s use of its spending power is mere persuasion, not coercion.

The first has to do with the distinction between a liberty and the “worth” of that liberty. Rawls makes this distinction in presenting his two principles of justice to show how the first principle, equal basic liberties, must be coupled with principles that ensure some minimum of material resources in order for citizens to exercise their liberties. Brettschneider similarly says that to ensure that the right of free expression has real worth, the state must ensure “the minimum means for citizens to be able to resist democratic persuasion.”

Under certain circumstances, the withdrawal of state funding from groups espousing hateful viewpoints is tantamount to pulling the “minimum means” out from under them. This is ultimately an empirical question, but such cases may be more widespread than Brettschneider seems to assume.

Consider the case of the Christian Legal Society (CLS) at U.C. Hastings College of Law. The public university denied funding to the student group on the grounds that CLS held discriminatory beliefs regarding homosexuality, but Hastings continued to permit the group to meet and speak on campus. Brettschneider endorses the U.S. Supreme Court’s decision that the university did not violate the right of free speech when it refused to subsidize the group, but he disagrees with the Court’s reasoning. The Court focused on which side respected the viewpoint neutrality requirement in funding. Brettschneider argues that the appropriate standard is the idea of free and equal citizenship, which is not viewpoint-neutral. As he puts it, there is no “entitlement for a particular viewpoint to be successful” and

---

27 Id. at 113.
“that hateful or discriminatory organizations have no right to be subsidized by the state.”\textsuperscript{29}

To raise some questions about the “worth” of the rights of free speech and association, let us stylize the example. What if CLS had been the only Christian student organization at the school and it would have ceased to exist as an organization without university support? Under such circumstances, in which state funding is necessary to ensure the minimum worth of basic liberties of Christian students, should CLS receive funding? From the neutralists’ perspective, withholding state subsidies under such circumstances would be objectionably coercive. But if Brettschneider’s approach were to permit state subsidy in such circumstances, would he thereby throw the ideal of free and equal citizenship under the bus, revealing just how extreme his commitment to robust speech protections is? The neutralists would cheer as the prohibitionists would recoil in horror.

My second argument aimed at casting doubt on Brettschneider’s claim that the state’s use of subsidies is mere persuasion is inspired by John Stuart Mill’s account of coercion. Mill’s account is more expansive than most contemporary accounts. He links coercion with the state’s power to punish lawbreakers, but he also suggests a wider range of tools, beyond the use or threat of force, with which agents can exercise constraining power over others. Coercion involves not only the means of “physical force in the form of legal penalties” but also “the moral coercion of public opinion.”\textsuperscript{30} Mill suggests that the power of social stigma is more potent than the use or threat of physical force:

For a long time past, the chief mischief of the legal penalties is that they strengthen the social stigma. It is that stigma which is really effective . . . . In respect to all persons but those whose pecuniary circumstances make them independent of the good will of other people, opinion, on this subject, is as efficacious as law . . . . Our merely social intolerance kills no one, roots out no opinions, but induces men to disguise them or to abstain from any active effort for their diffusion.\textsuperscript{31}

What Brettschneider calls “persuasion” is coercion on Mill’s account. Even if you reject Mill’s more expansive view of coercion, after re-reading these passages of Mill, it is hard to deny that the state’s use of social stigma and appeal to public opinion

\textsuperscript{29} Brettschneider, supra note 1, at 112, 114.


\textsuperscript{31} Id. at 30-31.
can be a more potent threat to individual liberty than legal penalties. Brettschneider explicitly draws on Mill as an inspiration for extending the powers of persuasion from the citizenry to the state itself,32 but it is important to remember that Mill himself favored an active citizenry, not the state, taking the lead in persuading fellow citizens to transform norms and practices. Viewed through this Millian lens, Brettschneider’s “transformative theory,” in which “the state seeks to transform religious belief” (and other personal beliefs) incompatible with free and equal citizenship, sounds more threatening to individual liberty than the use of direct bans and penalties.33

CONCLUSION: WALKING THE LIBERAL TIGHTROPE

Perhaps the greatest virtue of Brettschneider’s position is the source of its greatest weakness: he paves a middle path between the extremes of the Hateful Society and the Invasive State, but in doing so, he has given us an approach riddled by an uneasy tension between respecting liberty and promoting equality.

Both sides have reasons to fear that he gives with one hand what he takes away with the other. To neutralists, he appears to embrace robust rights protections, only to endanger them through democratic “persuasion” aimed at changing personal beliefs and practices. Religious believers may favor old-fashioned neutralists to Brettschneider's approach when they realize that the latter would devote tax dollars toward “attempting to transform” any discriminatory beliefs they may hold! To prohibitionists, including some feminists, Brettschneider advocates the state’s “active” role in criticizing and even condemning hateful viewpoints, only to permit groups espousing hateful viewpoints, such as the Westboro Baptist Church, to speak out and stage protests.

This tension between respecting liberty and promoting equality is endemic to liberalism. It is the tightrope liberals have to walk. No self-respecting liberal can afford to give up on either. While there are moments when Brettschneider’s approach seems to lean more to one side or the other, his book is a serious effort at reconciling these twin values and deserves to be widely read and discussed.

32 BRETTSSCHNEIDER, supra note 1, at 13.
33 Id. at 155, 157-58.
Democratic Persuasion and Freedom of Speech

A RESPONSE TO FOUR CRITICS AND TWO ALLIES

Corey Brettschneider†

INTRODUCTION

Liberalism demands robust rights to free expression. In American jurisprudence, the liberal state is bound by one of the world’s strictest rules protecting free speech, the doctrine of “viewpoint neutrality.”† This doctrine requires the state to protect all speech regardless of beliefs or political content. Viewpoint neutrality is commonly thought to be based on a neutralist theory of liberal democracy that requires the state not to favor any set of values.‡

Feminist critics of neutralist liberalism resist what they regard as an overemphasis on unlimited freedom of expression

† I want to thank the distinguished participants in this symposium for their outstanding and thoughtful contributions: Frank Michelman, Steve Calabresi, Josiah Ober, Andrew Koppelman, Robin West, and Sarah Song. Michelman and Calabresi presented versions of their papers at the annual meeting of the Law and Society Association, organized by Sonu Bedi, where they were joined by Annie Stilz. I thank Sonu and Annie for their exceptional contributions to the Law and Society Association roundtable on the book. John McCormick also organized a superb panel at the American Political Science Association where Song, Koppelman, and Ober were joined by Eric Posner. John Moore, the editor of the Brooklyn Law Review, kindly and skillfully saw this symposium through to completion. I thank him and Nelson Tebbe for their work in organizing it. For excellent substantive comments and research assistance, I thank Minh Ly.

rights. Catharine MacKinnon, for instance, claims that free speech and the value of equality are on a “collision course.” According to these critics, while rights to free speech matter, rights to equality are equally, if not more, important and should sometimes limit free speech rights when the two conflict. Almost all democracies outside of the United States follow this “prohibitionist” approach. They limit free speech when hateful expression attacks equal respect for minorities or the value of democracy itself. The prohibitionist approach tries to correct the alleged inability of liberalism to defend the core values of democracy.

In my book, When the State Speaks, I offer an account of liberal democracy that combines the neutralists’ protection of rights with the feminists’ and prohibitionists’ concern for the equal status of citizens. I call this third view of liberalism and free speech “value democracy.” It grounds viewpoint neutrality on an ideal of free and equal citizenship. On my account, the state should be neutral in protecting the right to express all viewpoints. But it should not be neutral in the values that it supports and expresses. Value democracy thus embraces viewpoint neutrality in protecting the right to free expression of all beliefs, but rejects neutralism as a theory of what the state should say. The state must favor some substantive values, namely the ideal that all citizens should be treated as free and equal.

It is not enough, however, to recognize this commitment in the abstract. Liberal democracy must also find a way to protect the substantive values on which it is based. Otherwise, it will run into the problem that neutralist liberals face of “being unable to take their own side in an argument” when the free and equal status of women, minorities, gays, and other citizens is attacked. In these cases, liberal democracy must be able to articulate the “reasons for rights” that justify respecting free speech rights and viewpoint neutrality in the first place.


4 See Erik Bleich, The Freedom To Be Racist?: How the United States and Europe Struggle to Preserve Freedom and Combat Racism 97-105 (2011); see also Adam Liptak, Outside U.S., Hate Speech Can Be Costly: Rejecting the Sweep of the First Amendment, N.Y. Times, June 12, 2008, at A1 (describing differences in the way the United States and other countries, such as Canada and Germany, treat potentially offensive speech).

When the State Speaks thus offers an account of “democratic persuasion” that requires the state to protect all viewpoints from coercion or prohibition. But when it “speaks” in statements by public officials, when it educates, when it uses its spending power, and when it confers the tax privileges of non-profit status, the state must affirmatively take the side of upholding free and equal citizenship. Democratic persuasion, I argue, is not just something that the state is permitted to do. It is a matter of political obligation. Our constitutional jurisprudence, including the doctrine of viewpoint neutrality, must be tailored to permit the state to pursue its duty of democratic persuasion. At the same time, democratic persuasion places limits on state speech. It prohibits the state from speaking in ways that undermine its commitment to the values of freedom and equality.

The proper place for viewpoint neutrality, I argue, is in preventing government coercion or censorship of viewpoints. Citizens should be allowed to hear and endorse all viewpoints, even hateful ones. While I think threats and speech that might incite imminent violence can be prohibited under the First Amendment, generalized viewpoints cannot be banned. I argue, however, that the state should not be viewpoint neutral in its own expression. The state should protect free speech out of respect for the freedom and equality of citizens. Citizens are free and equal in having the capacity to debate and decide on matters of personal and political principle. The state should find a way not only to uphold free speech, but also to defend the democratic values that justify protecting free speech in the first place. For example, the state has an obligation to advance civil rights through education and public holidays. We rightly dedicate a holiday to Martin Luther King, not to the southern segregationist Bull Conner. Likewise, the public schools are justified in teaching students racial equality.

An even stronger measure that the state should take is to use its spending power to advance democratic values. I therefore defend the IRS decision to deny the subsidies that come with non-profit 501(c)(3) status to Bob Jones University, which the Supreme Court upheld in Bob Jones v. United States. The IRS already requires that, for non-profits to receive the subsidies of tax-exemption, they must have a “public benefit.” That is, such organizations must provide services to the public that offset the cost of the tax-exemption.

In the book, I argue that the IRS should make its public benefit more specific, and explain that being a hate group is inconsistent with having a public benefit. This clarification of the meaning of public benefit should be made by Congress rewriting the 501(c)(3) statute. The IRS should thus be required by law to deny the tax subsidies of 501(c)(3) to hate groups that directly oppose the democratic values of free and equal citizenship. When the state uses its spending powers, it should promote democratic values and not be bound by viewpoint neutrality.

While the Court’s decision in *Bob Jones* is consistent with my view, the Court has since moved in the wrong direction in expanding viewpoint neutrality in other cases that concern state spending. In *Christian Legal Society v. Martinez*, the Supreme Court held that it was constitutional for Hastings Law School to withdraw funds from a student group that discriminated against gay students. In her majority decision, Justice Ginsburg claimed that the funding policy of requiring non-discrimination in admissions for student groups was consistent with the state’s viewpoint neutrality. She wrote that the policy was based on an ideal of toleration, which she claimed was a neutral value. Although I agree with the Court’s result in *Christian Legal Society*, I suggest its reasoning wrongly tried to show that requiring non-discrimination in admissions is consistent with the doctrine of viewpoint neutrality. Non-discrimination and toleration are non-neutral viewpoints. Their non-neutrality can be seen in how those viewpoints are attacked by discriminatory groups. The problem with the Court’s reasoning was that it assumed that the state must be viewpoint neutral in its expression. I argue in the book that non-discrimination and toleration are non-neutral viewpoints that the state should advance through its own speech. In sum, while viewpoint neutrality has a place in limiting government coercion, it should not limit the state’s ability to promote democratic values.

It should be emphasized, however, that democratic persuasion places limits on what the state can say. Democratic persuasion prohibits the state from speaking in ways that undermine the ideal of free and equal citizenship. For example, it would be wrong for the president, legislators, and the courts to speak in favor of racial discrimination. I therefore favor an expansive reading of the equal protection clause and the

---

8 130 S. Ct. 2971 (2010).
establishment clause to limit some forms of state speech. For instance, I argue that *Rust v. Sullivan*\(^9\) was wrongly decided, because the state did not have the right to deny information to women about to their rights to an abortion as guaranteed by *Roe v. Wade* and *Planned Parenthood v. Casey*. I also endorse Michael Dorf’s view that it would not be constitutional for states to fly the Dixie flag, because it would be a form of state speech on behalf of the discriminatory values that flag represents.\(^{10}\)

Several of the essays in this symposium attempt to push my view toward one of the opposing poles of neutralism or prohibitionism. On the more neutralist side, Steve Calabresi worries that I have abandoned traditional liberal commitments to respect the independence and autonomy of religious citizens.\(^{11}\) Although he endorses much of value democracy, he wants to see greater reticence in democratic persuasion to limit its application to religious organizations. While Calabresi accepts a central role in liberal democracy for a more reserved form of democratic persuasion, Andrew Koppelman denies that the state has a duty to pursue democratic persuasion. At most for him, it is a “second best” set of tools for the state to use in some circumstances, especially given his skepticism about whether it is necessary or effective.\(^{12}\)

Robin West\(^{13}\) and Sarah Song\(^{14}\) push in the opposite direction of prohibitionism. On their view, I am right to seek to persuade citizens to change hateful viewpoints. They share my commitment to a political theory that challenges discrimination and seeks to promote ideals of equality in the family and civil society. But they worry that I have not gone far enough in my account. West believes that democratic persuasion does not act strongly enough to protect equality.\(^{15}\) Song questions how I might respond to critics, like Susan Okin and other liberal feminists, who argue that the state should promote more extensive changes in civil society than is permissible in democratic persuasion.\(^{16}\)

---

\(^9\) 500 U.S. 173 (1991)
\(^{12}\) Andrew Koppelman, *You’re All Individuals: Brettschneider on Free Speech*, 79 BROOK. L. REV. 1023 (2014)
\(^{15}\) West, supra note 13.
\(^{16}\) Song, supra note 14
In response to these critics, I suggest that value democracy and democratic persuasion offer a third way forward in thinking about the role of values in liberalism. I attempt to show that value democracy strikes what Calabresi aptly calls “the golden mean” between neutralism and prohibitionism. My defense and elaboration of value democracy and democratic persuasion are aided by excellent essays from Frank Michelman and Josiah Ober. I begin by highlighting how Michelman’s essay underscores the strengths of my view that a legitimate state has an obligation to engage in democratic persuasion. His focus on legitimacy offers an important reply to theorists who want to see greater reticence in democratic persuasion, including Calabresi and Koppelman. Ober’s essay responds powerfully to concerns about the effectiveness of democratic persuasion and its respect for citizens. Finally, I address West’s and Song’s calls for a more radical form of democratic persuasion, suggesting that the cautionary arguments from Calabresi and Koppelman can be used to push back against West and Song.

I. MICHELMAN ON DEMOCRATIC PERSUASION AND LEGITIMACY

Frank Michelman explains both the uniqueness of my view and its place in the liberal tradition. He focuses on my claim that the duty of democratic persuasion stems from an ideal of legitimacy. Neutralist liberals might resist democratic persuasion on the grounds that the state should be neutral toward different values. However, the point that Michelman rightly emphasizes from When the State Speaks is that legitimacy is more fundamental to liberalism than neutrality. If the state is to be neutral in some cases, that neutrality must be justified in terms of legitimacy. What makes the state legitimate is its commitment to the status of all citizens as free and equal.

In my theory of value democracy, the state’s legitimate power must be justified by its respect for an ideal of free and equal citizenship. In When the State Speaks and in my first book, Democratic Rights: The Substance of Self-Government, I argued that state coercion must always be made consistent

---

19 Michelman, supra note 17.
with this ideal.\textsuperscript{20} As Michelman explains, my “thin” conception of legitimacy requires rights-based limits on government action.\textsuperscript{21} These limits include equal protection, robust free speech protections, and a broad right to privacy. As Michelman explains, I reject the notions of legitimacy based on the empirical question of whether there is widespread compliance with the state. Legitimacy in value democracy is instead a normative concept regarding whether the state has the right to rule or coercively enforce its laws.

My claim that the requirements of legitimacy limit coercion places my theory of value democracy squarely in the tradition of liberal democracy. But \textit{When the State Speaks} makes the case for an even broader conception of legitimacy than in other liberal theories. At times, this commitment requires the state to be neutral, as in its protection of the right of free speech for all viewpoints. The state does this to respect citizens as free, or as able to exercise their moral powers to debate, choose, and pursue conceptions of the good life and of justice. But at other times, the commitment to the status of citizens as free and equal might require the state not to be neutral. When confronted by hateful expression, the legitimate state cannot be neutral in its own speech. The legitimate state must pursue democratic persuasion to defend the free and equal status of the women, minorities, gays, and other citizens facing attack.

In addition to ensuring that the values of free and equal citizenship limit coercion, the legitimate state should use democratic persuasion to promote these values and work to have them adopted by citizens. Michelman anticipates the objection that legitimacy does not depend on democratic persuasion alone.\textsuperscript{22} Indeed, I agree that perhaps the most fundamental part of legitimacy is the content of the laws and whether it complies with the demand to respect rights. But, as Michelman reports, I argue that democratic persuasion might add to a government’s legitimacy, helping to move it closer to what he regards as the proper threshold for the right to rule. While Michelman views legitimacy as being “categorical”—the state either has it or lacks it entirely—liberal theorists who view legitimacy as being “scalar” or a matter of degree might endorse the idea that democratic persuasion contributes to a

\textsuperscript{20} COREY BRETTSCHNEIDER, DEMOCRATIC RIGHTS: THE SUBSTANCE OF SELF-GOVERNMENT (2007) [hereinafter BRETTSCHNEIDER, DEMOCRATIC RIGHTS].

\textsuperscript{21} Michelman, \textit{supra} note 17.

\textsuperscript{22} \textit{Id.}
state’s legitimacy. When hate speech attacks citizens, democratic persuasion defends their status as free and equal.

An important point raised by Michelman is that, in keeping with the fundamentality of legitimacy, conceptions of neutrality, specifically an account of viewpoint neutrality, must themselves be grounded in a liberal theory of legitimacy. Building on this point, When the State Speaks explains that there are three reasons why the legitimate state has an obligation to engage in democratic persuasion. Reviewing these reasons will set up my argument later in this essay against the belief that liberal reticence ought to override democratic persuasion.

First, a liberal democracy with widespread illiberal beliefs would have questionable stability. Without democratic persuasion, government decisions might respect free and equal citizenship, but these decisions would be at odds with the views of the populace. As Steve Calabresi points out, such a society would resemble Weimar Germany, where democracy failed because it was not supported by the widespread endorsement of democratic principles. This argument for stability is “instrumentalist” in my terms, or “consequentialist” as Michelman says. The legitimate state ought to pursue democratic persuasion to prevent the decline of liberal democracy. But Michelman correctly notes that this instrumental argument might not satisfy more deontological theorists who want an explanation of how democratic persuasion shows respect for individual persons.

Michelman thus highlights the deontological character of my second and third arguments for why the legitimate state ought to pursue democratic persuasion. The second is an argument from transparency for a state role in articulating and defending the reasons for rights. Citizens have an entitlement to be given a justification for the laws that bind them, out of respect for their autonomy or their capacity to reason about justice. This notion of transparency builds on the widely accepted principle that laws and the reasons for them must be promulgated publicly. Value democracy extends the transparency requirement from the law to the basic values of liberal democracy.

---

23 Id.
24 For an analysis of the state’s duty to pursue democratic persuasion, based on the arguments from stability, transparency, interconnection, and public trust, see Brettschneider, When the State Speaks, supra note 5. The kind of stability that I am invoking here is what Rawls calls “stability for the right reasons” in John Rawls, Political Liberalism xi, 391, 495 (Columbia Univ. Press expanded ed., 2005).
25 Calabresi, supra note 11.
26 Michelman, supra note 17.
27 Id.
The legitimate state must publicly justify and promulgate the reasons for democratic values. This argument does not mean, as Koppelman seems to suggest, that the legitimate state needs to rebut every random disagreement with the core values of free and equal citizenship. Rather, as Michelman explains it, I argue that there is a generalized obligation of the state to make clear the democratic values on which its legitimacy is based. This duty applies most strongly when hate speech or discriminatory beliefs are not random but affect people’s standing as free and equal citizens. At times democratic persuasion might take the form of broadly promulgating democratic values. At other times, when necessary, it might mean criticizing and giving reasons to reject hate speech and discriminatory beliefs.

A third set of reasons for democratic persuasion as a requirement of legitimacy concerns what I call the arguments from “interconnection” and “public trust.” According to the argument from public trust, government officials who make and implement policy have a great deal of discretion. The laws might conform formally to democratic values, but if some officials do not endorse these values, they may act in ways that undermine the freedom and equality of citizens. Democratic persuasion is needed to promote greater acceptance of democratic values among citizens, including current and future public officials. The ideals of free and equal citizenship are independent of what any public official says. If there is an official, such as a state governor, with discriminatory beliefs, my theory would require other officials, including officials from other U.S. states and the federal government, to criticize those discriminatory beliefs. Indeed, on my reading of the Constitution, hateful state speech by public officials would be prohibited by the equal protection clause.

The argument from public trust draws attention to the “vertical” relations between citizens and the state. The argument from interconnection focuses instead on the importance of the “horizontal” connection between citizens. The problem, which Robin West highlights, is that citizens might act in civil society to attack the free and equal status of women, minorities, and gays. I argue in the book that citizens can undermine equal status by advocating hateful beliefs, or by making discriminatory decisions about whom to hire, promote, and mentor. The law’s formal respect for free and equal

---

28 For an excellent study on the persistence of racism in modern America, see Elizabeth Anderson, The Imperative of Integration (2010). I would argue that Anderson’s study shows the continuing need for democratic persuasion.

29 West, supra note 13.
citizenship is insufficient to counter this problem. The argument from interconnection is that democratic persuasion is needed to promote more widespread acceptance of democratic values in civil society. Citizens have an obligation to treat each other in accordance with the democratic values of free and equal citizenship, and the state has an obligation to use its expressive capacities to encourage citizens to do so. Otherwise, these values will be undermined in the way that citizens treat each other. The argument from interconnection attempts to redress claims about what is called “structural racism” that takes place outside the formal boundaries of law.

As Justice Harlan argued in his famous dissents in the Civil Rights Cases and in Plessy v. Ferguson, citizenship must permeate the culture and cannot remain a merely formal ideal if we are to live in a true democracy in which people enjoy equal status. I agree with Harlan and argue that civil rights laws are important in prohibiting discrimination in the workplace and universities. The Civil Rights Act of 1964 is properly understood as being constitutional and compatible with free association rights. At the same time, I endorse current First Amendment jurisprudence, which distinguishes between the legitimate regulation of the workplace, schools, and public accommodations, where discrimination can be prohibited, versus private organizations, which have rights of free association. The Court has ruled that the right of free association and free speech protect private organizations and religious groups, such as the Boy Scouts of America and the Westboro Baptist Church, in their freedom to determine who is a member of their organization as well as what message they wish to embrace. Such groups, unlike businesses or public accommodations, enjoy the free speech right to promote illiberal ideals. That does not mean the state must stand idly by.

Even when free speech rights, privacy, and free association rights protect private groups from anti-discrimination laws, the liberal state should still address discriminatory and hateful viewpoints as being publicly relevant. Democratic persuasion aims to promote more widespread respect for the free and equal status of citizens, even when other rights keep the state from promoting and protecting such beliefs through coercive law. As Michelman writes, a “crucial component” of my theory is that citizens should affirm democratic principles.

30 Plessy v. Ferguson, 163 U.S. 537, 552-64 (1896) (Harlan, J., dissenting); The Civil Rights Cases, 109 U.S. 3, 26 (1883) (Harlan, J., dissenting).
Citizens as individuals move toward realization of an ideal that Michelman, following Rawls, calls “full autonomy” when the state they jointly support speaks up on behalf of civic obligations of respect for all as free and equal.

I take these deontological arguments to suggest that democratic persuasion does not disrespect citizens by disagreeing with them. Rather it respects the way citizens should be treated in democracy and the way they should treat others. Democratic persuasion criticizes the viewpoints of people who advocate racial and sexual discrimination, but it respects their status as citizens. This stance of criticizing the viewpoint but respecting the person and his or her ability to reason can be understood by a helpful analogy. If I hold a mistaken view, such as the earth being flat, it does not disrespect me to argue against my view and to provide evidence for the earth being round. Rather, it recognizes my ability to reason. Similarly, it does not disrespect citizens to argue against their hateful or discriminatory beliefs. If they use their free speech rights to attack the free and equal status of other citizens, they are undermining the very principle of legitimacy that gives them the right in the first place. To point this out is not to disrespect them. Rather, it upholds the status that all citizens should have in a legitimate democracy. What is crucial here when it comes to state speech is that democratic persuasion must not insult individuals, but rather give reasons that appeal to values central to democratic legitimacy. A state that promulgates the justification for its basic laws and rights engages in reason-giving speech that is respectful, not insulting, to citizens.

II. IS DEMOCRATIC PERSUASION TOO AGGRESSIVE? A RESPONSE TO CALABRESI

Recognizing the three reasons why democratic persuasion is required as a matter of legitimacy, and that legitimacy is prior to neutrality in liberalism, I will respond to several criticisms raised about my view. I begin with the concern that democratic persuasion is too aggressive in criticizing and seeking to transform views that oppose free and equal citizenship. The worry is that democratic persuasion strays too far from the traditional liberal distinction between the public and the private.

Steven Calabresi believes that I have successfully identified the structure of First Amendment free speech
He agrees that the grounding of viewpoint neutrality must be value based. He endorses a limited conception of democratic persuasion that includes a role for the state in defending the core values of freedom and equality. But Calabresi is worried that I overextend democratic persuasion in two areas that upset the traditional liberal reticence to intervene in the private sphere.  

First, he worries about my claim that the state should not extend the tax privileges of 501(c)(3) status to discriminatory groups, such as the Boy Scouts. He argues that nonprofits should be off-limits from government criticism because the diversity of those organizations supports a robust liberal society. He suggests that extending democratic persuasion to decisions about tax exemption will lead to the politicization of the IRS and even possibly a partisan cultural war, with conservatives seeking to revoke the non-profit status of pro-gay rights groups.

I agree with Calabresi in respecting the right of all groups in civil society to organize around any principle or viewpoint they wish. I go so far as to defend the right of groups such as the Ku Klux Klan and the American Nazi party to march and organize. Nothing in my proposal suggests banning discriminatory groups. To the contrary, the right to hold hateful viewpoints is encompassed by democratic citizenship and in particular the requirement that everyone should be entitled to have and decide upon their own beliefs, out of respect for their autonomy. I endorse Meiklejohn’s suggestion that democratic autonomy demands a viewpoint neutral protection of the right of free speech against government prohibitions of expression. Indeed, I am with the most liberal of free speech theorists in that I suggest my approach is compatible with the Supreme Court’s holding in Boy Scouts of America v. Dale, defending the right of private associations to discriminate in their membership policies based on sexual orientation. These rights protect a realm of civil society from coercive control.

31 Calabresi, supra note 11.
32 Id.
33 Id.
34 Id.
35 The Court allowed the Ku Klux Klan to burn crosses if there was no intent to intimidate in Virginia v. Black, 538 U.S. 343, 363 (2003). The Court held that the American Nazi Party had a free speech right to hold a march in National Socialist Party of America v. Village of Skokie, 432 U.S. 43, 43-44 (1977). On my defense of free speech for hateful viewpoints that are not threats, see BRETTSCHNEIDER, WHEN THE STATE SPEAKS, supra note 5, at 74-75.
But a right to free association does not entail either a right to be publicly subsidized or a right to be free from criticism by the state. Organizations that discriminate or advocate hateful beliefs should not enjoy the tax privileges that come with 501(c)(3) status. That status comes with publicly supported tax benefits, including the deductibility of donations. Status under 501(c)(3) is, in effect, an indirect subsidy to these groups. Such subsidies should be reserved for groups that pursue a public good.\textsuperscript{37} Discrimination groups such as the American Nazi Party, the Klan, or the Westboro Baptist Church (which has declared that God hates gays) are not pursuing, and are in fact attacking, the public good of free and equal citizenship. We can respect the right of discriminatory groups to exist without endorsing or publicly supporting them by revoking their tax advantages. Denying these subsidies does not deny pluralism in civil society. To be clear, the proposal is to end a privilege or subsidy that discriminatory groups receive. This would only lead to their paying the same taxes that ordinary citizens pay. Departing from that tax rate should require extra justification, as Justice Rehnquist recognized in \textit{Regan}, because subsidies place additional burdens on other citizens, who must make up for the lost tax revenue. Of course, these groups have the right to continue to exist without subsidy. In fact, \textit{Bob Jones} continues to operate without 501(c)(3) status.

In addition to his concern about pluralism, Calabresi also worries that my proposal may trigger a culture war over who receives 501(c)(3) status.\textsuperscript{38} Much of Calabresi's concern here could be answered by detailing how legislation could be written that would clarify the terms upon which non-profit status might be revoked for discriminatory groups. The potential for a culture war exists in the status quo because of the vagueness of the current statute. In my view, there would

\textsuperscript{37} \textit{Regan} v. Taxation with Representation, 461 U.S. 540, 550 (1983). Justice Rehnquist wrote in his unanimous decision that tax exemptions were forms of state subsidy, and that they should be reserved for groups that "promote the public welfare." \textit{Id.} at 544 (footnotes omitted) ("Both tax exemptions and tax deductibility are a form of subsidy that is administered through the tax system. A tax exemption has much the same effect as a cash grant to the organization of the amount of tax it would have to pay on its income. Deductible contributions are similar to cash grants of the amount of a portion of the individual's contributions. The system Congress has enacted provides this kind of subsidy to nonprofit civic welfare organizations generally, and an additional subsidy to those charitable organizations that do not engage in substantial lobbying. In short, Congress chose not to subsidize lobbying as extensively as it chose to subsidize other activities that nonprofit organizations undertake to promote the public welfare.").

\textsuperscript{38} Calabresi, \textit{supra} note 11.
be a way of codifying the non-profit proposal to provide the IRS with less discretion than they already have under the laws governing 501(c)(3). The 501(c)(3) requirements currently include a need to show “charitable purpose” or, as the Court has understood it, promotion of the “public good.” I argue that the statute might be amended to clarify that opposition to ideals of equal citizenship constitutes a failure to meet this public good requirement. Codifying the ideal in this way would offer a significantly clearer standard for the IRS in defining what constitutes a “public good.” Using respect for free and equal citizenship as the standard would posit a clear rule that the courts could use to counter government abuse. My proposal would thus reduce the risks of government abuse compared to the current, vague standard of groups having to promote the “public good” to qualify for non-profit tax-exemptions.

This proposal also might answer some of Calabresi’s worry that conservatives might begin to use the statute to revoke the status of Ivy League universities that are perceived as being too liberal. As Calabresi recognizes, this would be a blatant abuse of power. Clarifying the meaning of charitable purpose and public good would correct against that sort of abuse. Courts could be entrusted to push back on such abuse by relying on a reformed 501(c)(3) statute that clarifies the “charitable purpose” or public good requirement.

Although Calabresi wants to interpret 501(c)(3) along neutralist lines, all three legitimacy arguments for democratic persuasion support using free and equal citizenship as the standard for groups to receive the privileges of tax-exemption. On consequentialist grounds, the proposal would promote free and equal citizenship without denying rights. The clarification of the meaning of “charitable purpose” and “public good” would promulgate and make transparent the reasons for rights. The policy would also address concerns about interconnection and public trust by using subsidies to promote a culture based on respect for free and equal citizenship. The proposal helps to

---

39 See Regan v. Taxation with Representation, 461 U.S. 540, 544 (1983), and Bob Jones University v. United States, 461 U.S. 574, 587-88 (1983), both interpreting the text of 501(c)(3) as requiring an organization receiving tax exemption to have a “charitable purpose” or to “promote the public welfare.” In Bob Jones, the Court held that the Internal Revenue Service could discontinue 501(c)(3) tax status to Bob Jones University because the university’s policy of banning interracial dating was at odds with the public purpose of educational institutions. Id. at 605. The IRS made the right decision on my view, given Bob Jones’s explicit opposition to rights to interracial marriage. Opposition to rights of interracial marriage does not promote the public welfare, since it undermines the ideal of equal citizenship.
transform civil society and promote core values of legitimacy without violating the rights to free associations and speech of civil society groups.

A separate worry raised by Calabresi is that I do not exempt religious organizations from being subject to democratic persuasion when they oppose basic values of freedom and equality. In his view, religious organizations themselves are protected not only by the guarantee of free exercise but also on grounds of equal protection. The concern is that subjecting religious groups to democratic persuasion might violate both guarantees.

A test case for Calabresi, however, is *Bob Jones University v. United States*. In *Bob Jones* the IRS revoked the 501(c)(3) status of a racist university with a religious affiliation. The Court upheld that equal protection and free speech did not entitle the university to tax exempt status. The Court ruled that the IRS could discontinue tax subsidies to Bob Jones because the university banned interracial relationships. This opinion rightly made it possible for the state to engage in democratic persuasion in its tax exemption policies.

Calabresi and I both agree that the *Bob Jones* case was correctly decided, and that tax exemptions should not have been continued for a university that banned interracial relationships. We differ, however, on the analysis of the case. Calabresi argues that the university was not actually a religion, and so tax exemption could be conditioned on non-discrimination. By contrast, I am willing to grant that Bob Jones University is a religious organization, but I conclude that the religious nature of a discriminatory group does not automatically exempt it from criticism. Regardless of the details of this case, it seems clear that racist, sexist, and homophobic religions have attacked the equal status of minorities, women, and gays under the law. For example, the Westboro Baptist Church has picketed funerals for gay murder victims, including Matthew Shepard, declaring that they deserved to die. The Westboro Baptist church is a religious organization. But its core doctrine advocates capital punishment for gays. We therefore cannot escape the challenge here of whether they should receive 501(c)(3) status by denying they are a religion, or by denying, as the Kansas Court of Appeals did, that the church’s core activity of protesting military and gay funerals is central to their religious

---

40 Calabresi, *supra* note 11.
41 *Bob Jones Univ.* 461 U.S. at 605.
42 Calabresi, *supra* note 11.
mission. Hateful religious groups are still religions. While religious hate groups are entitled to protection from coercive interventions that ban them from speaking or organizing, they should not be exempted from democratic persuasion.

The argument for subjecting religious groups to democratic persuasion is grounded in an ideal of pluralism. As Michelman clarifies, the conception of pluralism for a legitimate democracy should guarantee freedom and equality for all. These are the democratic values that the state must defend against its opponents, religious or otherwise. The ideal of equal status is the value that best explains and justifies the protection of religious freedom in the first place. Free exercise and the ban on establishment are based in a deeper concern to offer people equal respect. This is the reason we allow them to exercise their religion as they wish. Without such a concern we would be hard pressed to explain why some religions must be limited when they seek to restrict the rights of other religions. Religious groups that want to curtail the rights of others have a right to speak, but they should be criticized for attacking the premises of the fundamental freedoms that they enjoy. Just as free speech is grounded in an ideal of free and equal citizenship, so too is the guarantee of free exercise.

Democratic legitimacy is based on protecting equal respect for all citizens. But for equal respect to be secured, clearly discriminatory or hateful religious groups should not be exempt from democratic persuasion, as Calabresi suggests. Democratic persuasion should apply to discriminatory religious groups in the consequentialist hope of bringing about the kind of pluralistic society in which all religions are respected, and in which there is transparency about how the protection of religious freedom is grounded in deeper values of legitimacy.

III. IS DEMOCRATIC PERSUASION EFFECTIVE AND RESPECTFUL? A RESPONSE TO KOPPELMAN DRAWING ON OBER

I welcome Andrew Koppelman’s commentary as an opportunity to respond to his consequentialist objections and to

---

43 In re Westboro Baptist Church, 189 P.3d 535, 541-42 (Kan. Ct. App. 2008) (holding that the Westboro Baptist Church could not receive a tax deduction for a truck that it used during protests). The Court ruled that the truck was being used for political and not religious expression. For discussion, see Brettschneider, When the State Speaks, supra note 5, at 154, 194 n.17.

44 Michelman, supra note 17.
explain why a purely consequentialist framework does not fully capture the legitimacy arguments for democratic persuasion. Koppelman pushes in the direction of neutralism on the grounds that democratic persuasion is supposedly ineffective and unnecessary. I respond that democratic persuasion is necessary for the state to meet its duty to be transparent about the democratic values that justify its own legitimacy. It is also the most effective means of transforming beliefs while respecting rights.

Part of Koppelman’s concern about effectiveness is based on a basic misunderstanding. He seems to think that democratic persuasion is my only response to racism, sexism, and homophobia. This overlooks the other strong measures that I endorse in the book to defend the free and equal status of citizens. As I write, “[v]alue democracy accepts that the state’s coercive power might be appropriate for protecting rights in certain cases, such as using the police to stop violence or enforcing the Civil Rights Act to uphold non-discrimination in the workplace.” Democratic persuasion should be understood as part of a larger liberal defense of citizens’ rights.

What democratic persuasion contributes to this defense is a way of protecting free and equal citizenship from attacks in a traditionally vulnerable flank. Civil society groups like families and churches are highly influential, but are not covered by anti-discrimination law. They are protected by robust rights to free speech and association that prohibit using coercion to ban hateful expression and discrimination in their membership. I defend liberal rights to free speech and association, but I criticize the purely naturalistic approach that would have the state remain silent in the face of hateful expression. Neutralism would allow a culture of inequality to emerge unchallenged. The status of citizens as free and equal would then be threatened by two problems. First, the problem of public trust is that state officials, such as police officers, who learn discriminatory beliefs in civil society might use their discretion to undermine free and equal citizenship. For example, they might fail to enforce laws against domestic violence or rape. Second, the problem of interconnection is that discriminatory beliefs would not be narrowly contained in civil society, but would undermine free and equal citizenship more broadly. With their discriminatory views unchallenged by the

45 For an example of this misunderstanding, see Koppelman’s claim that “[Brettschneider] presents this and nothing else as the solution to the dilemma he describes.” Koppelman, supra note 12, at 1025.

46 BRETT SCHNEIDER, WHEN THE STATE SPEAKS, supra note 5, at 25.
state, people might act on views that regard minorities, women, and gays as inferior, such as by making prejudiced choices in hiring, promoting, and mentoring. Anti-discrimination laws can help, but these laws can only go so far without widespread endorsement of equality by citizens. For the free and equal status of citizens to be protected, the state should use democratic persuasion to argue against discriminatory beliefs, while respecting the rights of free speech and association.

Koppelman ignores my defense of rights, even though I uphold free speech even more categorically than he does. He supports free speech rights only conditionally on consequentialist grounds. The consequentialist might be open to abandoning rights when faced with the threat of the Hateful Society. Koppelman seems to take this position, writing that free speech defenders need “to be open to the possibility that those grounds are overridden if the consequences are bad enough.” Unlike Koppelman, I make a non-consequentialist argument for free speech based on a legitimate democracy’s respect for the autonomy of citizens, or their capacity to reason about justice and the good.

Koppelman might reply that democratic persuasion is not needed, because discriminatory and hateful beliefs are marginal, like the views of a “single lonely crank” in his memorable language. He claims that “[o]vert racism has been nearly eliminated during the period when the First Amendment was construed to give strong protection to racist speech.” In contrast to Koppelman, Robin West shows in her essay how the contemporary United States still faces continuing harms from racism, sexism, and homophobia. The statistics on hate crimes support West’s argument. According to the FBI, nearly 6,000 hate crimes, including assaults, rapes, and murders, were committed in 2012. To cite just one example, the New York Times reports that last year, a gunman “followed his victim,

---

47 See Koppelman, supra note 12.
48 Koppelman takes a consequentialist approach to free speech and other rights. As he argues, “[a]ny consequentialist case for a given liberty is likely to rest on some combination of optimism (about what will happen absent regulation) and distrust (of the government actors who would regulate the liberty in question). The warrant for either of these will vary from one society to another.” Koppelman, supra note 12, at 1030 (footnotes omitted).
49 Id. at 1029.
50 Id. at 1024.
51 Id. at 1030.
Mark Carson, for several blocks, taunting him with antigay slurs, before killing him. As the Manhattan district attorney, Cyrus Vance Jr., said, “This young man’s tragic death serves as a reminder of the discrimination that many of our family members, co-workers, and friends still face.” These crimes did not take place in a vacuum, but were motivated by hateful beliefs. Even in cases without direct acts of violence, these beliefs teach women, minorities, and gays to hate themselves. These individuals internalize the message of contempt and inferiority. Instead of standing idly by, the state has a duty to engage in democratic persuasion to criticize the message of hate and abuse. As I noted in the last section, this generalized obligation does not, as Michelman correctly clarifies, mean the state is mandated to respond to each solitary “crank.” It is a generalized obligation to promote the ideal of free and equal citizenship when democratic values are undermined by hateful and discriminatory groups.

At this point, Koppelman might admit that democratic persuasion is needed, but question whether state officials should engage in it. He suggests relying entirely on groups in civil society to argue against hateful beliefs. This view frames democratic persuasion as a mutually exclusive alternative to robust discussion among citizens. Josiah Ober and I argue instead that democratic persuasion can promote discussion in civil society.

Ober, who is one of the most insightful theorists of classical rhetoric and politics writing today, explains in his essay how carefully crafted speech by public officials has succeeded in changing minds. Persuasive speech has transformed minds and practices, supporting significant cultural change in favor of equality. Ober gives the example of President Lyndon Johnson, who spoke in favor of the Civil Rights Act of 1964. Engaging in democratic persuasion was essential to convincing Congress, building public support, working with civil society groups, and ultimately passing legislation to protect the equal status of citizens. The state was not silent or neutral, but made persuasive arguments that justified the proposed laws in terms of the values of freedom and equality.

The passage of the Civil Rights Act demonstrates why state speech and civil society are not isolated zones that are separated by a no-man’s-land. There is considerable interaction.

53 Russ Buettner, Gunman Boasted of Killing Man He Taunted with Antigay Slurs, the Police Say, N.Y. TIMES, June 18, 2013, at A.24 (quotations omitted).
54 Ober, supra note 18.
55 Id.
For example, President Johnson and his Congressional allies worked with civil society leaders like Martin Luther King Jr. and the NAACP to lobby undecided votes and Congressional opponents of the legislation. They used democratic persuasion to criticize arguments by the Klan and other discriminatory organizations that opposed the passage of the Act and its goal of racial equality. They spoke persuasively for the Civil Rights Act against significant opposition. Democratic persuasion has long been crucial to enacting legislation that protects the rights of citizens, including laws against discrimination, sexual harassment, and domestic abuse. Neutralism would deprive political leaders of the persuasive tools they need to pass legislation like the Civil Rights Act. It would hardly be persuasive or even coherent for political leaders to propose protective laws without invoking the democratic values that justify those laws in the first place.

Democratic persuasion is therefore far from being a solitary solution to the problem of protecting the equal status of citizens, as Koppelman seems to think. The full and varied tools of democratic persuasion include speeches by political leaders, public education, the use of the state’s financial resources through grants, and the extension of non-profit status. Democratic persuasion works together with the laws that protect citizens from violence and discrimination. The point is that the diverse strategies of democratic persuasion, combined with the persuasive rhetoric that Ober describes, can effectively promote free and equal citizenship while respecting rights. Besides President Johnson’s successful speeches to pass the Voting Rights Act and the Civil Rights Act, another example of effective democratic persuasion was the discontinuation of tax subsidies to Bob Jones University for its ban on interracial dating. Bob Jones ended its ban in 2000, and apologized in 2008 for its previous discriminatory policies.

In the book, I identify a potential difficulty with the right of free speech, namely its “inverted” character. The reason for protecting free speech is to respect the freedom and equality of citizens. Citizens are equal in having autonomy, or the capacity to reason freely about justice and conceptions of what

---

56 Koppelman writes regarding democratic persuasion, “[d]oes [Brettschenider] really believe that this alone will guarantee that minority rights will be respected in democratic politics?” Koppelman, supra note 12, at 1025.


58 BRETTSCHEINER, WHEN THE STATE SPEAKS, supra note 5, at 71-108.
makes for a good life. The content of hate speech, however, is inverted or opposed to the reasons for protecting rights. Hate speech seeks to deny the freedom and equality of a targeted group of citizens. The danger is that citizens might mistake the state’s protection of the right to express hateful viewpoints with its indifference or even approval of those views. The state could then be seen as being complicit with the hateful viewpoints that it shelters. To clarify that it does not agree with the hate speech that it protects, the state should engage in democratic persuasion to show that it supports free and equal citizenship.

Koppelman might argue that this problem of complicity simply does not exist. He has asserted that citizens do not regard the state’s protection of free speech as signaling its indifference toward hateful expression. He ignores the evidence that citizens and even the courts often conflate the state’s viewpoint neutrality in protecting free speech rights with its neutrality toward the content of that speech. As I argued in a piece in Foreign Policy, riots broke out around the world to protest the U.S. government’s protection of the free speech rights of an anti-Muslim group. These protests conflated the state’s protection of free speech rights with its indifference to hateful, anti-Muslim values.

In the end, Koppelman agrees with value democracy’s claim that free speech rights are based on substantive values of freedom and equality. But he treats the state’s non-neutrality toward values as an obvious point when it is actually controversial in constitutional jurisprudence, the academic literature on free speech, and popular debate. Koppelman overlooks how neutrality in rights protection has too often been thought to signal the state’s neutrality toward the values expressed by hate speech. This conflation is made not only by the people who protested anti-Muslim hate speech, but by constitutional lawyers and the Court itself. Prominent constitutional lawyers, including Larry Alexander and Martin Redish, base their thinking about free speech on the idea that the state should be neutral in the values it expresses. The Court itself has repeated the view multiple times that “when it comes to

60 Koppelman, supra note 12.
61 Id. See supra note 2 and accompanying text.
the First Amendment there is no such thing as a false idea.”\textsuperscript{62} Court decisions like \textit{Christian Legal Society v. Martinez} mistakenly claim that neutrality about the state’s values is required by the right of free speech.\textsuperscript{63} The state needs to engage in democratic persuasion to clarify that it is not neutral toward the hateful values that it protects. Democratic persuasion would promulgate or make transparent the reasons for protecting rights, and the state’s own non-neutral promotion of equality.

Koppelman’s final worry is that democratic persuasion seems to disrespect the free thinking of citizens or their capacity to reason.\textsuperscript{64} The concern is that citizens might be expected to conform to official views. Value democracy has two responses to this worry. First, democratic persuasion provides standards to evaluate and criticize state speech itself. The “substance-based limit” requires the state to use state speech only in ways that are consistent with respect for citizens’ freedom and equality. If the state speaks in a way that promotes discriminatory views, that would be condemned by the normative standards that must guide democratic persuasion. Democratic persuasion includes not any act of state speech, but only speech that defends the democratic values of free and equal citizenship.

A second response is that democratic persuasion criticizes hateful viewpoints in a way that is consistent with respect for the capacity of citizens to reason. Josiah Ober’s excellent essay explains this point well. President Johnson’s speeches on civil rights were critical of the discriminatory viewpoints of the southern segregationists. Johnson argued that these views contradicted the core values of democracy and the American Constitution. Many segregationists disagreed with equality for African-Americans. But were the segregationists disrespected as citizens when their discriminatory viewpoints were met with reasons, evidence, and arguments for equality? Although President Johnson criticized the segregationists’ denial of African-Americans’ right to vote, this kind of criticism does not disrespect citizens. They are respected because democratic persuasion appeals to their reason, continues to include them in democracy, and protects their right of free expression.

President Johnson in his speech to Congress on the Voting Rights Act defended the importance of voting rights for all, even as he criticized the discriminatory views that would

\textsuperscript{62} REDISH, supra note 2, at 167 (“Under the First Amendment there is no such thing as a false idea.” (quoting Gertz v. Robert Welch, Inc., 418 U.S. 323, 339 (1974)).

\textsuperscript{63} Christian Legal Soc’y v. Martinez, 130 S. Ct. 2971, 3009 (2010).

\textsuperscript{64} Koppelman, supra note 12.
deny the vote to African-Americans. As he declared, it is wrong “to deny any of your fellow Americans the right to vote in this country.” President Johnson’s criticism of discriminatory views was not disrespectful. Rather, democratic persuasion vindicated a deeper idea of respect for the freedom and equality for all citizens that is at the heart of democratic legitimacy.

IV. IS DEMOCRATIC PERSUASION TOO RETICENT? A RESPONSE TO ROBIN WEST ON TORT AND HYPOCRISY

I have responded to concerns about democratic persuasion from a neutralist direction. Robin West raises concerns from the opposite side. She thinks that the logic of my argument should open room for even more extensive kinds of democratic persuasion beyond what I would endorse. In particular, she argues for using tort law to allow private citizens to sue and inflict civil penalties on hate groups. Jeremy Waldron and Catharine MacKinnon have similarly suggested that people who are subject to hate speech could sue for damages. A tort-based approach has recently been taken up by Justice Alito. Dissenting in *Snyder v. Phelps*, Alito acknowledges that the state cannot prohibit the Westboro Baptist Church’s hate speech, but he suggests that a private tort might be allowed against the church.

I recognize that these proposals are intended as ways of criticizing hate groups and promoting the values of free and equal citizenship. But I am skeptical about using private torts as a means of pursuing democratic persuasion for two reasons. First, it is crucial to my proposal that individuals be able to effectively exercise a right to dissent from democratic persuasion and to resist it. Unlike Michael McConnell and some of the Court’s jurisprudence, I have argued that this right does not entitle the advocates of hateful or discriminatory viewpoints to public subsidies of their views. However, torts

---


66 West, supra note 13.


69 On my view, doctrines such as unconstitutional conditions and the limited public forum are mistaken in creating positive rights for viewpoints to be subsidized by the state. See Corey Brettschneider, *Value Democracy as the Basis for Viewpoint Neutrality: A Theory of Free Speech and Its Implications for the State Speech and*
would go much further than non-subsidy. While non-subsidy refuses to give additional financial support for hateful viewpoints, torts risk taking away all resources from citizens who engage in hate speech, perhaps to the point of bankruptcy. The aim and likely effect of torts is that the right to free speech would become a mere formality without the resources to exercise it. I would thus reject proposals that aim to prohibit speech, such as a proposal to publicly tax or privately sue hate speech out of existence. Such proposals would violate the right to free speech in a way that would deliberately exclude the real possibility that citizens could dissent. By contrast, non-subsidy allows citizens to continue to dissent. For example, although their tax subsidies were discontinued, Bob Jones University and the Christian Legal Society continued to exist and exercise their right to dissent.

Second, I think that turning to private tort risks losing the clear notion that the state has an obligation to speak in democratic persuasion. Pursuing democratic persuasion through private torts might create the impression that hate speech is only an issue between individuals, rather than an attack on the public, democratic values of freedom and equality that the state has a duty to secure for all citizens. The state makes it clear that it is expressing public democratic values when it criticizes hate groups and refuses to subsidize them. I worry that the turn to private law moves in exactly the other direction, turning hate speech into a private matter rather than one of public relevance. At least as the proposal is stated, it is not clear how West’s plan could turn private lawsuits into the kind of public promulgation that I am advocating. It is also not clear how it would be compatible with the right to free speech.

While being sympathetic to democratic persuasion, West also raises the concern that the state might speak in favor of democratic values, while in reality working to undercut them. West calls this the problem of the “hypocritical state.” The risk is that democratic persuasion could function as a kind of ideology that masks the real injustices committed by the state. West is right to worry about the hypocritical state. A state that

---


70 Ben Zipursky has suggested in conversation that this objection might be met if the attorney general brought the suits to make clear that the public interest is at stake. This is an interesting proposal, and if developed, might indeed answer the objection about public expression, although my other concerns about abridging the right to dissent would remain.

attacks democratic values, no matter what it says, would undermine the basis for its legitimacy. Democratic persuasion is a necessary but not sufficient part of what a fully legitimate state must pursue. The legitimate state must respect the value of free and equal citizenship in both its acts and its speech.\textsuperscript{72}

West and I both regard the hypocritical state as illegitimate. But we differ in that she seems to regard the hypocritical state as worse than a state that consistently disrespects free and equal citizenship. I would argue that the hypocritical state is highly problematic, but it at least identifies the ideals of legitimacy that its actions fall short in accomplishing. These ideas can then be used by critics of the state to advocate for change. The public officials who verbally affirm democratic values can also be held to the standards that they have already admitted are valid. Jon Elster calls this tendency for verbal commitments to turn into public standards the “civilizing force of hypocrisy.”\textsuperscript{73} For example, the segregationist government in the United States was a hypocritical state. It espoused ideals of equality that it systematically violated. But pioneers in the Civil Rights Movement and supportive public officials were able to transform that state into one that was more legitimate. They pointed to America’s hypocrisy on race to argue for the country to move closer to the ideal of equality. Thus, the legitimacy of the democratic persuasion was actually a lever to bring the practices of the state in line with its own stated ideals. In this sense, the hypocritical state may be preferable to one that holds no pretense of supporting individual rights. At least in the hypocritical state, there is a foothold for reform.

The evolution of constitutional law is often characterized by movements to iron out the contradictions of the hypocritical state and to narrow the gap between rhetoric and reality. But this progress is only possible if the state at least acknowledges the validity of democratic values. Hypocrisy is the midway point on the path from injustice to legitimacy.

\textsuperscript{72} I have argued that the legitimate state must guarantee welfare rights. See BRETTSCHEIDER, DEMOCRATIC RIGHTS, supra note 20; Corey Brettschneider, Public Justification and the Right to Private Property: Welfare Rights as Compensation for Exclusion, in PROPERTY-OWNING DEMOCRACY: RAWLS AND BEYOND (Martin O’Neill & Thad Williamson eds., 2012).

\textsuperscript{73} Jon Elster, Deliberation and Constitution Making, in DELIBERATIVE DEMOCRACY 97, 111 (Jon Elster ed., 1998).
V. IS DEMOCRATIC PERSUASION TOO LIMITED IN SCOPE AND DOES IT IGNORE SUBTLE FORMS OF COERCION?: A RESPONSE TO SONG

Sarah Song, like West, thinks that, far from being too aggressive, democratic persuasion might be too weak. But whereas West focuses on expanding the tools of democratic persuasion, Song seeks to expand the scope of the views that are addressed by it. She argues that democratic persuasion should be committed to cutting off non-profit status to the Catholic Church. In her view, the only thing that seems to distinguish Bob Jones and the Catholic Church is that while the former is a racist institution, the latter is sexist.

In *When the State Speaks*, I suggest that democratic persuasion should argue against viewpoints, including those that are sexist, racist, or homophobic, that attack the public ideal of free and equal citizenship. The Church, however, has taken the position that restricting the priesthood to men does not imply that it opposes the equal status of women. It claims that it is expressing a purely religious doctrine that links the priesthood to Jesus Christ, a male, while welcoming women to serve in high offices in the professions and politics. Song’s response is that the case for the non-public nature of the priesthood restriction is “ambiguous” at best.

I agree with Song that the question of whether the Catholic Church opposes free and equal citizenship is complicated and that my argument in their defense might not be obvious. I take the position, though, that when the issue is ambiguous we should give the organization the benefit of the doubt about the consistency of its position with the ideal of free and equal citizenship. Such an approach is called for to avoid the charge of selectivity and partisanship that Calabresi raises. It is also needed to make the kind of pluralism that is central to value democracy as broad as possible.

The Westboro Baptist Church and Bob Jones University before 2000, on the other hand, were not ambiguous about their opposition to the equality of blacks and gays. Bob Jones University banned advocacy for the right to interracial marriage.

74 Song, supra note 14.
75 Bob Jones University dropped its ban on interracial dating in 2000, and later issued an apology in 2008 for its previous racially discriminatory policies. See supra note 57 and accompanying text.
murdered. These are not hard cases of reasonable disagreement and clearly differ from the example of the Catholic Church.

While Song pushes in some parts of her essay toward a more extensive form of democratic persuasion, in other parts she worries that democratic persuasion might be coercive. While Song invokes the nineteenth-century English philosopher John Stuart Mill’s warnings of the coercive dangers of majority opinion to note that some forms of persuasion might feel coercive to those on the receiving end of it. I think that Mill could not have meant to say, however, that any kind of persuasion is coercive. To say that would be inconsistent with his defense of free speech, in which he rules out coercion, but allows for “remonstrating with [a person], or reasoning with him, or persuading him.” Mill makes this remark in the context of allowing others to reason with a person for the sake of his own good. If Mill permits that, he would surely allow others to persuade a person for the sake of protecting the standing of others as free and equal citizens. His distinction between prohibited coercive measures and permissible argument also can only make sense if persuasion is not equated with coercion. It is not coercive to allow good arguments from others to persuade a free citizen to change a belief.

Similarly, I would resist the idea, which Song finds in the work of Robert Nozick, that anything that makes a choice less desirable counts as coercive once that consequence is communicated. She concludes that the policy of withdrawing the privileges of tax exemption for discriminatory groups is coercive, because it makes a choice (continuing to undermine free and equal citizenship) less desirable (it will result in the discontinuation of tax exemption). This definition of coercion is overbroad. As Song herself mentions, it allows any way of making an action “less desirable” count as coercive. For instance, it seems to imply that it would be coercive to have a policy that students should maintain a certain G.P.A. to receive a state-funded college scholarship, since the policy makes having low grades less desirable. I use the term coercion instead to refer to acts that aim

---

76 Song, supra note 14.
77 J. S. Mill, On Liberty, in ON LIBERTY AND OTHER WRITINGS 1, 13 (Stefan Collini ed., Cambridge Univ. Press 1989) (1859). Although Song claims that I downplay Mill’s worries about the coercive aspects of majority opinion, he did not think that people had a duty to never criticize hate speech. As I wrote in WHEN THE STATE SPEAKS, Mill believes “that when ‘the acts of an individual may be hurtful to others or wanting in due consideration for their welfare, without going to the length of violating any of their constituted rights . . . the offender may be justly punished by opinion, though not by law.” Id.
at prohibiting a choice, not merely acts that make a choice harder to make. I employ this narrower definition of coercion as aiming at prohibition to distinguish proper from improper means of democratic persuasion.

Attempting to force people to change their minds violates their right of free expression. But attempting to convince them can be made compatible with their free choice. The use of subsidy and non-profit status also leaves this choice intact. Withdrawing a tax privilege is not coercive in this definition because the state does not aim at prohibiting the existence of these groups. There is no denial of rights in ending these subsidies, since there is no entitlement for groups to receive government subsidy or exemption from common taxation. I have argued at length against a misplaced doctrine of constitutional conditions that reframes free speech as a mistaken kind of positive right to tax subsidy.78 The continued existence of Bob Jones University and groups like the Christian Legal Society, despite the discontinuation of government subsidy, serves to illustrate that there is no coercion in my sense of the term that comes with denial of subsidy.

In her reply, Song raises an additional argument for a free speech right to receive subsidies. She cites John Rawls’ notion of the worth of liberty to suggest that groups might claim 501(c)(3) status and subsidies as a matter of rights.79 According to Rawls, fundamental rights, to avoid being mere formalities, must be backed by resources that allow individuals to exercise those rights. The literature on unconstitutional conditions and the related “limited public forum” doctrine regard it as coercive when the state denies money on the basis of the viewpoint being expressed.80 According to this literature, it violates rights to condition funds on private actors endorsing a set of beliefs.81

I want to resist, however, any excessively broad defense of a right to receive government subsidies for discriminatory and hate groups based on the worth of liberty. The fact that many of these groups have access to non-government resources to continue expressing their views suggests that ending government subsidies does not deny a substantive right. As I

78 Brettschneider, Value Democracy, supra note 69, at 639-40. For an example of such a doctrine, see, e.g. Kathleen M. Sullivan, Unconstitutional Conditions, 102 HARV. L. REV. 1413, 1424-25 (1989).
79 Song, supra note 14.
80 See, e.g., Sullivan, supra note 78, at 1428.
81 Id. at 1428, 1506.
have noted above, Bob Jones University and the Christian Legal Society continue to thrive without these subsidies, and so would the Boy Scouts. I have argued in other work that distributive justice in material resources should be an important concern for citizens in a democracy. Citizens should have a right to the basic material resources needed to exercise their rights and liberties. But the right to basic resources does not imply any right to be subsidized in expressing one’s viewpoint. Positive rights exist as an entitlement of poor and needy citizens, separate from the issue of non-profit status or government grants for specific purposes. The logic of the worth of liberty animates much of the flawed jurisprudence on unconstitutional conditions and the limited public forum.

I do agree with Song and argue in the book that when the government has a monopoly on a resource, such as that which the U.S. Postal Service once possessed over mailing letters, the denial of that resource might result in the denial of a right. For example, the Supreme Court has ruled that the Postal Service cannot condition mailing letters based on the viewpoints they express. In this case, the denial of a resource denied the substantive right to free speech, given how much the citizenry depended on the Postal Service to send letters. On these same grounds, I would oppose measures that would have libraries ban books or block access to certain internet sites in the supposed name of democratic persuasion. These libraries are the sole access that some users have to information and the internet. To give another example, consider a hypothetical based on the Christian Legal Society case. If the Christian Legal Society, as a student group, were barred from meeting on campus, there could be a worth of liberty argument for a free speech violation. Meetings in campus spaces might be the main way to communicate for students, such that denying access to those spaces would deny free speech. However, the right for a student group to meet on campus does not mean that they have a right to the active support of an official state subsidy paid by public taxation.

In sum, I agree with Song that in a limited set of cases there might be a free speech argument based on a concern for the worth of liberty. But these cases are limited to government monopolies. They do not apply to government subsidy of discriminatory or hateful groups. The concern about monopolies

82 See Brettischneider, Democratic Rights, supra note 20, at 5-6.
serves as yet another bulwark to protect free speech in my theory of democratic persuasion.

CONCLUSION

I began this response essay by highlighting Frank Michelman’s arguments for democratic persuasion as a part of the state’s legitimacy. He suggests the best way to understand this relationship is that democratic persuasion, while not sufficient on its own for legitimacy, should be seen as contributing to whether a state meets the threshold for being democratically legitimate. What makes a state legitimate is whether it upholds the democratic values of freedom and equality for all its citizens. I emphasized why stability, the duty of transparency, and the arguments for public trust and interconnection all supported a role for the state in promoting and defending its own underlying values.

With these arguments in mind I turned to critics who thought democratic persuasion should be tempered. Steve Calabresi argued that while the theory is essentially correct, it should not discontinue the privileges of tax-exemption for hateful religious groups. I responded by highlighting why a right to be free from coercive bans on speech or religion does not entail a right to be free from criticism or an entitlement to a public subsidy. I also argued that my proposal clarifies the criteria for receiving tax exemption, reducing the potential for abuse of government power compared to existing law.

Koppelman worried that democratic persuasion might be unnecessary, ineffective, and disrespectful to the proponents of hateful viewpoints. I argued that Koppelman confused criticism of hateful viewpoints with disrespect for the reasoning capacity of citizens who express those viewpoints. Democratic persuasion shows respect, even while challenging beliefs, by upholding those citizens’ free speech rights, by appealing to their reason, and by continuing to include them in our shared democracy. Drawing on Josiah Ober’s essay, I gave examples of the effectiveness of democratic persuasion from the arguments that President Johnson and other political leaders made for the Civil Rights Act and Voting Rights Act. Forcing political leaders to be silent or neutral about democratic values, as Koppelman seems to suggest, would deny them of one of their most powerful tools to pass legislation. I explained that democratic persuasion is part of a larger defense of free and equal citizenship, including laws against discrimination and domestic violence.
While in different ways Calabresi and Koppelman find democratic persuasion to be too strong, Robin West argued I should widen the theory to include tort law. West’s proposal would allow private individuals to sue the advocates of hateful viewpoints. Much of what she says suggests an important reply to some of Koppelman’s claims that racism may not be much of a problem in civil society. Her arguments reinforce my concern about ways that citizens in their relationships with each other could threaten values of free and equal citizenship. But I resisted her specific proposals to use tort law on the grounds that it would deny the right of free speech and the right to dissent from democratic values.

Song in turn highlighted why I should expand the scope of the theory to criticize the Catholic Church for restricting the priesthood to men. While agreeing that democratic persuasion should criticize beliefs that discriminate against women, I resisted her proposal to condemn the Catholic Church. I distinguished between the Church’s position, which welcomes women to serve in professional and political office, from the viewpoint of groups, like the Westboro Baptist Church, that attack free and equal citizenship. Although the strength of the Church’s claim to support public equality for women might be questioned by Song, democratic persuasion is limited to clear cases of opposition to democratic values. This distinction allowed me to demonstrate some of the reticence Calabresi calls for without completely exempting religious groups like the Westboro Baptist Church from democratic persuasion.

In sum, I propose democratic persuasion as a “golden mean” between the prohibitionists who would ban hateful viewpoints and the neutralists who would have the state say nothing to criticize discrimination. It defends the standing of all citizens as free and equal while respecting their expressive rights.
ARTICLES

Discovery and Darkness

THE INFORMATION DEFICIT IN CRIMINAL DISPUTES

Ion Meyn†

INTRODUCTION

Does a defendant have the right to investigate the crime for which he is charged? Courts would say yes, and in fact impose a duty upon defense counsel to investigate.¹ The prevailing scholarship would also say yes, as it presumes a defendant will perform a pretrial investigation that uncovers evidence the State does not.² Yet despite these duties and presumptions, a criminal defendant is not structurally assigned an investigatory role in his case.

The typical discovery statute only permits a criminal defendant to view fragments of the State’s evidence against him, deeming him a passive recipient of information.³ In contrast, the

† Assistant Clinical Professor, University of Wisconsin Law School. I am indebted to the members of the University of Wisconsin Law School Junior Faculty Group and to Alex Huneeus, David Schwartz, Howard Wasserman, Mary Leary, Jennifer Laurin, Babe Howell, Gary Sloboda, Keith Findley, and Cecelia Klingele for their close reads and insight. I thank Leslie Kuhn, Alisha McKay, and Monica Mark for their tireless work and contribution.

¹ Williams v. Taylor, 529 U.S. 362, 396 (2000) (holding that a criminal defense attorney has a duty to investigate, which implies that the criminal defense attorney has the power to do so).


³ See, e.g., FED. R. CRIM. P. 16 (delineating what is and is not subject to disclosure); STANDARDS RELATING TO DISCOVERY AND PROCEEDURES BEFORE TRIAL pt. 2
State is assigned a central role in conducting a formal investigation and, in fulfilling that duty, exercises powerful police powers to search property, seize evidence, and interrogate witnesses. Granted the discretion to compel information from any source, civil litigants on both sides of the dispute are likewise structurally assigned pretrial roles to assess liability. A criminal defendant, having no discretion to compel pretrial discovery and permitted but a keyhole view of the State’s evidence, is the only litigant relegated to darkness. To grant a criminal defendant the discretion and power to conduct an independent inquiry into the incident would be to recast a defendant as having a formal role to play in the criminal investigation.

In recognizing some degree of information disparity, scholars advocate for more resources (adequate staffing and sufficient funds for investigators and experts) and for open-file policies that increase access to prosecutorial files. These reforms would go some distance in mitigating the existing information gap. An open-file policy has the laudatory goal of encouraging a more informed outcome. But affording a better view of a prosecutor’s file will rarely permit an alternative view of the crime. The prosecutor’s file is populated with police reports narrated by authors who have determined that the defendant is guilty. Further, an open-file policy is not as open as the term suggests. The policy only calls for documents from

(Advisory Comm. on Pretrial Proceedings, Tentative Draft 1969) [hereinafter ABA STANDARDS] (enumerating limited categories of evidence that prosecutor must disclose); see also Jenny Roberts, Too Little, Too Late: Ineffective Assistance of Counsel, the Duty to Investigate, and Pretrial Discovery in Criminal Cases, 31 FORDHAM Urb. L.J. 1097, 1098 (2004) (noting the disparity between discovery rights afforded to civil and criminal litigants); infra note 57 and Figure 3.


5 See, e.g., Laurence A. Benner, The Preemption of Guilt: Systemic Factors that Contribute to Ineffective Assistance of Counsel in California, 45 CAL. W. L. REV. 263, 277 (2009) (“The most important finding from our study is the discovery that indigent defense providers in many California counties lack the resources necessary to conduct adequate defense investigations.”); Brown, supra note 2, at 1602 (defense attorney’s ability to investigate is limited by budgetary constraints); Máximo Langer, Rethinking Plea Bargaining: The Practice and Reform of Prosecutorial Adjudication in American Criminal Procedure, 33 AM. J. CRIM. L. 223, 236 n.42 (2006) (same).

6 Alafair S. Burke, Revisiting Prosecutorial Disclosure, 84 IND. L.J. 481, 514 (2009) (arguing for an open-file policy); see also Brown, supra note 2, at 1637 (stating open-file discovery is best option); Peter A. Joy, Brady and Jailhouse Informants: Responding to Injustice, 57 CASE W. RES. L. REV. 619, 641 (2007) (favoring open-file policy); Langer, supra note 5, at 276 (open-file policy may diminish the coercive nature of plea bargaining); Roberts, supra note 3, at 1153-55 (concluding open-file discovery is best solution).
the prosecutorial file, not the larger investigative file where any reports that do not support the State’s theory likely reside. More fundamentally, an open-file policy does not alter the nature of the discovery disparity. Because the policy does not grant any discretion to compel discovery, it does not assign an investigatory role to a defendant. In merely expanding a defendant’s entitlement to discrete categories of information from one source, the State, the policy maintains the status quo. The defendant remains a passive recipient of information.

This article contends that the existing dynamic in criminal cases is inconsistent with the design of the adversarial system and results in a factual deficit that undermines the legitimacy of outcomes. Even assuming a defense attorney is well trained and well resourced, she is structurally precluded from compelling information from witnesses, which hinders her ability to conduct an independent inquiry into the question of who committed the crime, and, if a defendant did commit a crime, the degree of culpability. Only entitled to limited disclosures of the State’s evidence, the defense counsel is instead forced to suggest a counter-narrative based largely on documents selected and prepared by the opposing party. Part I of this article explores the difference between an informal investigation, to which a criminal defendant is confined, and that of a formal investigation, from which a criminal defendant is excluded. Part II identifies common formal investigative powers that are extended to civil litigants, and it employs a case study to ascertain what is lost in the absence of formal powers to investigate. Part III surveys potential arguments against assigning to a criminal defendant a central role in the investigation, and responds to these concerns.

I. A STRUCTURAL EXCLUSION FROM THE CRIMINAL INVESTIGATION

Once the State initiates its investigation, some portion of the universe of relevant facts becomes known to the State—these facts make up the State’s investigatory file. Some smaller subset of these facts is forwarded to the prosecutor, making up the prosecutorial file. In the typical jurisdiction, a defendant is only entitled to a limited view of the prosecutorial file—these are statutorily required disclosures.
The typical criminal discovery statute does not grant a defendant formal pretrial investigatory power, defined as the discretion to compel facts from multiple sources. Formal investigatory powers may be expressed in various ways. In civil litigation, these powers take the form of depositions, interrogatories, and document requests. A criminal defendant, however, is rarely afforded such tools. He is instead entitled to discrete categories of opponent-sourced information found in the prosecutorial file. A defendant is rarely authorized to view the entire prosecutorial file or to view any part of the more expansive investigative file. Alternate routes are closed off, as most open-record laws prohibit access to documents that are part of an open investigation. A criminal defendant does not typically have the power to compel information from evidentiary territory uncharted by the State that may provide an alternative theory of liability. Rather, a defendant’s statutory role is limited to receiving what is forwarded to him by the prosecutor. In contrast, the State yields

extraordinary formal powers in collecting information from any source it deems relevant to the investigation.

These asymmetrical privileges to information create a dynamic unique to criminal law. The prosecutor assesses the particular facts that executive agents forward to her, releases facts she determines a defendant should view, and adjudicates the dispute through a plea offer that is supported by facts she selects. Though a criminal defendant has no structurally assigned role in the investigation, he is subjected to an adversarial process. If the integrity of the adversarial system depends on testing the pretrial conclusion made by the executive in its investigation, the failure to create the conditions for a counter-investigation undermines that integrity. To be clear, no statute prohibits a defendant from engaging in an informal investigation. Nothing precludes a defendant (with the exception of pretrial custody) from asking around for names of individuals who might have heard something about the crime, to ask the manager of a gas station for surveillance tapes, or to press a detective for information about who told him what during the course of the investigation. But while there is no prohibition to asking these questions, there is also no right to a response.

A. Informal Investigation: Best When the Stakes are Low and There Are Multiple Sources of Information

A body at rest will remain at rest unless it is subject to an outside force. Information, too, tends to remain undisturbed in the absence of an outside force. The more force applied to a source of information by an investigatory tool, the more is revealed. If statutorily granted tools of investigation backed by subpoena power and the threat of judicial sanction define what information is subject to discovery, statutory power not afforded defines what information tends to remain protected. In remarkably uniform fashion, civil litigants who meet low jurisdictional minimums can

---

8 Bennett L. Gershman, Preplea Disclosure of Impeachment Evidence, 65 VAND. L. REV. EN BANC 141, 142 (2012) ("As a former state prosecutor, I recall the issues surrounding preplea disclosures in practice. The give and take of the relatively informal bargaining process typically focused on how much information about the case I was willing to share with defense counsel.").

9 Gerard E. Lynch, Our Administrative System of Criminal Justice, 66 FORDHAM L. REV. 2117, 2120 (1998) ("[F]or most defendants the primary adjudication they receive is, in fact, an administrative decision by a state functionary, the prosecutor.").

10 This is Newton’s first law of motion, the law of inertia. 1 ISAAC NEWTON, THE MATHEMATICAL PRINCIPLES OF NATURAL PHILOSOPHY 19 (Andrew Motte trans., 1729) (1687); see also STEVE HOLTZNER, PHYSICS FOR DUMMIES, 64-65 (2006).
utilize a powerful array of formal investigatory tools.\textsuperscript{11} In criminal law, the State maintains a monopoly over investigative choices\textsuperscript{12} and is afforded formal investigatory tools that in some respects eclipse those available to civil litigants.\textsuperscript{13} Yet, a criminal defendant, unassigned a formal investigatory role, is left to initiate an investigation by \textit{informal} means.

Anyone has the ability to conduct an informal investigation. It is a method we use daily. By definition, an informal investigation consists of asking a question with the hope of receiving an answer. When we ask a passerby for the time, we almost expect an answer. If we fail in our attempt, we can usually turn to another passerby and obtain the same information. In this scenario, the stakes are typically low. The responding party faces little consequence in providing an answer, maybe annoyance, maybe a fleeting sense of satisfaction for engaging in an act of civility. Any negative repercussions for failing to obtain an answer are mitigated because the requested information is not limited by the source. The asking party can ask any number of individuals for the time of day. In these low-stakes, multiple-source scenarios, the informal method is adequate to satisfy the asking party’s objectives.

The informal method’s potential limitations begin to emerge as the stakes increase and the sources of information begin to decrease. Take the Cabbage Patch doll shortage in 1982. The hysteria to secure a doll for one’s child and bring holiday happiness to the home gained national attention. Other than resorting to bribes or violence, parents had at their disposal only informal powers of investigation to learn whether a store had a doll in stock.\textsuperscript{14} If sales associate #1 denied having dolls for sale, a

\begin{footnotesize}
\begin{enumerate}
\item See, e.g., rules cited supra note 4; infra Figure 3. Civil litigants in federal court are entitled to seek discovery of “any nonprivileged matter that is relevant to any party’s claim or defense.” FED. R. CIV. P. 26. States typically adopt statutes that are similarly broad in scope. See, e.g., CAL. CIV. PROC. CODE § 2017.010 (West 2012) (stating that parties may obtain discovery regarding any relevant, non-privileged matter, that is admissible or “reasonably calculated to lead to the discovery of admissible evidence”).
\item Such powers include a threat of a probation hold and revocation, the power to arrest, the power to search a person or place, the power to seize evidence, and the opportunity to falsely assert that the failure to cooperate will lead to negative consequences.
\item See, e.g., JAKKS Pacific, \textit{Cabbage Patch Kids Craze!}, YOUTUBE (May 1, 2008), http://www.youtube.com/watch?v=9s0lIvx7Pvs.
\end{enumerate}
\end{footnotesize}
motivated parent might seek to undermine or corroborate this answer by questioning sales associate #2. But a parent could not compel production of inventory records and sales receipts or inspect inventory to verify these representations. And the parent would have a difficult time sufficiently testing the personal knowledge of any sales associate. We may not trust the answer we receive in an informal investigation, but such a method does not readily permit the sustained inquiry that is required to gather information, assess critical facts, and test credibility.

The informal method presents unique challenges for a criminal defense attorney. The stakes are high with a felony charge—a defendant’s long-term liberty interest is at issue, and any witness to the crime may face negative consequences for furnishing an answer. Sources of information tend to be limited—and may be restricted to one person, as in the case of a sole eyewitness. From the perspective of an experienced defense counsel, the eyewitness may be a potential alternate suspect. In such a scenario, the eyewitness will likely be reluctant, even hostile, to the idea of voluntarily disclosing information to the defendant. Unlike turning to another passerby to ask the time, a defense attorney cannot seek out another source if only one person witnessed the crime. Where the informal method may succeed in a single-source, high-stakes scenario, the formal method will provide a better chance of success in a criminal case: either the witness will comply with a pretrial subpoena, or, if a witness refuses, the party seeking information can request appropriate procedural or evidentiary sanctions. If the witness is the opposing party’s sole witness, for example, a court might prohibit that witness from being able to testify at trial due to the failure to comply with a pretrial subpoena.

**Figure Two: Where Informal and Formal Powers Are Most Effective:**

![Diagram showing the effectiveness of informal and formal powers in low and high stakes scenarios with single and multiple sources of information.](image-url)
This is not to say that the informal method lacks value in a criminal investigation. The informal method, for example, avoids a strategic shortcoming inherent in a deposition—the cost of transparency. One must notify the opposing party before taking a deposition, and what the deponent reveals is revealed to the opposing party. A litigant must consider the risk that the opposing party will benefit from the deposition more than he will. (The State does not incur these risks, as it is the only litigant in the common law system privileged to conduct interrogations and secret grand jury hearings without notice.\textsuperscript{15}) The use of informal methods to interview a witness, in contrast, occurs in the deep woods—if the opposing party does not hear a tree fall, the tree has not fallen. An attorney is under no pretrial obligation to share harmful revelations with the State. The informal method is also potentially more efficient than the formal method. There is no need to serve a witness or to schedule a deposition, and there are no costs associated with a court reporter’s transcript.\textsuperscript{16} Interviewing a witness on her own stoop may uncover valuable information. In this setting, she may be more candid, provide unsolicited information, and criticize actions of police that in the presence of a prosecutor she may not express.

Still, witnesses from neighborhoods where the line between being a witness and a suspect is viewed as arbitrary tend to be reluctant to divulge information. Even where a witness initially cooperates, it is not uncommon for her to evade questions as they become more probative and implicitly confrontational. In an informal interview, the litigant is subject to the will of the witness. In a formal investigation, a witness is subject to the will of the asking party, and ultimately, the judicial sanctions that accompany any non-compliance. Depositions permit unyielding examination. Any obfuscation is

\textsuperscript{15} This information is subject to any discovery statute or \textit{Brady}, which, as further discussed \textit{infra}, will result in little daylight for a defendant.

\textsuperscript{16} Though not required to do so, civil litigants generally use a stenographer in a deposition, which arguably provides the most accurate record. Stenographers generally charge an hourly rate and then charge the requesting party more per page than a non-requesting party who wants a transcript copy. A legal blog in Minnesota surveyed local transcription services—the commentary to the blog provides color to an otherwise dry survey. Jack Smith, \textit{Do You Know What Your Court Reporter is Charging Your Clients for Depositions?}, MINNESOTA LITIGATOR (Feb. 28, 2012), http://www.minnesota-litigator.com/2012/02/28/court-reporter-charging-clients-depositions/. At 40 pages per hour, an eight-hour deposition using a court reporter might cost a requesting party $1,600 ($40 hourly charge + $4 per deposition page). Other options, however, are available—like a “dirty deposition” in which a party records the deposition and has its own office prepare a transcript, which after review is deemed accurate by party stipulation.
on the record and may demonstrate the witness’ bias. Use of formal tools permits a more ordered implementation of an investigative strategy; to move deliberately from peripheral witnesses (to gather background information) to critical witnesses (to expose inconsistencies and challenge credibility). One can attempt the same using informal tools—but in the absence of subpoena power, it is difficult to schedule witness testimony according to any preconceived plan. Use of the formal method to compel documentation and testimony potentially provides a nuanced and comprehensive pretrial understanding of the facts at issue and the motivations of witnesses. This sort of sophisticated approach is difficult to implement with only informal methods, where, despite the high stakes at issue and limited sources of information, growling dogs and refusals to open a door may leave a defendant with no recourse.

In contrast to the informal method that relies on voluntary compliance, the power to compel a person to appear at a place and time to answer questions under oath may be in many circumstances the only feasible way to capture critical pretrial information. Other than the State—which historically has had at its disposal police powers to conduct a formal investigation—the power to compel information from any source was not extended to private litigants until the advent of the modern discovery era. Modern-era reforms afforded formal investigatory tools to civil litigants, but not to criminal defendants.

B. Modern Era: Extending Formal Powers to Private Litigants

In 1938, Congress ushered in the modern era of pretrial fact development and testing in civil litigation. Before this time, plaintiffs were first required to conduct an informal investigation to substantiate the complaint, and only after this burden was met could they petition the court to compel pretrial information. The new rules directly afforded litigants formal

---

17 See, e.g., TMZ, Lil Wayne Deposition—I Don’t Recall!, YOUTUBE (Sept. 25, 2012), http://www.youtube.com/watch?v=YQsMqRvPzRw.
18 See infra Part I.B.
20 Beisner, supra note 19, 554.
tools of investigation. A formal investigatory tool grants discretion to compel relevant information from any source, protecting only privileged information. Discretion is given directly to the litigant—a court order is not needed to wield the power. Underscoring the significance of these reforms to civil litigation, Martin Redish likened these rights to the gift of fire.

A formal investigatory tool does not guarantee consideration of every relevant fact. A host of privileges protects against the production of information—what one tells his priest during confession may result in penance but not in paying up or doing prison time. The low value of a case may negate the feasibility of a full-throttled formal investigation. Other case-specific circumstances may prevent disclosure. A witness might live in the litigant’s zip code—she is easy to find, serve, and depose. But if the witness lives in rural Portugal, it may be prohibitively expensive to find her (third gravel road after apple tree), serve her (one must refer to the Hague Convention on Taking of Evidence Abroad in Civil or Commercial Matters), and to question her (travel, translation, and lodging).

Though the causal basis for this sweeping shift in civil procedure is subject to debate—“exactly why this dramatic change was made was never fully clarified by any of the key actors”—there is some agreement that reforms were “premised in some sense on the notion that ‘mutual knowledge of all the relevant facts gathered by both parties is essential to proper litigation.’” There were also aspects of the old civil litigation system that invited reform—plaintiffs who did not have access to facts were somehow expected to allege facts with specificity before gaining access to the courts. Reform lowered the entrance fee (liberalizing pleading requirements) and provided pretrial mechanisms to discover, collect, and test otherwise

---

21 Fed. R. Civ. P. 26 advisory committee’s note (1937) (“This rule freely authorizes the taking of depositions under the same circumstances and by the same methods whether for the purpose of discovery or for the purpose of obtaining evidence. . . . The more common practice in the United States is to take depositions on notice by the party desiring them, without any order from the court, and this has been followed in these rules.”).


24 Before engaging in civil discovery across international borders, a prudent attorney determines whether this multi-lateral treaty applies and whether the particular request is covered by its guidelines.


26 Redish, supra note 23, at 847.

27 Id. (citation omitted).

28 Id. at 870-71.
unattainable information. These conditions, however, are not wholly unique to civil plaintiffs. A criminal defendant squares off against an opponent in possession of critical facts necessary to assessing guilt. Nevertheless, criminal procedure reforms have not attempted to counter this systemic imbalance. Commentators have also observed procedure’s connection to the intended expression of the substantive law. In improving procedural fairness, civil reforms attempted to “employ procedure as a more effective means of implementing substantive law.”

C. Criminal Procedure Reform: The Birth of Disclosure

Federal criminal law has retained fidelity to the pre-modern conception of discovery—neither party is statutorily granted discretion to compel the production of pretrial information. Yet the need for reform is arguably most acute in criminal disputes. Creating a dramatic disparity, the State exercises police powers to compel information through non-transparent methods during an unregulated, pre-complaint period. In 1944, when it was a foreign concept for the State to furnish a criminal defendant with any facts, a new federal rule permitted a defendant the post-complaint, pretrial right to inspect any of his things the government had impounded. Where civil procedure reform had already granted robust investigative powers to litigants during this pretrial period, a criminal defendant was merely afforded the right to inspect what was once his. Subsequent federal reforms did not disrupt a criminal defendant’s passive-recipient status. Rather, additional reforms resulted in a list of disclosures—entitlements to discrete categories of evidence. In 1966, a criminal defendant was granted access to his own statement, his grand jury testimony, and to reports of scientific tests—all disclosures. A defendant was also entitled to documents “material” to presenting a defense—a disclosure intended to “limit the scope of the government’s obligation to search its files while meeting the legitimate needs of the defendant.”

29 Id. at 870.
30 FED. R. CRIM. P. 16 advisory committee’s note to 1944 amendment.
31 FED. R. CRIM. P. 16 advisory committee’s note to 1966 amendment.
32 Id. Subsequent reform entitled defendant to the disclosure of anticipated expert opinion testimony. See FED. R. CRIM. P. 16 advisory committee’s note to 1993 amendment.
Likewise, constitutional rights do not disrupt the passive-recipient status of a criminal defendant. Premised on the idea that access to certain information is essential to due process, the *Brady* right requires a prosecutor to turn over information that is exculpatory (favorable to defendant) and material (consideration of this information would lead to a reasonable likelihood of different outcome). According to one casebook, *Brady*’s obligation “to disclose exculpatory evidence overrides any limitations on discovery provided for by a jurisdiction’s discovery statutes or rules.” This characterization overstates *Brady*’s impact—though a constitutional right indeed overrides any statutory provision that impedes its application, *Brady* is a weak taskmaster. *Brady* confers no discretion to a defendant to compel the production of information he believes relevant to the dispute. *Brady* in fact may not provide a defendant with any pretrial discovery rights at all. *Brady* is arguably available to only those who advance to trial, excluding 90% of defendants who resolve the dispute before trial.

In the remaining instances where *Brady* applies, eligible information must satisfy a demanding test. The prosecutor must determine whether any information in the State’s possession favors a defendant, would be *admissible* at trial, and would have a reasonable likelihood of *changing the outcome*. Under this test, the prosecutor has ample room to undervalue evidence that might otherwise be exploited by a

---

34 RUSSELL, WEAVER ET AL., CRIMINAL PROCEDURE: CASES, PROBLEMS & EXERCISES 888 (3d ed. 2007).
35 United States v. Ruiz, 536 U.S. 622, 633 (2002) (holding *Brady* right does not attach until trial). There is a split regarding whether *Ruiz* applies only to impeachment evidence, or to any and all exculpatory evidence. United States v. Conroy, 567 F.3d 174, 179 (5th Cir. 2009) (rejecting argument that “the limitation of the Court’s discussion [in *Ruiz*] to impeachment evidence implies that exculpatory evidence is different and must be turned over before entry of a plea”); cf. McCann v. Mangialardi, 337 F.3d 782, 788 (7th Cir. 2003) (stating, in dicta, “*Ruiz* indicates a significant distinction between impeachment information and exculpatory evidence of actual innocence. Given this distinction, it is highly likely that the Supreme Court would find a violation of the Due Process Clause if . . . [state] actors have knowledge of a criminal defendant’s factual innocence but fail to disclose such information to a defendant before he enters into a guilty plea”).
37 *Brady*, 373 U.S. at 87-89,
defendant.\textsuperscript{38} She may think that the evidence is favorable but inadmissible. She may believe the evidence admissible, but not likely to change the outcome. A prosecutor may also misapprehend the standard, thinking it narrower than it already is. In a major Wisconsin case, an experienced special prosecutor announced to the circuit court that \textit{Brady} constitutes “evidence that clearly indicates, if you will, the guilt of a third party or absolutely minimizes the guilt of the defendant”\textsuperscript{39}—leaving one to wonder what prosecutor would pursue a case against a defendant where evidence established the guilt of a third party.

\textit{Brady} is also susceptible to being undermined by investigators. Law enforcement may deliver an investigatory file that, unknown to the prosecutor, excludes exculpatory and material evidence. In these circumstances, the prosecutor herself will be unaware of information that should be turned over to the defendant. As to the 10% of those defendants who advance to trial and fall under \textit{Brady’s} purview, \textit{Brady} applies to and benefits only the select few who in post-conviction proceedings manage to find a hidden document that the defendant did not know existed. Such efforts are typically thwarted in jurisdictions that protect the State’s files from open-record requests during the pendency of a direct appeal. Once a defendant has lost his direct appeal, he is likely permitted access to law enforcement files—but he no longer has any right to counsel. Even if such documents are discovered, courts tend to forgive prosecutorial neglect and favor finality.\textsuperscript{40}

Where civil litigants are granted statutory power to compel relevant information from any source, federal rules dictate that a criminal defendant is merely entitled to limited

---


\textsuperscript{39} \textit{See} Brief for Defendant-Respondent at 7-8, State v. Vollbrecht, 820 N.W.2d 443 (Wis. Ct. App. 2012) (No. 2011AP425) (“The case law makes it very clear that the defense is entitled to exculpatory evidence, and there’s a fairly high standard for what that means. It’s evidence that clearly indicates, if you will, the guilt of a third party or absolutely minimizes the guilt of the defendant. We don’t see that sort of evidence in our files.” (internal quotation marks omitted)).

\textsuperscript{40} \textit{See}, e.g., \textit{Smith}, 132 S. Ct. at 630-31. \textit{Smith} split the Court’s conservative wing and strengthened the \textit{Brady} doctrine by providing a \textit{per se} right to a new trial in narrow circumstances. \textit{Id}. Justice Thomas’ dissent reveals how far a judge will go to give the State a pass for its failure to turn over evidence. \textit{Id}. at 640 (Thomas, J., dissenting).
disclosures of State’s evidence.\textsuperscript{41} The federal rules influence a significant number of states.\textsuperscript{42} Of equal influence is the ABA Standard, the alleged liberal bookend to the federal model’s conservative approach.\textsuperscript{43} Unlike the federal rules, the ABA Standard provides for the pretrial disclosure of the prosecutor’s witness list.\textsuperscript{44} More importantly, the ABA Standard requires that disclosures occur immediately, establishing a dynamic that emphasizes the disclosures’ relationship to informing a plea.\textsuperscript{45} In contrast, the federal disclosures are predominantly oriented toward trial preparation, despite trials’ rare occurrence. In the end, however, there is little daylight between the two standards—the ABA Standard does not change the nature of the discovery disparity. Both standards envision the defendant as a passive recipient of information and provide for \textit{prix fixe} menus of State’s evidence—limited disclosures from one party, the State, to its opponent, the defendant.\textsuperscript{46} The ABA Standard does not extend any discretion to a defendant to compel information from multiple sources. The ABA Standard explicitly rejects the idea that the prosecutor should be required to turn over information relevant to the investigation, observing that the

\textsuperscript{41} These observations are drawn from the Federal Rules of Criminal Procedure, the ABA’s standard, and a sampling of civil and criminal procedures adopted by the federal government, as well as ten states that account for more than half of the nation’s population and are geographically diverse: Alabama, California, Florida, Illinois, New York, Ohio, Pennsylvania, Texas, Virginia, and Wisconsin. See \textit{infra} Figure 3. Of the ten states in the sampling, six share significant similarities with the federal rule, whereas three are more closely wedded to the ABA Standard. See \textsc{Yale Kamisar et al.}, \textsc{Modern Criminal Procedure: Cases, Comments, and Questions} 1200-01 (13th ed. 2012) (surveying criminal discovery nationwide).

\textsuperscript{42} \textsc{Charles H. Whitebread & Christopher Slobogin}, \textsc{Criminal Procedure: An Analysis of Cases and Concepts} 671 (5th ed. 2008) (discussing the adoption of Rule 16 of the Federal Rules of Criminal Procedure and “the proliferation of similar rules at the state level”); Lissa Griffin, \textit{Pretrial Proceedings for Innocent People: Reforming Brady}, 56 N.Y.L. Sch. L. Rev. 969, 980-81 n.69 (2011) ( remarking that as to the \textit{Brady}-based language of Fed. R. Crim. P. 16, the state must turn over evidence “material to the preparation of the [defendant’s] defense”); Roberts, \textit{supra} note 3, at 1122 (stating that almost a fourth of the states adopt the federal standard).

\textsuperscript{43} ABA STANDARDS, \textit{supra} note 3; see also Roberts, \textit{supra} note 3, at 1122 (noting that the ABA standard has influenced roughly a quarter of states).

\textsuperscript{44} ABA STANDARDS, \textit{supra} note 3, at §§ 2.1(a)(i), (iii) (providing for disclosure of witness lists and for “those portions of grand jury minutes containing testimony of the accused and relevant testimony of persons whom the prosecuting attorney intends to call as witnesses at the hearing or trial” respectively). Given the infrequent use of the grand jury, this latter requirement would have little practical effect. Even if it did apply, the prosecutor would be able to make strategic adjustments that would minimize any disclosures required under this standard.

\textsuperscript{45} \textit{Id.} at § 5.1(a)(i–iii) (identifying three particular pretrial stages relevant to discovery obligations); \textit{id.} at § 5.3(a)(ii) (providing for judicial hearing to ensure compliance with discovery obligations during initial exploratory stage).

\textsuperscript{46} The most meaningful differences are that under the ABA standard the prosecution must turn over a witness list, and the timing of disclosures. \textit{Id.} at § 2.1(a)(1).
relevancy standard “seems too far outside the mainstream of current American practice and expression.” Yet, the relevancy standard has governed the exchange of discovery in civil litigation since 1938.

D. Disclosures versus Formal Powers of Investigation

A formal investigatory tool permits a litigant the discretion to compel the production of information—she decides what source is potentially significant and what information she will seek. In order to refuse to comply with a formal request for information a responding party must persuade a court that the requested information has no potential to lead to admissible evidence. Refusals to at least partially comply with a request are accordingly rare. A responding party will invariably make objections, but these in operation serve to protect against any subsequent allegation that the response is not adequate. Objections tend to be long-winded—comments like “hopelessly overbroad,” “unapologetically vague and ambiguous,” “truly burdensome,” “mocks rules of grammar,” or “a crushing blow for humanity” might, for example, all appear in one exuberant sentence. Through this noise, the sought after document or witness is typically produced.

A litigant only entitled to disclosures, however, has no discretion to make an independent inquiry into the universe of facts that might inform liability. A statute or opponent binds that discretion. A statutorily defined disclosure might require the State to turn over to defendant any statement he made to police. The defendant cannot exceed that particular statutory constraint and request statements made by others that reference the defendant. An opponent-defined disclosure might require the prosecutor to turn over any document she intends to use at trial. Discretion sits with the prosecutor; she may elect to turn over nothing. And if she plans to use documents at trial, they will favor the State, like photos of wounds. In jurisdictions influenced by federal

47 Id. at § 2.1(d) (exculpatory material).
48 See, e.g., Fed. R. Civ. P. 26(b)(1) (“Parties may obtain discovery regarding any nonprivileged matter that is relevant to any party’s claim or defense—including the existence, description, nature, custody, condition, and location of any documents or other tangible things and the identity and location of persons who know of any discoverable matter.”).
49 See, e.g., Fed. R. Crim. P. 16(a)(1)(A), (B), (D).
constraints,\textsuperscript{51} pretrial disclosures are so limited that a defendant has no discretion to obtain witness lists, police reports, or names of investigating detectives.\textsuperscript{52} Further narrowing the significance of disclosures, a defendant is only authorized to obtain information from one source, the State.\textsuperscript{53}

Disclosures granted in criminal law typically constitute the beginning and end of statutorily permitted discovery.\textsuperscript{54} In contrast, civil procedure statutes that provide for disclosures serve a different purpose.\textsuperscript{55} In federal court, the mandatory disclosures civil litigants make at the lawsuit’s inception are intended to “accelerate the exchange of basic information,” “focus the discovery that is needed,” and “guide further proceedings in the case.”\textsuperscript{56} Though scholarship has recognized a discovery disparity between civil and criminal litigants,\textsuperscript{57} what is most significant is the disparity’s nature—a criminal defendant is a passive recipient of information whereas a civil litigant exercises discretion to compel information from multiple sources. Based on a sampling of jurisdictions, the following table underscores how a criminal defendant depends on disclosures from one party, the State. In contrast, though

---


\textsuperscript{52} California, Illinois, and Ohio are states in the sampling that require the State to turn over the witness list. Cal. Penal Code § 1054.1(a) (West 2013); Ill. Sup. Ct. R. 412(a)(i); Ohio R. Crim. P. 16(b). California and Ohio provide for disclosure of exculpatory information. Cal. Penal Code § 1054.1(e) (West 2013); Ohio R. Crim. P. 16(B)(5). Florida and Ohio require disclosure of all police reports and Florida requires disclosure of contact information of witnesses and the interviewing detectives. Fla. R. Crim. P. 3.220(b)(1)(A)-(B); Ohio R. Crim. P. 16(B)(1).


\textsuperscript{54} See infra Figure 3. Typically, jurisdictions in the sample only provide for disclosures at the request of defendant. The vast majority of discovery available to the criminal defendant is not mandatory, but only occurs via request.

\textsuperscript{55} See infra Figure 3; Fed. R. Civ. P. 26(a)(1)(A)(i)(ii) (requiring disclosure of individuals likely to have information and certain documents); Ill. Sup. Ct. R. 213(f) (requiring disclosure of witness information if requested); N.Y. C.P.L.R. 3101(d) (McKinney 2012) (requiring disclosure, if requested, of information pertaining to expert witnesses); Pa. R. Civ. P. 4003.4 (allowing discovery of statements from parties, nonparties, and witnesses that pertain to the action); Tex. R. Civ. P. 194.1.2 (providing wide range of disclosure, including names of those with relevant information).

\textsuperscript{56} Fed. R. Civ. P. 26 advisory committee’s note to 1993 amendment.

\textsuperscript{57} Robert L. Fletcher, Pretrial Discovery in State Criminal Cases, 12 Stan. L. Rev. 293, 294 (1960) (noting “a long and deeply imbedded practice designed to keep the defendant in the dark as long as possible”); Jean Montoya, A Theory of Compulsory Process Clause Discovery Rights, 70 Ind. L.J. 845, 855-56 (1995) (noting that no criminal discovery procedures match the broad discovery possibilities in civil procedure); Roberts, supra note 3, at 1098 (noting disparity between discovery rights afforded to civil and criminal litigants).
civil litigants sometimes are provided disclosures to seed investigations, they are as a matter of course permitted discretion to compel information from any source.\(^56\)

**Figure Three\(^{59}\): Discovery Mechanisms: Civil Litigant**

<table>
<thead>
<tr>
<th></th>
<th>Fed</th>
<th>AL</th>
<th>CA</th>
<th>FL</th>
<th>IL</th>
<th>NY</th>
<th>OH</th>
<th>PA</th>
<th>TX</th>
<th>VA</th>
<th>WI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disclosures</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Interrogatories</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Depositions</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Documents from Parties</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Documents from Non-Parties</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**Discovery Mechanisms: Criminal Defendants**

<table>
<thead>
<tr>
<th></th>
<th>Fed</th>
<th>AL</th>
<th>CA</th>
<th>FL</th>
<th>IL</th>
<th>NY</th>
<th>OH</th>
<th>PA</th>
<th>TX</th>
<th>VA</th>
<th>WI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disclosures</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Interrogatories</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Depositions</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Documents from Parties</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Documents from Non-Parties</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Not afforded any formal investigatory powers, a criminal defendant is not assigned an investigatory role. The typical statute permits him to passively receive pretrial evidence from one source. Collected and summarized by his

---

\(^{56}\) See infra Figure 3.

\(^{59}\) (X - Broad Discovery Right) (/ - Limited Discovery Right). There is no bright line test in determining what discovery rights afforded are “broad” versus “limited.” It is a comparative analysis.
opponent, this evidence is weighted against him. A defendant may conduct an informal investigation, but this method is not well suited for the high-stakes, limited-source scenario that tends to define criminal cases. As the sole party assigned an investigatory role, the State uses formal powers to develop facts and establish the narrative. What law enforcement turns over to the prosecutor is typically a subset of the investigative file. And what a prosecutor turns over to a defendant is a subset of the prosecutorial file. This limited disclosure will constitute “the facts of the case” for a defendant who is unable to extract further information through informal means. The conditions that triggered reform to civil procedure are thus present in criminal law; a criminal defendant is situated against a motivated opponent who controls the facts. Without the formal power to develop a counter-narrative, a criminal defendant is consequently subject to an adversary who controls all aspects of pretrial litigation.

II. WHAT IS LOST FOR CRIMINAL DEFENDANTS

Although the formal investigatory tools exercised by civil litigants may not be tailored to the criminal law, an exploration of what it means to have such tools—depositions, document requests, interrogatories—reveals what is lost in their absence.

A. Formal Investigatory Tool: Depositions

All 50 states authorize civil litigants to depose witnesses. By this power, an attorney may compel any person to appear and answer questions under oath. Questions are not constrained by rules of evidence that apply at trial. By design, depositions permit inquiry into inadmissible hearsay (“you heard Joey say that Frank hated the victim?”), other acts (“you heard that the Frank had been convicted for battery six months

---

60 See supra Figure 1.
61 See supra Part I (noting that disclosures only provide limited access to the State’s, and only the State’s, evidence).
62 Langer, supra note 5, at 250 (stating in many cases the prosecutor, in control of the evidence, successfully plays the role of sole adjudicator in plea negotiations); Lynch, supra note 9, at 2120; Natapoff, supra note 12, at 968 (stating “the investigative sphere is the most powerful adjudicative arena, in which police and prosecutorial decisions about information and potential liability determine the circumstances under which individuals must confront the coercive powers of the state”).
before the crime?”), and character evidence (“is Frank easily provoked to violence?”); all are ingredients to an effective investigation. In some jurisdictions there are no time limits to the deposition. Even if limited by time, a deposition frees an attorney from the pressure of eliciting in-court testimony from a sophisticated and recalcitrant witness. The cost of witness prevarication is shifted from the questioning party to the witness and opposing party. Absent are the theatrics that tend to undermine the legitimacy of trial. In trial, a witness can play dumb and take a long time to refresh his recollection or understand the meaning of a simple passage. This tactic interrupts tempo. The trier of fact’s interest wanes during periods of silence. In a deposition, such tactics are mere antics—the requesting party will note on the record the inordinate time the witness is taking to answer questions and, if necessary, will seek sanctions and secure permission to extend time to finish the deposition. The deposition is designed to collect a broad understanding of motives and facts that define a dispute. In operation, a deponent will be compelled to answer most questions. Any objection to a question that does not invoke a privilege typically accomplishes little—the witness must eventually answer. An attorney may attempt to suspend a deposition because the asking party is badgering the witness or going beyond the scope of permissible questions. But hell hath no looks of annoyance like a judge drawn into a petty discovery dispute. The civil deposition power is permissive in theory and unrestrained in fact. This liberal application of the

64 David Young, A New Theory of Relativity: The Triumph of the Irrelevant at Depositions, 36 UWLA L. REV. 56, 59 (2005) (“The concept of relevance is still the primary focus at depositions in determining the permissible scope of discovery.”).

65 Florida, New York, Ohio, Pennsylvania, and Wisconsin statutes place no time limits on depositions. See supra Figure 3; see also FLA. R. CIV. P. 1.310; N.Y. C.P.L.R. 3106 (McKinney 2013); OHIO R. CIV. P. 30; PA. R. CIV. P. 4007.1; WIS. STAT. ANN. § 804.05 (West 2010). Federal Rules, Alabama, California, Illinois, Texas, and Virginia place time limits on depositions. FED. R. CIV. P. 30 (seven hours); ALA. R. CIV. P. 30 committee comments to Aug. 1, 2004 amendment (five hours per day); CAL. CIV. PROC. CODE § 2025.290 (West 2013) (seven hours); ILL. SUP. CT. R. 206(d) (three hours); TEX. R. CIV. P. 199.5(c) (six hours, though case and corresponding discovery level may demand a longer period of time); VA SUP. CT. R. 4:5(b)(3) (allowing court discretion to set time limitation).

66 Although the Federal Rules of Civil Procedure dictate that “[t]he examination and cross-examination of a deponent proceed as they would at trial under the Federal Rules of Evidence,” there are distinctions between trial practice and depositions. FED. R. CIV. P. 30(c)(1). Objections during depositions can be made and noted, but the deponent must respond unless the objection relates to the need to preserve a privilege or enforce a court order. FED. R. CIV. P. 30(c)(2).

67 FED. R. CIV. P. 37(a)(5)(A) (a person subject to a successful motion to compel disclosure faces the prospect of paying the “movant’s reasonable expenses
deposition power is not a given. A deposition, like any investigatory tool, may be tailored to address concerns particular to any type of dispute.

Four states extend deposition power to criminal defendants in a manner approaching equivalence to the civil deposition: Vermont, Missouri, Indiana, and Florida. These states permit parties to depose broad categories of individuals, police officers and victims included. In New Mexico, parties may issue a pretrial subpoena and take a recorded statement—an affordable “dirty deposition” subject to wide use, more cost-effective than a traditional deposition, and a tool that demonstrates how innovations to formal investigatory tools might respond to concerns particular to the criminal justice system. Even in states that nominally allow depositions to criminal defendants, some require a defendant to make the formidable showing that the witness is “material and necessary.” In more restrictive jurisdictions, a defendant may only petition the court to take a deposition to preserve testimony—like that of a key witness on her deathbed. The

incurred in making the motion, including attorney’s fees”); W. R. Grace & Co. v. Pullman, Inc., 74 F.R.D. 80, 84 (W.D. Okla. 1977) (“The harm caused by being required to take additional depositions of a witness who fails to answer a question based on an improperly asserted objection far exceeds the mere inconvenience of a witness having to answer a question which may not be admissible at the trial of the action.”); Banco Nacional De Credito Ejidal v. Bank of Am. N.T. & S.A., 11 F.R.D. 497, 499 (N.D. Cal. 1951) (“Basically the propriety of probing any matter within the knowledge of deponent is dependent upon relevancy—and relevancy, especially at the pre-trial stage, is very liberally construed.”).

Interview with Katherine Judson, Innocence Project Litigation Fellow, in Madison, Wis. (Oct. 23, 2012) (“Dirty deposition” is the nom de guerre assigned by author); see also, Rule 5-503 of New Mexico District Court Rules of Criminal Procedure (“Any person, other than the defendant, with information which is subject to discovery shall give a statement. A party may obtain the statement of the person by serving a written ‘notice of statement’ upon the person to be examined and upon each party not less than five (5) days before the date scheduled for the statement. The notice shall state the time and place for taking of the statement.”).

See, e.g., MONT. CODE ANN. § 46-15-201 (West 2013) (stating a deposition of a prospective witness may be taken if the witness will be unable to attend trial or the witness “is unwilling to provide relevant information to a requesting party and the witness’s testimony is material and necessary in order to prevent a failure of justice”).

ALA. R. CRIM. P. RULE 16.6; ALASKA R. CRIM. PROC. 15; ARK. CODE ANN. § 16-44-202 (West 2013); COLO. CRIM. P. 15; CONN. GEN. STAT. ANN. § 54-86 (West 2013); DEL. SUPER. CT. CRIM. R. 15; GA. CODE ANN. § 24-13-130 (2013); HAW. R. PENAL P. 15; IDAHO CRIM. R. 15; ILL. SUP. CT. R. 414; KY. R. CRIM. P. 7.10; ME. R. CRIM. P. 15;
federal statute is one of these jurisdictions, providing that a “party may move that a prospective witness be deposed in order to preserve testimony for trial. The court may grant the motion because of exceptional circumstances and in the interest of justice.”74 In these jurisdictions, the deposition is explicitly not intended to be a pretrial discovery tool but is instead oriented toward the trial moment. Remaining jurisdictions deny deposition power by omission (the lack of any provision granting defendant the right to request one).75

B. Formal Investigatory Tool: Production of Documents

Civil litigants have the pretrial power to request the production of documents from the opposing party76 that are relevant to any claim or defense.77 A responding party must make a “reasonable effort to assure that the client has provided all responsive information and documents available to him . . . .”78 Federal criminal procedure, too, purports to make documents and objects “subject to disclosure.”79 These procedures, however, have very different DNA. Federal criminal procedure does not afford a defendant any discretion to compel documents of a defendant’s choosing, but rather designates distinct categories of documents subject to disclosure. The first category requires the State turn over documents it intends to use “in its case-in-chief at trial”—any documents revealed by this disclosure will favor the State’s case. The second category requires the disclosure of any item that was “obtained from or belongs to the defendant,” a tell-me-what-I-already-know disclosure.81 The last category provides that the


75 Criminal procedure statutes in Louisiana and Virginia do not address depositions.
76 See e.g., Fed. R. CIV. P. 34; ALA. R. CIV. P. 34(a); CAL. CIV. PROC. CODE § 2031.010 (West 2013); FLA. R. CIV. P. 1.350(a); ILL. SUP. CT. R. 214; N.Y. C.P.L.R. 3120 (McKinney 2013); OHIO R. CRIV. P. 34(A); PA. R. CIV. P. 4009.1; TEX. R. CIV. P. 196.1(a); VA. SUP. CT. R. 4-9(a); WIS. STAT. ANN. § 804.09 (West 2013).
77 Fed. R. CIV. P. 26(b); see also Fed. R. CIV. P. 34.
78 Fed. R. CIV. P. 26 advisory committee’s notes to 1983 Amendment.
79 Fed. R. CRIM. P. 16(a)(1).
State must turn over items “material to preparing the defense.”82 Some courts maintain the Brady standard does not govern this provision83 whereas others look to Brady for guidance.84 The debate underscores the cautious nature of the provision’s language. In determining what is “material” to the defense, the State is prone to undervalue evidence helpful to the defendant.85 These three categories permitting limited disclosures from a single source fall well short of the formal power and discretion to request documents that are extended to civil litigants.

Providing more robust disclosure rights for a criminal defendant than either the federal or ABA standards, Florida is again an outlier. Florida requires that, upon request, a prosecutor turn over all investigative reports.86 Florida's statute does, however, protect notes of investigators from disclosure.87 Yet these notes provide an unedited version of what later is altered or omitted in the resultant police report. The potential significance of an agent’s notes is underscored in the United States Supreme Court’s Smith v. Cain decision.88 Eyewitness Larry Boatner implicated Defendant Smith in a New Orleans shooting. Boatner testified that gunmen entered his friend’s home and began shooting.89 At trial, Boatner identified Smith as a shooter. After trial, the defense learned of a detective’s notes that stated “Boatner ‘could not . . . supply a description of the perpetrators other then [sic] they were black males.’”90 This case demonstrates how the non-transparent method employed by the State provides an opportunity for

83 See supra notes 33-38, for a discussion of Brady. The Brady right only applies to admissible evidence; it does not provide for any right to investigate, but rather is animated by the much narrower concept of due process. Brady v. Maryland, 373 U.S. 83, 87 (1963). Under Brady, the “materiality” standard is rigorous; a document is only material if it has a reasonable probability of changing the outcome. Kyles v. Whitley, 514 U.S. 419, 435 (1995).
84 Robert M. Cary et al., Federal Criminal Discovery 95 (2011) (“Courts sometimes equate the Rule 16(a)(1)(E) materiality standard with the Brady rule, which also has a materiality component. Other courts have disagreed, and rightly so.” (footnotes omitted)).
85 See supra Part I.C. (discussing Brady and issues surrounding the exculpatory and material standard); see also Findley & Scott, supra note 38, at 351 (“Brady demands too much of prosecutors when it simultaneously asks them to act as advocates charged with prosecuting a defendant and as neutral observers responsible for assessing the value of evidence from the defendant’s perspective.”).
86 Fla. R. Crim. P. § 3.220(b)(1).
87 Fla. R. Crim. P. § 3.220(b)(1)(B).
89 Id. at 629-30.
90 Id. at 629. The State also failed to disclose Boatner’s statement that he “could not ID anyone because [he] couldn’t see faces’ and ‘would not know them if [he] saw them.’” Id. at 630.
abuse of power—the detective should not have omitted such evidence from the police report. These notes in the Smith case would not be discoverable under Florida’s statute. Without statutory protection, a defendant must rely on Brady. But the Brady right only attaches once a defendant advances to trial. For the 90% of criminal defendants who enter pleas, the State would potentially be given impunity for its officers engaging in such behavior.

Civil litigants are authorized to subpoena documents from non-parties for the purpose of conducting a pretrial investigation.\textsuperscript{91} A criminal defendant may not typically use a third-party subpoena (subpoena \textit{duces tecum}) as a pretrial investigatory tool.\textsuperscript{92} Rather, the subpoena \textit{duces tecum} in criminal disputes is usually limited to expediting any trial or evidentiary hearing by requiring the disclosure of documents close in time to the upcoming hearing.\textsuperscript{93} Even then, the standard for its use is typically stringent. Under federal law, a criminal defendant must in a court proceeding demonstrate: (1) each item sought is likely admissible, (2) “not otherwise procurable through due diligence prior to trial,” (3) that defendant would be unable to “properly prepare for trial without such [pre-trial] production and

\textsuperscript{91} The Federal Rules of Civil Procedure as well as each state in Fig. 3 allow for civil litigants to obtain documents and things from both parties and nonparties. \textit{Fed. R. Civ. P.} 34, 45(c); \textit{Ala. R. Civ. P.} 34(a), 45(a)(3); \textit{Cal. Civ. Proc. Code} \S\S 2031.020(b), 2020.410, 2025.280(b) (West 2013); \textit{Fla. R. Civ. P.} 1.350(b), 1.351(a), 1.410(c); \textit{Ill. Sup. Ct. R.} 214; \textit{N.Y. C.P.L.R.} 3120, 3111; \textit{Ohio R. Civ. P.} 34(A)-(C), 45(A)(1)(b)(iii)-(vi); \textit{Pa. R. Civ. P.} 4009.1, 4009.12(a)-(1)-(2), 4009.21(a), 4009.23(a); \textit{Tex. R. Civ. P.} 196.1(a), 196.2(a), 205.1(c)-(d), 205.3(a); \textit{Va. Sup. Ct. R.} 4:9(a)-(b), 4:9A(a)-(b); \textit{Wis. Stat.} \S\S 804.09(1)-(3), 805.07(2)(a) (2013).

\textsuperscript{92} In jurisdictions surveyed in Figure 3, a federal criminal defendant, as well as a criminal defendant in Alabama, Ohio, Illinois, New York, Pennsylvania, Wisconsin, and Texas may not use the subpoena \textit{duces tecum} to conduct a pretrial investigation. \textit{See}, e.g., \textit{Fed. R. Crim. P.} 17(c)(1) (under \textit{Bowman Dairy Co. v. United States}, 341 U.S. 214, 220 (1951), the subpoena \textit{duces tecum}’s purpose "was not intended to provide an additional means of discovery. Its chief innovation was to expedite the trial by providing a time and place before trial for the inspection of the subpoenaed materials"); \textit{Ala. R. Crim. P.} 17.3 (stating in statute’s commentary that “[t]his rule is not intended to be a discovery device”); \textit{Ohio R. Crim. P.} 17(c) (according to \textit{In re Subpoena Dues Tecum Served Upon Attorney Potts}, 796 N.E.2d 915, 918 (Ohio 2003), use of the subpoena \textit{duces tecum} is restricted to the trial moment—and any effort to use it in advance of trial requires a showing that the requested document(s) will be admissible); \textit{Pa. R. Crim. P.} 107 (comments stating that subpoena only used for hearings or trial); \textit{People v. Hart}, 552 N.E.2d 1, 2 (Ill. App. Ct. 1990) (defendant may not use subpoena \textit{duces tecum} as a discovery tool); \textit{People v. Magliore}, 679 N.Y.S.2d 267, 270 (N.Y. Crim. Ct. 1998) (purpose of subpoena \textit{duces tecum} is limited to compelling "specific documents that are relevant and material to facts at issue in a pending judicial proceeding"); \textit{Cruz v. State}, 838 S.W.2d 682, 686 (Tex. App. 1992) (consistent with federal rule); \textit{State v. Schaefer}, 746 N.W.2d 457, 461 (Wis. 2008) (holding a defendant may not use subpoena \textit{duces tecum} as a discovery tool).

inspection,” and (4) that the request “is made in good faith.”

A criminal defendant in more liberal jurisdictions will still encounter significant hurdles, like requirements that a subpoena duces tecum require a court hearing in which defendant must demonstrate a material need for the requested documents.

C. Formal Investigatory Tool: Interrogatories

Interrogatories—written questions to secure investigative leads—are valuable at a dispute’s inception; a party can require the opponent to list facts in support of the opponent’s allegations, along with the identity of individuals with information and documents that provide the basis for those assertions. Responses must “represent the collective knowledge of the opponent.”

Granted to civil litigants, interrogatories are not extended to either the prosecution or the defense in jurisdictions influenced by federal and ABA standards. Florida again distinguishes itself; the equivalent of a “form interrogatory” (a list of judicially or legislatively authorized questions that in civil litigation may be propounded and are not subject to any objection) is embedded in its statute. Florida requires the prosecutor to disclose “a list of the names and addresses of all persons known to the prosecutor to have information that may be relevant to any offense charged or any defense thereto, or to any similar fact evidence to be presented at trial,” an evidentiary category that captures

---

94 Id.
95 In jurisdictions surveyed in Figure 3, to use a subpoena duces tecum in Virginia and Florida, court intervention is required and a “materiality” standard must generally be satisfied. FLA. R. CRIM. P. § 3.220(f) (“[i]n a showing of materiality, the court may require such other discovery to the parties as justice may require”); VA. SUP. CT. R. 3A:12(b) (requiring that requesting party include an affidavit “that the requested writings or objects are material to the proceedings,” and in Commonwealth v. Faulkner, 82 Va. Cir. 417, 422-23 (Va. Cir. Ct. 2011), the court stated that subpoena duces tecum’s use for limited discovery is permissible); Millaud v. Superior Court, 227 Cal. Rptr. 222, 224 (Cal. Ct. App. 1986) (stating that “the lack of specific statutory authority for the procedure is not conclusive” and that the court has “an inherent power ‘to develop rules of procedure aimed at facilitating the administration of criminal justice and promoting the orderly ascertainment of the truth’”).
96 Edward L. Miner & Adrian P. Schoone, The Effective Use of Written Interrogatories, 60 MARQ. L. REV. 29, 30 (1976) (“Interrogatories are often preferable to depositions for identifying such things as witnesses, documents, the dates and substance of transactions and conversations.”).
97 Id. at 29.
98 See, e.g., FED. R. CIV. P. 33; ALA. R. CIV. P. 33; CAL. CIV. PROC. CODE § 2030.030 (West 2007); FLA. R. CIV. P. 1.340(a); ILL. SUP. CT. R. 213; N.Y. C.P.L.R. 3130 (McKinney 2013); OHIO R. CIV. P. 33; PA. R. CIV. P. 4005(a); TEX. R. CIV. P. 197; VA. SUP. CT. R. 4.8(a); WIS. STAT. § 804.08 (2013).
eyewitnesses, alibi witnesses, investigating officers, and witnesses the prosecutor does not intend to call.\textsuperscript{100}

\textbf{D. In the Neighborhood: What it Would Mean to Have Formal Investigatory Tools}

How would having access to the formal investigatory tools available to a civil litigator—interrogatories, document requests, depositions—impact a criminal case? Would doing so reduce false positives, increase accuracy in outcomes, gain efficiencies, or improve conceptions of procedural justice and prosecutorial integrity? A recent case involving the homicide of Rodolfo Jimenez in Racine, Wisconsin provides a starting point to examine these questions.\textsuperscript{101} Shawn Milton was tried for shooting Jimenez. Milton’s trial, which took only two days, resulted in a guilty verdict and life imprisonment. At trial, Milton’s defense counsel did not call a single witness. In a well-resourced post-conviction inquiry that resulted in a new trial, the appellate team discovered information that under typical rules of criminal procedure would have remained undisclosed. The appellate team, for example, secured the State’s investigative file through open-record requests, a method unavailable before trial. Every piece of information discovered in the post-conviction inquiry would have been subject to pretrial production had Milton had the investigatory tools available to civil litigants.

At 10 PM on January 16, 2006, Rodolfo Jimenez was gunned down in the street. From an apartment window, an eyewitness observed two individuals facing Jimenez run south. No weapon or bullet casings were found. Two years later an individual told police he had heard that seven individuals were involved in the Jimenez shooting. He named, among others, Milton. Canvassing the neighborhood generated more heat than light—individuals who did not witness the crime nonetheless implicated various people. Because typical discovery statutes provide for little discovery, these complexities and contradictions typically remain unknown to the defense after a prosecutor presents charges in such a case.

One of the young men implicated was Matthew Roth. Police questioning of Roth resulted in a detailed written statement. According to his statement, Roth approached “a

\textsuperscript{100} \textit{Id.}

\textsuperscript{101} An actual case, names of individuals have been changed, as well as dates and other identifying details. A redacted copy of the appellant’s brief is on file with the author and the \textit{Brooklyn Law Review}. 
Mexican” for a cigarette and said, “Dame un cigarillo.” Jimenez responded, “Fuck you, get a job.” Roth heard a gunshot. Jimenez doubled over. Roth looked back to see a gun in Milton’s hand. After signing the statement on every page—a tactic to aid the prosecutor at trial (for impeachment purposes, a prosecutor will make it clear to a jury that the witness signed each and every page)—Roth was released from custody.

The State filed charges against Milton. Milton sat in custody for seven months. A criminal defense attorney would typically receive little pretrial information during this time. But, a civil litigator wielding formal investigatory powers would immediately serve interrogatories (accompanied by a request for documents):103

1. State all facts that support allegations in the Complaint, describing documents and providing contact information of individuals who have information supporting these facts.

2. Provide contact information of suspects and of those implicated in the shooting, describing documents that reference these individuals.

3. Provide a description of all items collected in the investigation, along with all forensic documents.

4. Provide contact information of all individuals interviewed by law enforcement in the investigation, describing all related documents.

From this first formal investigatory act, the civil litigator would learn of Roth’s written statement implicating Milton. In any jurisdiction following the federal rules, a prosecutor would not turn this statement over until after Roth testified at trial.104 The civil litigator’s initial round of discovery would also reveal that Roth implicated Milton only after police subjected Roth to a 10-hour interrogation in the middle of the night. A civil litigator would file additional discovery requests

---

102 In the Milton trial, Roth started to change his story—the prosecutor blew up every page of his statement and surrounded Roth with a life-sized version of his statement.

103 These interrogatories are compressed and do not follow the traditional format.

104 In the Milton case, the prosecutor did turn this statement over to the defense before being obligated to do so; the prosecutor was only required to turn it over a few weeks before trial. Wis. STAT. ANN. §§ 971.23(1), 971.23(1)(e) (West 2011). Under federal law, the prosecutor would not have to turn this statement over until after Roth testified at trial. 18 U.S.C. § 3500 (2012).
(discovery begets discovery), and, as the massive post-conviction investigative file illustrates, such efforts likely would have uncovered the following before trial:

1. Reports indicating detectives picked up Roth and told him he was a suspect for Jimenez’s murder.

2. Reports that at the beginning of the 10-hour interrogation, Roth denied any knowledge of the shooting.

3. Reports that during the 10-hour interrogation, officers falsely suggested to Roth that Milton had implicated Roth in the shooting. Reports indicate that only after hearing this falsehood did Roth implicate Milton.

4. Reports that on the same block, just six weeks before the Jimenez shooting, Roth shot at an unarmed man. The man barely escaped (Roth shot out the victim’s back window; a bullet had lodged in the front passenger seat). Roth was not prosecuted—perhaps, as a civil litigator might have learned, because Roth’s sister was a Racine police officer.

5. Reports that Roth was found in possession of a large cache of ammunition and a handgun before and after the Jimenez shooting.

In the typical jurisdiction, these documents would not be subject to pretrial disclosure—in post-conviction proceedings in the Jimenez case, for example, the prosecutor argued he had no obligation to turn over reports of Roth’s prior shooting.

To guide deposition choices, a civil litigator will use a key document like Roth’s written statement. A civil litigator would attempt to undermine or confirm the statement’s representations by deposing detectives who interrogated Roth, any person mentioned in the statement, and Roth. According to Roth’s statement, Dante Randall cut Roth’s hair on Randall’s front porch on the day of the shooting. But meteorological data indicated it was 23 degrees below freezing that day, and the house described by Roth on Green Street turned out to have an open porch—hostile conditions for a haircut.

By informal investigative means, it took the appellate team three months to persuade Roth’s alleged barber, Randall, to meet at Burger King. He entertained questions for 20
minutes—an insufficient amount of time in an inappropriate venue to discuss police reports, neighborhood history, any criminal history, and a homicide. In a pretrial context, having the fortune to secure a critical witness’ momentary attention does not qualify as a discovery plan. Randall’s reluctance to meet would likely be more acute at a pretrial stage, given the warranted fear in neighborhoods like his that knowledge about a crime makes one a suspect. But a civil litigator, compelling Randall’s attendance, would have found that Randall did not cut white people’s hair (Roth is white) and would not cut hair on an open porch in winter. At a deposition, the civil litigator would learn that Randall was not even certain that he lived on Green Street in January of 2006. This uncertainty was significant. Roth had stated that after Randall gave him a haircut, Roth walked to the rear garage of Randall’s Green Street home and observed Milton with the same revolver he observed in Milton’s hand at the shooting. The State argued that this fact showed Milton had the opportunity to shoot Jimenez. But if Randall had not lived on Green Street, then Roth’s recollection of events was mistaken or a work of fiction.

A civil attorney would issue subpoenas to compel production of documents from third parties. Criminal defendants typically do not have this pretrial investigatory power. Formal discovery tools available to a civil attorney would have required the Green Street home’s owner to search boxes in her attic—her receipts indicated Randall had moved out in September 2005, four months before the shooting. A civil attorney would compel the utility company to release information—she would learn that in September, Randall’s bill had been transferred to a house on Lakeside Avenue. A subpoena to the Lakeside Avenue owner for his rental records would have confirmed Randall’s move. A visual inspection of the Lakeside home indicated that it did not have a front porch or a rear garage. Roth had manufactured a significant fact.

105 In jurisdictions surveyed in Figure 3, a federal criminal defendant, as well as a criminal defendant in Alabama, Ohio, Illinois, New York, Pennsylvania, Wisconsin, and Texas may not use the subpoena duces tecum to conduct a pretrial investigation. See Fed. R. Civ. P. 34, 45(c); Ala. R. Civ. P. 34(a), 45(a)(3); Cal. Civ. Proc. Code §§ 2031.020(b), 2020.410, 2025.280(b) (West 2013); Fla. R. Civ. P. 1.350(b), 1.351(a), 1.410(c); Ill. Sup. Ct. R. 214; N.Y. C.P.L.R. 3120, 3111; Ohio R. Civ. P. 34(A)-(C), 45(A)(1)(b)(iii)-(vi); Pa. R. Civ. P. 4009.1, 4009.12(a)(1)-(2), 4009.21(a), 4009.23(a); Tex. R. Civ. P. 196.1(a), 196.2(a), 205.1(c)-(d), 205.3(a); Va. Sup. Ct. R. 4:9(a)-(b), 4:9A(a)-(b); Wis. Stat. § 804.09(1)-(3); 805.07(2)(a) (2013); United States v. Nixon, 418 U.S. 683, 699-700 (1974).
The surface was scratched. A civil attorney, however, would go further than merely undermining the credibility of the State’s key witness; she would use formal investigatory powers in search of an alternative theory of liability. To establish a link between the two shooting incidents, she would depose eyewitnesses to the shooting Roth had committed just weeks before the Jimenez incident. She would discover that Roth needed little provocation to unleash a stream of bullets—that Roth emerged from the alley and shot into the victim’s car because the victim verbally defended his girlfriend against insults hurled by Roth’s friends.

A civil attorney would also depose Antonio Hernandez, who was mentioned in police reports but never questioned by police. Under informal investigatory methods, it took the appellate team months to arrange a meeting. Hernandez’s reluctance to cooperate was attributable in part to his status as a confidential informant. If he turned against the State, the State could deem him a liar and reinstate drug charges against him. In providing any information exculpatory to Milton, Hernandez faced the prospect of losing his union job and going to prison. A civil attorney, however, would be entitled to compel Hernandez’s attendance and testimony—and would learn that, on the night of the shooting, Roth had confessed to killing Jimenez.106

After a yearlong investigation, Milton’s appellate team presented 10 witnesses and 50 exhibits in an eight-day evidentiary hearing. Confronted with new facts, Roth claimed his right against self-incrimination and refused to answer. Moved by a cohesive narrative that suggested that Roth, and not Milton, was the perpetrator, the court granted Milton a new trial. Months later Milton pled no contest to greatly reduced charges that capped incarceration at five years. He walked free on a “time-served” sentence. Had Milton been afforded pretrial power to compel documents and testimony, information developed by the appellate team’s investigation would have emerged before trial, not after Milton’s conviction. On the one hand, greater resources (and time) helped compensate for the absence of any formal investigatory tools. On the other hand, more resources did not compensate for the absence of the formal power to compel information. Key witnesses undermined the appellate investigation through

106 When the prosecutor heard about Hernandez’s statement, he deemed Hernandez a liar, and stated his value as a confidential informant (CI) was now worthless.
their refusal or reluctance to cooperate. The Jimenez post-conviction investigation suggests: (1) that the power to compel pretrial attendance and testimony, along with documents that should be subject to scrutiny, would be significantly more efficient than conducting an informal investigation, (2) that the appellate team would have found critical information with the aid of formal investigatory powers, and (3) that the idea that the defense has nothing to offer as a party to the investigation overestimates the likelihood that the State got the case right.

III. RESPONSES TO ARGUMENTS AGAINST FORMAL DISCOVERY

The resistance to granting a criminal defendant the power to investigate has deep roots. In 1960, Professor Robert Fletcher wrote:

Historically, discovery was unavailable in either civil or criminal cases, and, despite the full development of discovery in civil cases, denial in criminal cases has persisted. Even as recently as 1927, Mr. Justice Cardozo, then Chief Judge of the New York Court of Appeals, could see only the faint beginnings of a doctrine which would allow discovery in a criminal case. To achieve the degree of liberality that recent cases show, the courts have had to overcome the inertial force of a long and deeply imbedded practice designed to keep the defendant in the dark as long as possible.107

If the hand that rocks the cradle forms our world-view, then the law school experience establishes lasting impressions of what information is sufficient to resolve litigation in a civil or criminal law forum. Civil procedure casebooks provide a comprehensive treatment of formal discovery rights available to litigants. One casebook dedicates 60 pages to the subject.108 But in criminal procedure casebooks there is little discussion of discovery challenges faced by a defendant. One textbook states a criminal defendant’s “discovery provisions uniformly are broader than prosecution discovery provisions”109 as if to suggest he is somehow entitled to more information than the prosecutor, who sits at the helm of a massive law enforcement apparatus. With few exceptions,110 casebooks are silent on the discovery disparity between criminal and civil litigants. If

107 Fletcher, supra note 57, at 294.
109 KAMISAR ET AL., supra note 41, at 1201.
casebooks are intended to inform a student about the legal world as it is, not pointing out the discovery disparity is a failure to do exactly that. Yet, criminal casebooks miss this opportunity to educate law students about the nature of this disparity and whether it has a legitimate justification.

This insularity between disciplines continues into practice. There is little cross-pollination between criminal and civil practitioners. On the rare occasion civil practitioners step into the criminal arena they “tend to be stunned and often outraged by their inability to depose government witnesses or even to file interrogatories or requests for admissions.” A colleague teaching criminal procedure recently broke with this tradition of segregation. Knowing students had taken a semester of civil procedure, he introduced a hypothetical complaint and asked students how they would investigate if they represented the criminal defendant. Hands went up and depositions were scheduled, interrogatories drafted, requests for documents propounded. My colleague let fall the hammer: Padawans, you have none of these formal discovery tools available to you and your anticipated investigation has just been rendered impossible.

Confronted by the inequity from an advocate’s point of view, a sense of injustice emerged. To remedy the informational asymmetry in criminal law, scholarship tends to focus on giving a defendant more access to the prosecutorial file. But this solution merely expands the disclosure of State-authored reports. In a study funded by the Pew Foundation, the Justice Institute proposed:

Mandatory and open-file discovery, in which prosecutors make their entire case file available to the defense and disclose particular items at required times, leads to a more efficient criminal justice system that better protects against wrongful imprisonment and renders more reliable convictions.

---

112 Id. at 714-15.
114 Interview with Byron Lichstein, Associate Clinical Professor, University of Wisconsin Law School, in Madison, Wis. (Nov. 15, 2012) (notes on file with author).
115 Joy, supra note 6, at 641 (“The surest way to meet and exceed Brady disclosure obligations is to adopt an ‘open file’ discovery policy-essentially making available to the defense all of the information in the prosecutor’s possession.”).
This proposal does not correct for the State’s immense power to collect facts that favor its own position. Expanding a defendant’s access to the State’s file remains a solution anchored in the pre-modern discovery era—that the defendant should be a passive recipient of information as opposed to having an essential, formal role in the investigation of the case.

The term “open-file policy” is misleading. The policy provides only a vague degree of access to the prosecutorial file and no access to the police department’s investigatory file. Providing a criminal defendant with a single investigatory tool—the right to compel documents—would exceed the value of any open-file policy. Use of the tool would override the common law rule that investigatory files are protected from disclosure and provide access to prosecutorial and police files. Given that a defendant might seek documents about witnesses and incidents law enforcement did not consider, use of the tool would potentially educate the prosecutor about reports her own detectives failed to consider.

An open-file policy does not authorize a defendant to question any information the State discloses. Given that 90% of criminal defendants plead guilty, an open-file policy does not prevent the State from representing the strength of its case based on a narrative that is free from adequate scrutiny. Yet, despite the slight narrowing of the information gap that would be accomplished by an open-file policy, resistance to its implementation continues. To go beyond an open-file policy and to grant a criminal defendant formal investigatory power meets fiercer resistance. Would the power give defendant an unfair advantage? What would he do with such power? Would he subvert the truth and harass or threaten witnesses? Would he waste resources that are better used to control crime?

A. A Concern that a Defendant Will Have an Unfair Advantage

In criminal law disputes, neither party has use of those particular formal investigatory tools that civil litigants wield—

117 See supra Figure 1.

118 A defense attorney who wishes to call a witness to testify may be dissuaded by his inability to conduct a prior interview. To unintentionally elicit information harmful to the defendant can expose the attorney to allegations of ineffective assistance of counsel. Montoya, supra note 57, at 862.

119 See supra note 36 (regarding percentage of criminal defendants who plea versus advance to trial).
depositions, requests for documents, and interrogatories. If neither party has access to these formal tools, how can one party claim unfair treatment? And if both parties have equal status, then wouldn’t giving one party more power constitute an unfair advantage? As a preliminary matter, denying investigative opportunities to both parties does not improve the quality of facts that inform a dispute. As to fairness, it is not necessarily secured by ensuring that all parties are similarly deprived or empowered. Parties to a criminal dispute are not in any case similarly situated—“today’s defense counsel must meet the prosecutor’s particularly formidable and unprecedented arsenal of fact-gathering methods, including the use of an organized police force to marshal the evidence prior to trial.”

Nor are parties to a criminal dispute similarly deprived—the State in fact exercises its own brand of formal investigatory powers. To treat parties the same during the pretrial period is to leave the criminal defendant at a distinct disadvantage.

Unlike a criminal defendant, the State conducts a formal investigation through the exercise of police powers. Cloaked in state authority, agents have the formal power to arrest, search a person or place, seize evidence, interrogate potential suspects, and, in some instances, threaten dire consequences to ensure cooperation of potential witnesses—such as a probation hold and revocation to prison or the loss of custody of one’s child. In some jurisdictions, prosecutors convene a grand jury, a proceeding that in many respects

120 Montoya, supra note 57, at 870, 862 (“Professor Stanley Fisher has documented a pro-prosecution bias in police investigation and reporting.”).

121 See WILLIAM E. RINGEL, SEARCHES AND SEIZURES, ARRESTS AND CONFESSIONS § 23:9 (2013) for a discussion of the police and other state officials who are given statutory authority to perform arrests.

122 This power is limited by the Fourth Amendment, which protects against unreasonable searches by government agents. U.S. CONST. amend. IV.

123 Officers often work with probation agents to place holds on probationers to facilitate investigation of a crime. See Howard P. Schneiderman, Conflicting Perspectives from the Bench and the Field on Probationer Home Searches—Griffin v. Wisconsin Reconsidered, 1989 Wis. L. REV. 607, 615; see also Wagner v. State, 277 N.W.2d 849, 853-54 (Wis. 1979) (holding that a probation hold of approximately twenty-eight hours to investigate Wagner’s potential involvement in a serious crime was not inappropriately long).

124 A common threat women experience in poor neighborhoods is that the failure to cooperate will result in the loss of their children. This sort of pressure can also be used against a defendant. Regina Kelly was caught up in “a big drug sweep based on the word of a confidential informant who later would be proven unreliable.” Because Kelly was being held in jail, and there was no one to care for her kids, she pled guilty, receiving 10 years of probation. Those who advanced to trial sat for months—those cases were thrown out for lack of any reliable evidence. The Plea: Erma Fay Stewart and Regina Kelly, PBS FRONTLINE (June 17, 2004), http://www.pbs.org/wgbh/pages/frontline/shows/plea/four/stewart.html.
resembles a civil litigant’s pretrial deposition power.\textsuperscript{125} None of these formal investigatory mechanisms are available to a criminal defendant.

The State’s formal discovery tools are in some instances more robust than tools afforded to civil litigants. A civil litigant may have the power to inspect documents of her opponent at a mutually reasonable location, but she does not have the State’s power to conduct a physical search of a person’s premises and seize property. And though the State does not have the power to depose witnesses, prosecutors have the power to convene a grand jury, which extends pretrial subpoena power to a prosecutor in order to compel testimony and conduct a secret investigation.

The State also has the authority to conduct a pretrial interrogation, which is similar to and potentially more powerful than a deposition. Unlike an attorney taking a deposition, an officer is permitted to repeat a question forcefully and express an opinion that the defendant is guilty, a powerful tactic. (Use of this tactic, however favorable to the State’s position, may lack integrity—approximately 25\% of DNA exonerations arose out of cases in which the defendant falsely confessed to the crime.\textsuperscript{126}) Depositions tend to take place in a pleasant enough room while interrogations take place in cinderblock cells. Attendance at a deposition is compelled by the threat of judicial sanction for failure to comply with a subpoena. Participating in an interrogation is compelled by the request of a law enforcement officer or through arrest. The refusal to answer to authority is rare even after arrest, which unlike a subpoena requires no notice to any opposing counsel.\textsuperscript{127}

In a deposition, the witness typically has counsel. In an interrogation, the witness is typically alone and answering to one or more officers. In an interrogation, there are no significant time restraints. An interrogation may be conducted in multiple rounds by different interrogators at any time—often in

\textsuperscript{125} The Fifth Amendment provides that “[n]o person shall be held to answer for a capital or otherwise infamous crime, unless on a presentment or indictment of a Grand Jury.” U.S. CONST. amend. V. There is no requirement, however, that states employ the use of a grand jury. Hurtado v. California, 110 U.S. 516, 534 (1884). Most states have procedures to convene grand juries, and about a third require grand jury indictments for felonies. U.S. DEPT. OF JUSTICE, OFFICE OF JUSTICE PROGRAMS, NCJ 212351, STATE COURT ORGANIZATION 2004 215-17 (Aug. 2006), http://cdm16501.contentdm.oclc.org/cdm/ref/collection/juries/id/180.

\textsuperscript{126} DNA Exonerations Nationwide, INNOCENCE PROJECT http://www.innocenceproject.org/Content/DNA_Exonerations_Nationwide.php.

\textsuperscript{127} Only 22\% of those placed in custody invoke their Miranda rights and refuse to speak to police during an interrogation. Richard A. Leo, The Impact of Miranda Revisited, 86 J. CRIM. L. & CRIMINOLOGY 621, 653 (1996).
the middle of the night—and may potentially transpire throughout a long duration of custody. The subject is isolated from his support system. Officers may threaten the witness/suspect, falsely suggest others are implicating him, or manufacture a nonexistent case against him to obtain information. In contrast, civil attorneys taking a deposition are ethically barred from engaging in deception and deterred from doing so. Opposing counsel is not only present, but interactions are transcribed. Any ethical violation would be on the record and could lead to a disciplinary review that the opposing party may be happy to set into motion.

A key advantage to the interrogation is the resulting report. State agents interpret a witness’s responses and summarize the interview. After a deposition, a civil litigator may believe he neutralized a witness. A later review of the transcript, however, might suggest his line of questioning proved something less or

128 In State v. Masch, Milwaukee Case No. 2003CF004581, defendant was held in custody over a period of 58 hours and subject to four separate interrogations conducted by four different pairs of officers that totaled approximately 14 hours. The final pair of interrogators, two experienced homicide detectives, were called in by special request to break the defendant.

129 See, e.g., BRANDON L. GARRET, CONVICTING THE INNOCENT: WHERE CRIMINAL PROSECUTORS GO WRONG 39 (2011) (noting a case in which 17-year-old Paula Gray, who was borderline mentally impaired, inculpated herself and four other innocent people in a double murder. "Gray testified that she was asked, 'Did they emphasize what would happen if you did not tell this story?' and answered, 'That they would kill me.'").


131 See, e.g., Garrett, supra note 129, at 22-23. David Vasquez, for example, was told by police that his fingerprints were found at the scene of a murder and eventually confessed. Id. at 22. He was exonerated after the real perpetrator was found; he had served four years in prison by that time. Know the Cases: David Vasquez, INNOCENCE PROJECT, www.innocenceproject.org/Content/David_Vasquez.php.

132 Since the purpose of an interrogation, generally, is to cause the subject to confess, police often use “persuasive techniques comprising trickery, deceit and psychological manipulation.” Gisli H. Gudjonsson, The Psychology of Interrogations and Confessions, in INVESTIGATIVE INTERVIEWING: RIGHTS, RESEARCH, REGULATION 123, 124 (Tom Williamson ed., 2006). In contrast, several of the Model Rules of Professional Conduct would be implicated if an attorney engaged in deception during a deposition. See MODEL RULES OF PROF’L CONDUCT R. 4.1 (2011) (governing truthfulness in statements to others); MODEL RULES OF PROF’L CONDUCT R. 8.4(c) (2011) (prohibiting deceitful behavior); MODEL RULES OF PROF’L CONDUCT R. 3.5 cmt. 5 (2011) (“The duty to refrain from disruptive conduct applies to any proceeding of a tribunal, including a deposition.”); MODEL RULES OF PROF’L CONDUCT R. 3.3 cmt. 1 (2011) (indicating that the rule governing candor to the tribunal includes conduct during a deposition).

133 Police fabrication of reports is a significant problem, since police reports are often “dispositive in a case resolved through plea bargaining.” Christopher Slobogin, Testifying: Police Perjury and What to Do About It, 67 U. COLO. L. REV. 1037, 1044 (1996).
nothing at all. This “morning after” disappointment does not occur for detectives who obtain an inculpatory statement or author the resulting summary of their efforts. The State and its agents almost always retain rights over the narrative.\footnote{As of 2010, 17 states and the District of Columbia required recording of suspect confessions under certain circumstances. See Alan M. Gershel, A Review of the Law in Jurisdictions Requiring Electronic Recording of Custodial Interrogations, 16 Rich. J. L. & Tech. 9 (2010).} The interrogation in these ways serves as a “shadow deposition” that potentially provides the State an investigatory tool with greater advantages than a formal deposition.

The view that the parties to a criminal dispute are equally situated is as erroneous as it is pervasive. Ignoring the deep asymmetry between an empowered State and a frequently detained defendant, the misapplied concept—to equate equal treatment with fairness—is embedded in the rules of criminal procedure. The Federal Rules of Criminal Procedure, for example, adopt the concept of “reciprocity.”\footnote{Fed. R. Crim. P. 16 advisory committee’s notes to 1974 amendment (“The House version of the bill provides that the government’s discovery is reciprocal. If the defendant requires and receives certain items from the government, then the government is entitled to get similar items from the defendant. . . . The Conference adopts the House provisions.”).} Only after a criminal defendant requests certain information from the State may the prosecutor then request the same category of information from defendant—a statutory expression of “I’ll show you mine if you show me yours.” If, for example, a criminal defendant requests documents that the prosecutor plans to use at trial, only then may the prosecutor request the same disclosure from the defendant.\footnote{Fed. R. Crim. P. 16(b)(1)(A)(ii).} This concept of reciprocal exchange suggests equal burden and measure, masking the deep disparity that exists between the parties.

The state apparatus brings to bear impressive formal investigatory powers. A criminal defendant must sip from the cup of his opponent. Affording formal investigatory tools to a defendant would begin to correct for this existing imbalance yet leave intact the State’s inherent investigatory and trial advantages. The State would still conduct a formal, unregulated pre-complaint investigation. The State’s officers would still summarize their own interviews. The State would still control a police force that has the qualified right to search and seize, to engage in deception and threats to secure information, and to employ suggestive lineup procedures that target a suspect. The State would still have at its disposal detectives to authoritatively

The interrogation in these ways serves as a “shadow deposition” that potentially provides the State an investigatory tool with greater advantages than a formal deposition.
explain the course of the investigation to a jury, providing a theatrical display of self-validation. And the State would still pull from its stable of law enforcement experts to testify about the results a scientifically untested but quite persuasive analysis.\textsuperscript{137}

B. A Concern That a Defendant Will Threaten Witnesses

In 1974, an effort to require pretrial disclosure of the State’s witnesses in federal criminal disputes met vigorous opposition. The United States Department of Justice reported that allowing this disclosure would be “dangerous and frightening in that government witnesses and their families will even be more exposed than they are now to threats, pressures, and physical harm.”\textsuperscript{138} There is little empirical guidance on the issue. Since 1974, some state jurisdictions have permitted a criminal defendant access to witness lists—one would expect that any significant uptick in witness intimidation would have led to a repeal of this policy. Yet these provisions remain on the books.\textsuperscript{139} Justice William Brennan, observing that particular circumstances might warrant concern, opined “the proper response . . . cannot be to prevent discovery altogether; it is rather to regulate discovery in those cases in which it is thought that witness intimidation is a real possibility.”\textsuperscript{140} Beyond the issuance of protective orders in cases where there is reason to believe that a particular witness is in danger, there are sufficient existing deterrents to prevent witness intimidation. Pretrial custody reduces a defendant’s ability to communicate with the outside world. Communications from the jail are monitored. A jailhouse call revealing an attempt to intimidate a witness may be used against a defendant as affirmative evidence of guilt.\textsuperscript{141} Under federal law, anyone who attempts to dissuade a witness from testifying faces 20 years in prison.\textsuperscript{142}

\textsuperscript{137} National Research Counsel, Strengthening Forensic Science in the United States—A Path Forward 22 (2009) (“Some forensic science disciplines are supported by little rigorous systematic research to validate the discipline’s basic premises and techniques.”); id. at 42 (“The fact is that many forensic tests—such as those used to infer the source of toolmarks or bite marks—have never been exposed to stringent scientific scrutiny.”).

\textsuperscript{138} Brennan, Jr., supra note 12, at 6 (quoting Justice Department testimony before Congress) (footnotes omitted).

\textsuperscript{139} See, e.g., CAL. PENAL CODE § 1054.1(a) (West 2013); ILL. SUP. CT. R. 412(a)(i); OHIO R. CRIM. P. 16(I).

\textsuperscript{140} Brennan, Jr., supra note 12, at 14.

\textsuperscript{141} See, e.g., United States v. Miller, 276 F.3d 370, 373 (7th Cir. 2002) (“Evidence that the defendant threatened a potential witness or a person cooperating with a government investigation is relevant to show the defendant’s consciousness of guilt.”).

\textsuperscript{142} 18 U.S.C. § 1512(b) (2012).
In permitting a defendant deposition power, there is concern over subjecting a victim or witness to a defendant’s presence at a deposition “without the security provided by the courtroom setting.”\textsuperscript{143} A criminal defendant’s presence at a deposition, however, is not constitutionally required.\textsuperscript{144} In those jurisdictions that permit depositions in criminal cases, “only defense counsel need be present … .”\textsuperscript{145} For example, Missouri provides a default rule that a criminal defendant “shall not be physically present at a discovery deposition except by agreement of the parties or upon court order for good cause shown.”\textsuperscript{146} Florida provides protections for “sensitive witnesses.”\textsuperscript{147} A related argument is that a victim of domestic violence or sexual assault might be deterred from cooperating. She would face the anxiety of being subjected to hours of deposition testimony that, by design, allows for questioning that exceeds what could be explored at trial. Precautions would mitigate these concerns: ensuring that the defendant is not present, limiting the time to depose the victim, making certain subjects, like past sexual history, off limits, and restricting the distribution of testimony to attorneys.

Left unexplored by those concerned about witness intimidation is the fact that a criminal defendant is powerless to counter state-initiated efforts to intimidate witnesses. State-initiated threats can be serious in nature and effective.\textsuperscript{148} In 2010, after 15 years of prison, Jabbar Collins was exonerated of

\begin{itemize}
\item \textsuperscript{143} KAMISAR ET AL., supra note 41, at 1206.
\item \textsuperscript{144} The right to confront witnesses against the defendant only ripens at trial. See Sarah A. Stauffer & Sean D. Corey, Sixth Amendment at Trial, 87 GEO. L.J. 1641, 1647-48 (1999). In Florida, a defendant is not allowed to be present at a deposition absent court approval or stipulation by the parties. FLA. R. CRIM. P. 3.220(b)(7). The Florida Supreme Court concluded a discovery deposition does not provide for meaningful cross-examination of the deponent because a discovery deposition is not a device designed to gather testimony for later use at trial. State v. Lopez, 974 So. 2d 340, 347 (Fla. 2008).
\item \textsuperscript{146} MO. SUP. CT. R. 25.12(c).
\item \textsuperscript{147} FLA. R. CRIM. P. 3.220(b)(4) (providing, “[d]epositions of children under the age of 16 shall be videotaped unless otherwise ordered by the court. The court may order the videotaping of a deposition or the taking of a deposition of a witness with fragile emotional strength to be in the presence of the trial judge or a special magistrate”).
\end{itemize}
murder after a key State witness testified in a post-conviction proceeding that the prosecutor had threatened to hit him and incarcerate him for any failure to testify in accordance with the State’s theory.\textsuperscript{149} The witness was jailed for a week before he eventually agreed to testify falsely for the State.\textsuperscript{150} The State may also engage in threats to suppress potentially exculpatory evidence—“[p]olice and prosecutorial improprieties take on several different forms: [including] making threats against potential witnesses for the accused.”\textsuperscript{151} In a Milwaukee homicide case, a potential alibi witness told police that, at the time of the shooting across town, he thought the defendant had arrived at the witness’s residence. Officers threatened to tell the witness’s track coach that the witness was lying in a homicide investigation and warned that he would lose his scholarship.\textsuperscript{152} Without the power to compel pretrial answers from officers, witnesses, and those who may have observed the exchange, a defendant cannot overcome the threat’s effectiveness.

State witnesses are not the only individuals subject to intimidation. A witness with exculpatory or inculpatory information may refuse to come forward because he fears retribution from an alternate suspect or the State. The potentiality of witness retaliation is real, but the source of that retaliation may be State agents, a State witness, or an alternate suspect. Both parties should have sufficient powers to expose incidents of intimidation. The State has that power through the use of its police force and witness protection programs; a defendant does not.

C. \textit{A Concern That a Defendant Will Misuse Formal Powers}

A perennial concern over providing formal investigatory power to a criminal defendant is that he will misuse it to delay the State’s case.\textsuperscript{153} The predicate conditions to committing

\textsuperscript{149} Sulzberger, supra note 148, at A18.

\textsuperscript{150} Id.

\textsuperscript{151} See C. RONALD HUFF ET AL., CONVICTED BUT INNOCENT: WRONGFUL CONVICTION AND PUBLIC POLICY 71 (1996).

\textsuperscript{152} This example is based on an investigation conducted by the Wisconsin Innocence Project. Interviewed five years after the event, the young man, now working at a bank and running a non-profit to assist inner-city kids, confirmed detectives made these threats. Information on file with author.

\textsuperscript{153} Brennan, Jr., supra note 12, at 6 (detailing arguments made by Chief Justice Vanderbilt of the New Jersey Supreme Court against liberal discovery for criminal defendants); Brown, supra note 2, at 1590 (“defense counsel’s commitment is not to accuracy; it is to his or her clients, many of whom want inaccuracy to mask their
discovery abuse, however, are typically absent from criminal law. Discovery abuse in civil litigation is associated with document-heavy cases in which well-resourced opponents engage in dilatory practices by propounding unnecessary discovery requests or dumping truckloads of marginally responsive documents on opposing counsel. Other than in white-collar disputes (which represent a very small fraction of criminal cases), these tactics would not typically be available to a criminal defendant—the vast majority of criminal defendants are indigent. Any attempt by a defense attorney to unnecessarily subpoena law enforcement officers would reach the ears of the prosecutor and court. That an overworked public defender would have the time, interest, or resources to engage in discovery abuse is unlikely.

The concern that a criminal defendant would only use discovery rights to subvert the process is likely grounded in the presumption that the State generally “gets it right”—that the criminal defendant is guilty. But this resistance to affording a presumptively liable party with investigatory power appears to be reserved for a criminal defendant. There is no movement to deprive a civil defendant of formal powers where the defendant is most certainly liable—Exxon in the Puget Sound oil spill litigation, for example. Rather, we leave it to a civil defendant, however liable, to use its sometimes impressive resources (resources a criminal defendant might never possess) to protect its interests. In this sense, the potentiality that a liable party would be motivated to interfere with an investigation is not unique to criminal law. A civil defendant may face overwhelming liability that leads to damage control

guilt”). It should be noted that Brown nevertheless proposes more robust discovery, advocating for an open-file policy. Id. at 1637-38.

154 See, e.g., Kawamata Farms v. United Agri Prods., 948 P.2d 1055, 1090-91 (Haw. 1997) (“The record shows that DuPont engaged in a pattern of discovery abuse by, among other things, violating the circuit court’s discovery orders, ‘dumping’ forty boxes of documents pursuant to one of the Plaintiffs’ interrogatory requests, and intentionally withholding information and documents that DuPont should have produced during discovery. This inexcusable behavior by DuPont is very disturbing.”); Class Action starring Gene Hackman and Mary Mastrantonio. CLASS ACTION (20th Century Fox 1991).

155 Exxon’s big boat spilled 11 million gallons of oil into a pristine bay. Exxon argued it was free from liability, but was found to be “worse than negligent but less than malicious.” The Supreme Court, however, did agree with Exxon that maritime common law principles applied, reducing a $2.5 billion punitive award. Under maritime law punitive damages should not exceed a 1:1 ratio with compensatory damages ($507 million). See David Savage, Justices Slash Exxon Valdez Verdict, L.A. TIMES, June 26, 2008, http://articles.latimes.com/2008/jun/26/nation/na-valdez26.
behaviors, from destroying documents\textsuperscript{156} to making examples of employees who breach conceptions of loyalty.\textsuperscript{157} But these potentialities do not result in calls to preclude a civil defendant from testing plaintiff’s theories.

At the same time, should formal investigatory powers be extended to criminal defendants, some defense attorneys \textit{would} inevitably engage in delay or subterfuge. Civil procedure provides for checks on dilatory practice, including the imposition of protective orders and judicial sanctions.\textsuperscript{158} In Florida, where formal investigatory rights are extended to the criminal defendant, due process does not preclude a court from sanctioning him for violating discovery rules, including the power to prohibit him from calling a witness or introducing documentary evidence.\textsuperscript{159}

On the other hand, the concern that a criminal defendant will use discovery tools to subvert the truth ignores the possibility that law enforcement will use police powers to subvert the truth. The lack of formal investigatory power renders a criminal defendant particularly vulnerable to a law enforcement officer who falsifies witness testimony in police reports. A former San Francisco Police Commissioner recently

\textsuperscript{156} Charles R. Nesson, \textit{Incentives to Spoilate Evidence in Civil Litigation: The Need for Vigorous Judicial Action}, 13 CARDOZO L. REV. 793, 793 (1991) (“[O]ne half of litigators believe that ‘unfair and inadequate disclosure of material information prior to trial [is] a ‘regular or frequent’ problem . . . [and] 69% of surveyed antitrust attorneys [have] encountered unethical practices, in including, most commonly, destruction of evidence.’”).

\textsuperscript{157} Whistleblower statutes—those statutes that protect people who expose wrongdoing by either incentivizing their decision to speak or protecting them from retaliation—reflect policymakers’ attention to this problem in the civil sphere. \textit{See generally} Elletta Sangrey Callahan & Terry Morehead Dworkin, \textit{The State of State Whistleblower Protection}, 38 AM. BUS. L.J. 99 (2000) (discussing legislative and judicial protections available to whistleblowers in the United States); \textit{see also} Paul Sullivan, \textit{The Price Whistle-Blowers Pay for Secrets}, N.Y. TIMES, Sept. 21, 2012, http://www.nytimes.com/2012/09/22/your-money/for-whistle-blowers-consider-the-risks-wealth-matters.html?pagewanted=all (“If you look at the field of whistle-blowers, you see a high degree of bankruptcies. You may find yourself unemployed. Home foreclosures, divorce and depression all go with this territory.”) A spokesman for Taxpayers Against Fraud stated that, for whistleblowers, “[t]here is a 100 percent chance that you will be unemployed—the question is, Will you be forever unemployed? . . . The other 100 percent factor is the person who fired you, the person who designed and implemented the fraud, won’t be fired. He’ll probably be promoted again.” \textit{Id.} (internal quotation marks omitted). Statutes like Title VII’s retaliation provisions are meant to prevent such results. \textit{See} 42 U.S.C. § 2000e-3(a) (2006).

\textsuperscript{158} \textit{Fed. R. Civ. P. 26 advisory committee’s note to 1983 amendment} (“Excessive discovery and evasion or resistance to reasonable discovery requests pose significant problems.”). In response, the Advisory Committee struck the language stating “the frequency of use of the various discovery methods was not to be limited” with the intent that parties, when appropriate, would file a protective order. \textit{Id.} (internal quotations omitted).

\textsuperscript{159} \textit{Fla. R. CRIM. P. 3.220(n)}. 1
complained of the incidence of narcotics officers falsifying evidence. In another instance, Bronx Assistant District Attorney Jeannette Rucker determined that “it had become apparent that the police were arresting people even when there was convincing evidence that they were innocent,” and found that officers had provided “false written statements” to justify the arrests. Despite such concerns, a criminal defendant has no formal power to depose officers or the witnesses subject to State incentives, threats, or police falsification.

Not all civil litigants are angels and not all civil disputes are just about money. Like in the criminal law, civil cases can be of great significance. At issue might be a critical question about whether a health condition is covered by insurance, whether a bay fouled by oil will be restored, or whether minority students should be subjected to a segregated school system. We provide all potentially serious civil wrongdoers—whose decisions have led to discrimination, bankruptcies, injury, death, and environmental degradation—robust pretrial power to potentially undermine the truth. There is little reason that a criminal defendant, presumed innocent, should not be privileged to sit among such company. We might reframe the issue: irrespective of whether a criminal defendant is innocent, guilty, or ultimately something in between, he is motivated to challenge the State’s theory of liability. The fact that a criminal defendant faces significant consequences best positions him to check mistakes made by his opponent.

In the end, it is not the significance of the issues that distinguishes a criminal defendant from his fellow tortfeasors. Rather, the distinguishing characteristic is the unique burden

161 Id.
162 Abram Chayes, The Role of the Judge in Public Law Litigation, 89 HARV. L. REV. 1281, 1281 (1976); Linda S. Mullenix, Resolving Mass Tort Litigation: The New Private Law Dispute Paradigm, 33 VAL. U. L. REV. 413, 415 (1999) (stating that the “aggregative private dispute resolution paradigm resembles nothing so much as private legislation with wide-reaching effects, carrying the imprimatur of judicial oversight and approval, but frequently accompanied by troubling questions about fairness, adequate representation, and the subtle merger of legislative, administrative, and judicial functions”); see also Roscoe Pound, The Causes of Popular Dissatisfaction with the Administration of Justice, 40 AM. L. REV. 729, 730 (1906) (read by Roscoe before the American Bar Association in St. Paul, Minnesota on August 29, 1906) (“The rules which define those invisible boundaries within which each may act without conflict with the activities of his fellows in a busy and crowded world, upon which investor, promoter, buyer, seller, employer and employee must rely consciously or subconsciously in their every-day transactions, are conditions precedent of modern social and industrial organization.”).
borne by the criminal defendant. He faces severe punishment—the loss of liberty and damage to reputation, inhumane conditions of prison, and the assured lack of opportunities following release. All of these concerns underscore the importance of providing the criminal defendant with formal investigatory tools to defend himself. Added to this is the desire that a criminal defendant be able to defend himself in accordance with the “categorical ex ante judgment that society would prefer to let a guilty person go free rather than send an innocent person to prison.”

And a criminal defendant may in fact be innocent. To deprive a criminal defendant of formal investigatory power underestimates how difficult it may be for a factually innocent defendant to secure an acquittal. An innocent defendant likely has no idea who committed the crime. Alibis tend to be hard to prove. An alibi may be unsophisticated and, when tested by a talented prosecutor, appear suspect. A crime may occur at night (it often does) and an alibi that asserts the defendant went to sleep could not confirm he did not slip out. The theory that the State’s high burden of proof will set an innocent defendant free is a weak insurance policy against an erroneous eyewitness account, for example, which is highly resilient to cross-examination. In some cases, only affirmative evidence demonstrating that an alternate suspect likely committed the crime is sufficient to overcome the State’s case. To demonstrate innocence in such circumstances, the defense attorney cannot rely on the State’s evidence but must instead conduct an independent inquiry that leads to a different result.

D. A Concern that Trial is the Proper Forum for Testing the Case

Providing criminal defendants with formal investigatory tools might make the pretrial period more adversarial, and there is a concern that doing so would waste resources during a period in which most criminal defendants voluntarily admit

---

164 Leipold, supra note 12, at 1152 (“Because it has the burden of proof, the prosecutor collects most of the evidence.”).
165 Brown, supra note 2, at 1602 (“Once an eyewitness’s memory has been affected by suggestive identification procedures, it is hard to undo the damage.”).
There is a significant body of scholarship, however, that challenges the assumption that all pleas are voluntary. And for those criminal defendants who are overcharged or innocent, the lack of formal discovery tools leaves a criminal defendant ill-prepared by the time he advances to trial.

Trial provides an inappropriate forum to conduct an investigation of the case. By the time a jury is impaneled, litigants should not be exploring alternative theories of liability. Pretrial motions have been decided. Litigants have determined what narrative they want to convey. Any absence of a trial strategy at this juncture would constitute deficient performance—a trial strategy to “figure out what to do after we investigate at trial” is per se deficient. Even if a defendant attempted to conduct an investigation at trial, the rules of evidence would inhibit the effort; hearsay, inadmissible at trial, is essential to establishing investigative leads. Cross-examination is designed to cement, not uncover, a narrative. Trial does not provide the optimum forum to refresh a witness’ recollection, a process that can result in long periods of silence as a witness reviews documents. Trial is in part a public spectacle, roles have already been assigned, the script finalized. If a defendant has not adequately investigated the incident by the eve of trial, it is too late for defendant. He will lose.

And trial is not where disputes are predominantly resolved. As to the few disputes that advance to trial, the quality of facts informing strategy and witness selection is inextricably tied to the quality of the pretrial investigation. A pretrial investigation also protects the record. It is not uncommon for a witness to miss a trial appearance due to unstable living arrangements, mental health and substance abuse issues, or

---

166 Gershman, supra note 8, at 145 (“[W]hereas a fair trial involves a forced settlement of a factual dispute in a fair adversarial contest before a judge and jury, a fair plea typically does not involve a factual dispute.”).  
167 Langer, supra note 5, at 276 (suggesting more discovery rights like open-file policies to counteract the conditions that lead to coerced pleas).  
168 See, e.g., Silva v. Woodford, 279 F.3d 825, 846 (9th Cir. 2002) (“[A]n attorney’s performance is not immunized from Sixth Amendment challenges simply by attaching to it the label of ‘trial strategy.’ Rather, ‘certain defense strategies may be so ill-chosen that they may render counsel’s overall representation constitutionally defective.’”) (citing United States v. Tucker, 716 F.2d 576, 586 (9th Cir. 1983)).  
169 See, e.g., Fed. R. Evid. 802-04. The hearsay rule alone precludes conducting an adequate investigation—the question, “who told you that?” being central to any investigation. In addition, the “other acts rule” precludes inquiring about what the witness has done, and his knowledge of what others have done, in the past. See Fed. R. Evid. 404.  
170 Whereas 90 percent of criminal disputes resolve in a plea deal, only 10 percent of criminal litigants advance to trial. See supra note 36 and accompanying text.
trouble with the law. A formal investigatory tool like the deposition provides an insurance policy against no-shows. Unlike the one-shot opportunity presented at trial, a party has time to make pretrial attempts to secure a witness’ attendance at a deposition. If a witness fails to appear for trial, any recorded or written prior statement will typically constitute inadmissible hearsay. A deposition, however, ensures that in these circumstances admissible portions of the witness’s transcript will be admitted at trial under the “former testimony” exception to the hearsay rule. A jury should consider what all critical witnesses have to say. Substantive deliberation should not be undermined by the unavailability of a witness on the day of trial.

E. A Concern That Defendant Already Has Enough Rights

A criminal defendant has constitutional protections designed to check abuses by a police force that is otherwise authorized to search, seize, arrest, interrogate, lie, and threaten serious consequences for non-cooperation. These constitutional rights are defensive in nature. They grant no affirmative power to a defendant to develop evidence (with the notable exception of the Sixth Amendment right to counsel). They also do little to check the intentions of the state apparatus. Deference typically afforded to an officer’s account of conditions on the ground and the exceptions to the warrant requirement render the suppression of evidence rare. Individuals tend to cooperate in custodial interrogations despite a right to remain silent. Regarding constitutional efforts to prevent police coercion, in jurisdictions that do not require officers to record the interrogation the judge weighs the credibility of a criminal defendant who allegedly confessed to criminal activity against the credibility of officers. Because these constitutional protections are defensive in nature, they do not provide a criminal defendant any affirmative right to engage in fact-finding. Yet some might find it particularly unfair to give the criminal defendant a right to compel information from the State’s witnesses when the

171 Fed. R. Evid. 804(b)(1).
172 Craig D. Uchida & Timothy S. Bynum, Search Warrants, Motions to Suppress and “Lost Cases:” The Effects of the Exclusionary Rule in Seven Jurisdictions, 81 J. CRIM. L. & CRIMINOLOGY 1034, 1044-45 (1991) (citing to a 1979 study of almost 3,000 cases conducted by the General Accounting Office that found successful motions to suppress were made in 1.3% of the prosecuted cases).
173 Leo, supra note 127, at 653 (finding that only 22% of those placed in custody invoke their Miranda rights and refuse to speak to police during an interrogation).
defendant can elect to invoke the Fifth Amendment. This argument does not acknowledge that the existing informational asymmetry between the parties would persist even if formal discovery powers were only extended to a criminal defendant. A criminal defendant will never match the State’s exercise of police powers through a sophisticated police force.

The concern also overstates the Fifth Amendment’s limitation on the State’s power. It is common for civil litigants to be foreclosed from making inquiries into privileged information, however probative that evidence might be. In a shareholder suit, a plaintiff can expect to be prevented from inquiring into an executive board meeting held in the presence of the board’s attorney. Yet, these privileges do not ultimately inhibit broad and intrusive inquiry into the opposing party’s theory of the case. Likewise, that the State may be foreclosed from deposing a defendant would not preclude the State from compelling testimony from a defendant’s friends, family, alibi, former employers, landlords, and anyone else with relevant information. And the State may not be the only party hindered by the Fifth Amendment. The constitutional protection would also pose a challenge to a defendant pursuing an alternate perpetrator defense, as an alternate suspect deposed by defense counsel is also privileged to claim a right to silence.

F. A Concern That Formal Discovery Would Be Too Costly

One concern over extending formal discovery is “that depositions are very costly, and with the state footing the bill for indigent defendants, there is no financial sacrifice that would provide a restraint against appointed counsel conducting unnecessary depositions.” Providing deposition power, however, will not relieve public defenders of a relentless caseload. In Wisconsin, for example, a public defender must meet an annual quota of 200 points—receiving a half-point for a misdemeanor.

174 See United States v. Garsson, 291 F. 646, 649 (S.D.N.Y. 1923), for Judge Learned Hand’s position against expanding discovery to a criminal defendant.

175 Florida’s approach to discovery depositions in criminal proceedings provide an example of how discovery depositions can be used in the criminal justice system without running afoul of the confrontation clause. FLA. R. CRIM. P. 3.220(b)(1)(A)-(D) (allowing both the defendant and the prosecution to depose certain categories of witnesses).

176 KAMISAR ET AL., supra note 41, at 1206.

177 Peter A. Joy, Ensuring the Ethical Representation of Clients in the Face of Excessive Caseloads, 75 Mo. L. Rev. 771, 777 (2010) (“Most commentators and bar leaders agree that the major factors contributing to poor quality of defense services are excessive caseloads, lack of funds for expert witnesses and investigators, and extremely low pay rates for court-assigned lawyers and contract defense services.”).
and upwards of 20 points for a homicide. An entry-level public
defender must dispose of 400 cases in one year to meet her
minimum. Those litigating on the felony calendar and
predominantly taking Class A felony cases “reduce” their
workload to two or three homicide cases a month. Wisconsin
public defenders actually fare better than public defenders of
other jurisdictions where “felony caseloads of 500, 600, and 800
or more are common.” Caseload realities effectively provide a
built-in deterrent to a deposition’s overuse.

In serious criminal cases, the stakes are high and the
sources of information tend to be limited. In the absence of
interrogatories to seed the investigation and subpoena power to
compel production of testimony and documents, conducting an
informal investigation may prove not only less effective, but
more costly than conducting a formal investigation. To the
extent that instances of discovery abuse surface, judicial
intervention provides a moderating role, and subsequent reform
efforts in the civil forum have sought to increase cooperative
behavior in exchanging information. Operationally, the
complexity of a dispute tends to govern the use of discovery. Even
where litigants have an arsenal of discovery tools at their
disposal, simple disputes—the majority of civil cases, in fact—are
resolved in the absence of discovery. In Florida, the legislature
considered this concern and has codified that “[n]o deposition
shall be taken in a case in which the defendant is charged only
with a misdemeanor . . . unless good cause can be shown to the
trial court.” A party in New Mexico may subpoena a witness,
record the interview, and direct an assistant to prepare a
transcript to which the opposing party typically stipulates, a
measure that greatly reduces costs associated with securing a

---

178 Interview with Michele LaVigne, Clinical Professor of Law, Univ. of Wis.
Wainwright, NATIONAL LEGAL AID & DEFENDER ASSOCIATION, http://www.nlada.org/
180 FED. R. CRIM. P. 26 advisory committee’s notes to 1993 amendment
(requiring mandatory disclosures, the “purpose of the revision is to accelerate the
exchange of basic information about the case and to eliminate the paper work involved
in requesting such information, and the role should be applied in a manner to achieve
those objectives”).
181 David M. Trubeck et al., The Cost of Ordinary Litigation, 31 UCLA L. REV. 72,
89-90 (1983) (”Our data [analyzing civil litigation trends] suggest[s] that relatively little
discovery occurs in the ordinary lawsuit. We found no evidence of discovery in over half our
cases. Rarely did the records reveal more than five separate discovery events.”).
court reporter in a deposition. Formal discovery tools are subject to customization. Properly configured, they are potentially more effective and efficient than the informal discovery methods currently available to the criminal defendant.

CONCLUSION

Rules that govern the exchange of information ultimately reflect the quality of information society agrees to afford litigants. A limited grant of discovery power would suggest an unwillingness to disrupt daily life to resolve a dispute. A small claims court, for example, does not permit litigants to depose witnesses to determine the exact value of damage done to a personal printer. In contrast, invasive discovery tools are permitted in disputes deemed significant. In civil disputes, litigants are afforded investigatory tools that disrupt the lives of others. In this light, a criminal defendant has more in common with a small claims litigant.

Precluding a criminal litigant from engaging in formal investigation means that the quality of facts informing resolutions is, relative to civil law outcomes, inferior. Entitled to only discrete information, negotiations in criminal disputes are based on allegations in the complaint, evidence favorable to the State, and the raw power to threaten sobering penalties in exchange for reduced punishment. One cannot imagine a civil dispute in which a defendant’s settlement position is informed by allegations in the complaint and documents selected by a plaintiff. Yet, most criminal defendants are entitled to just that, facilitating complaint-based outcomes that credit prosecutorial hunches.

These pretrial deficiencies—affecting 90% of defendants—are not cured by trial. The information that informs a criminal trial, relative to civil trials, is also inferior. The overwhelming source of information originates from the State’s file—sources of potentially exculpatory evidence remain unexplored and witnesses who have not been deposed are freer to prevaricate. The Supreme Court acknowledged in Kyles v.

---

183 Interview with Katherine Judson, Innocence Project Litigation Fellow, in Madison, Wis. (Oct. 23, 2012) (notes on file with author).
184 Small claims courts are characterized by the “lack of opportunity to conduct discovery.” Bruce Zucker & Monica Her, The People’s Court Examined: A Legal and Empirical Analysis of the Small Claims Court System, 37 U.S.F. L. Rev. 315, 347 (2003).
185 See supra Part I.C.
186 See supra note 36 and accompanying text.
Whitley\textsuperscript{187} that it is a legitimate defense to argue to the jury that the State’s investigation was flawed. But a defendant cannot discern the existence of this defense without the ability to conduct an independent investigation to show what law enforcement missed.

Civil discovery rules are designed to fuel a broad \textit{pretrial} investigation, one that uncovers and tests the credibility of evidence before trial.\textsuperscript{188} The orientation of criminal discovery rules toward the trial moment further diminishes a defendant’s pretrial access to information.\textsuperscript{189} These disclosure provisions are wedded to an event that rarely occurs. Some jurisdictions are trial-centric \textit{in toto}; Wisconsin does not require any disclosure of State’s evidence until “a reasonable time before trial.”\textsuperscript{190}

A criminal defendant is not entitled to play an essential role in the investigation of his case. His right to be informed is in every respect inferior to all other parties in the common law system.\textsuperscript{191} In 1974, when changes were made to federal criminal procedure to ensure that pretrial disclosure was mandatory upon request, it was done because:

\begin{quote}
\begin{itemize}
  \item broad discovery contributes to the fair and efficient administration of criminal justice by providing the defendant with enough information to make an informed decision as to plea; by minimizing the undesirable effect of surprise at the trial; and by otherwise contributing to an accurate determination of the issue of guilt or innocence.\textsuperscript{192}
\end{itemize}
\end{quote}

This was the Advisory Board’s “Mission Accomplished” moment—the finish line is a long way off.

The criminal defendant’s role in an investigation should be reevaluated. Currently, he is a passive recipient of information. Statutorily, he is afforded no role in the investigation of his own case. Greater access to the prosecutorial file and more resources would go some way to mitigate existing deprivations. But in light of the State’s formal power to investigate, the failure to provide a criminal defendant formal investigatory tools leaves potentially

\textsuperscript{187} 514 U.S. 419, 445 (1995) (observing that disclosure of the informant’s statements would have given the defense grounds to attack “the thoroughness and even the good faith of the investigation”).

\textsuperscript{188} \textit{See, e.g.}, FED. R. CIV. P. 26(a).

\textsuperscript{189} Langer, \textit{supra} note 5, at 275 (stating that federal criminal procedure in particular “establish[es] a mainly trial-centric approach to discovery rules”).

\textsuperscript{190} WIS. STAT. ANN. § 971.23(1) (West 2013) (requiring that certain disclosures be made at a “reasonable time before trial”).


\textsuperscript{192} FED. R. CRIM. P. 16 advisory committee’s note to 1974 amendment.
unsubstantiated narratives untested. Compared to the major disputes resolved in civil litigation that are informed by multiple sources and careful examination of witnesses and documents, the criminal system remains shielded from the light of adversarial testing. Over 50 years ago, the United States Supreme Court stated, “[m]utual knowledge of all the relevant facts gathered by both parties is essential to proper litigation.” Criminal law has been spared of this wisdom.

Cruel and Invisible Punishment

REDEEMING THE COUNTER-MAJORITARIAN EIGHTH AMENDMENT

Aliza Cover†

INTRODUCTION: REDEEMING THE EIGHTH AMENDMENT

Our criminal justice system is in constitutional crisis—a crisis that the courts have yet to recognize. Over the past generation, America has waged an increasingly punitive war on crime,¹ and the casualties of that war have been disproportionately people of color.² Even a casual observer of the American system of punishment would be struck by its racial disparities. Yet the Supreme Court has failed to see a problem of constitutional dimension.³ This judicial blindness is the product of a deficient construction of the Eighth Amendment—a construction that takes its shape from majority norms rather than counter-majoritarian principles.

† Westerfield Fellow, Loyola New Orleans College of Law; J.D., Yale Law School. I am indebted to Hilary Allen, Andrea Armstrong, Trey Drury, Janey Hoeffel, Johanna Kalb, Sandy Mayson, Bidish Sarma, Rob Smith, and Anna VanCleave for their support and feedback as I developed these ideas. Many thanks, as well, to the participants of the Loyola Junior Faculty Forum Works-in-Progress series for their helpful comments. Most of all, I am grateful to Benjamin Plener, my partner in all things, including this.

¹ See, e.g., Craig Haney, Riding the Punishment Wave: On the Origins of Our Devolving Standards of Decency, 9 HASTINGS WOMEN'S L.J. 27 (1998); Bryan A. Stevenson, Confronting Mass Imprisonment and Restoring Fairness to Collateral Review of Criminal Cases, 41 HARV. C.R.-C.L. L. REV. 339, 340-41 (2006) (“In the last thirty-five years, the number of United States residents in prison has increased dramatically—from 330,000 people in jails and prisons in 1972 to almost 2.3 million imprisoned people today. The United States now has the highest rate of incarceration in the world. Almost five million people are on probation and parole in this country. . . . Corrections spending by state and federal governments has risen from $6.9 billion in 1980 to $57 billion in 2001. During the ten year period between 1985 and 1995, prisons were constructed at a pace of one new prison opening each week.” (footnotes omitted)).

² “No other country in the world imprisons so many of its racial or ethnic minorities. The United States imprisons a larger percentage of its black population than South Africa did at the height of Apartheid.” MICHELLE ALEXANDER, THE NEW JIM CROW: MASS INCARCERATION IN THE AGE OF COLORBLINDNESS 6-7 (2010).

With President Reagan’s official declaration of the war on drugs in 1982, America began an unprecedented experiment with mass incarceration. In 1980, American jails and prisons housed just over 500,000 people. Thirty-one years later, that number had reached more than 2.2 million. Drug offenses accounted for a significant portion of this increase. In 1980, 41,000 people were serving time in prisons and jails for drug offenses; by 2011, that number had risen more than tenfold to nearly 500,000. America now imprisons more people, with a higher per capita rate of incarceration, than any other country in the world.

The impact on communities of color has been profound. Nationally, at the end of 2007, one in 11 black adults, but only one in 45 white adults, was under correctional supervision. In 2011, more than three percent of all black males, but only 0.5 percent of white males, were serving time in state and federal prisons. The incarceration rates of young African American males are particularly striking. Nationwide, “between 6.6% and 7.5% of all black males ages 25 to 39 were imprisoned in 2011,” and among 18- to 19-year-olds, “black males were imprisoned at more than 9 times the rate of white males.” In certain parts of the country, the incarceration rate of African Americans is much higher. Wisconsin leads the nation with 12.8%

---

8 Id.
12 Id.
13 Id.; see also id. (“Excluding the youngest and oldest age groups, black males were imprisoned at rates that ranged between 5 and 7 times the rates of white males . . . . In comparison, Hispanic males were imprisoned at 2 to 3 times the rate of white males in 2011 . . . . In 2011, blacks and Hispanics were imprisoned at higher rates than whites in all age groups for both male and female inmates . . . .”).
of black working-age men incarcerated, compared to 1.2% of white men.  

In Milwaukee County, “[o]ver half of African American men in their 30s and half of men in their early 40s have been incarcerated in state correctional facilities.”

This racially disparate impact of mass incarceration and the war on drugs cannot be explained by an easy correlation between race and conduct. In fact, “[s]tudies show that people of all colors use and sell illegal drugs at remarkably similar rates. If there are significant differences to be found in the surveys, they frequently suggest that whites, particularly white youth, are more likely to engage in drug crime than people of color.” Nevertheless, “[i]n some states, black men have been admitted to prison on drug charges at rates twenty to fifty times greater than those of white men.” In 2007, only 14% of regular drug users—but “37% of those arrested for drug offenses and 56% of persons in state prison for drug offenses”—were African American. Moreover, “African Americans serve[d] almost as much time in federal prison for a drug offense (58.7 months) as whites d[id] for a violent offense (61.7 months), largely due to racially disparate sentencing laws such as the 100-to-1 crack-powder cocaine disparity . . . .”

These inequities—though dramatic—are not new. Disparate racial impact is an age-old feature of the American criminal justice system. Punishment was once imposed differentially through overt, legalized racial caste systems, from slavery and the Black Codes to Jim Crow. Today, the source of the racial inequality is more elusive: it is the result of layers of discretionary decision-making and complex socioeconomic and cultural dynamics, both within and without the criminal justice system. As the causality has become less visible and harder to remedy, the scope of the inequality has expanded. The sheer number of people in prison has increased by an order of magnitude, and the corresponding burdens now touch entire communities.

15 Id.
16 ALEXANDER, supra note 2, at 7 (footnotes omitted).
17 Id. (footnotes omitted).
18 MAUER & KING, supra note 7, at 2.
19 Id.
20 See infra Part II.B.
21 High incarceration rates affect not only those imprisoned, but the families and communities they leave behind—depriving children of fathers, parents of their partners, and households of much-needed income. Additionally, many minority and
The stark disparate racial impact of our criminal justice system is a predictable result of majoritarian policymaking and law enforcement without a constitutionally adequate counter-majoritarian check. The Constitution creates a careful balance between majoritarian and counter-majoritarian principles. It establishes a representative democracy accountable to the will of the People, but guards against a purely majoritarian political process that is insensitive, and at times overtly hostile, to the needs of the minority. The founders embedded within the constitutional framework an interconnected set of structural and substantive counter-majoritarian checks to protect against untrammeled majoritarian rule, particularly in the application of criminal punishment. The scope of the Eighth Amendment must be construed in light of this structural counter-majoritarian principle, synthesized with the principle of racial egalitarianism embodied in the Reconstruction Amendments.

Applying these original principles, I argue that the cruel and unusual punishments clause proscribes severe punishments disproportionately imposed upon minorities. From this structural perspective, the word “unusual” not only serves as a license to look at majoritarian “evolving standards of decency” or to curb the “wanton[] and . . . freakish[]” imposition of certain harsh penalties. The word “unusual” also signifies that when a punishment is imposed irregularly against certain suspect classes of people, the Constitution demands action from the judiciary to scrutinize the punishment and, if necessary, correct the distortion wrought by the majoritarian system.

My theory of the Eighth Amendment is motivated by a redemptive approach to constitutional interpretation that resonates with the work of Jack Balkin, among others. In his theory of “framework originalism,” Balkin argues that faithful low-income neighborhoods exist as virtual police states, aggressively monitored in ways that are entirely out-of-keeping with expectations of privacy and security in white, affluent areas. See generally M. Chris Fabricant, War Crimes and Misdemeanors: Understanding “Zero-Tolerance” Policing as a Form of Collective Punishment and Human Rights Violation, 3 DREXEL L. REV. 373 (2011).

22 See infra Part I.A.


24 Furman v. Georgia, 408 U.S. 238, 309-10 (1972) (Stewart, J., concurring) (“These death sentences are cruel and unusual in the same way that being struck by lightning is cruel and unusual. For, of all the people convicted of rapes and murders in 1967 and 1968, many just as reprehensible as these, the petitioners are among a capriciously selected random handful upon whom the sentence of death has in fact been imposed . . . . [T]he Eighth and Fourteenth Amendments cannot tolerate the infliction of a sentence of death under legal systems that permit this unique penalty to be so wantonly and so freakishly imposed.” (footnotes and citations omitted)).
constitutional interpretation is not rigidly tied to the expected application of the original framers, but it must remain tethered to the text of the Constitution and the higher-order principles motivating that text. The framers bequeathed to future generations a document that leaves room for interpretation—and they did so deliberately, so as to enable it to endure over time. The Constitution provides a framework, not a meticulous blueprint. As such, it is possible—and, indeed, necessary—to both keep faith with the original meaning of the Constitution and engage in a project of “constitutional construction” across generations. With the opportunity to construct constitutional meaning comes the responsibility to redeem the motivating principles behind the Constitution to “meet[] the challenges of changing conditions in ways that seek to further the promises and commitments of the plan . . . .”

Under my redemptive interpretation of the Eighth Amendment, the term “unusual” signifies a counter-majoritarian distrust of punishments that bear an irregular impact on those who have been systemically underrepresented in the political process. Because the Eighth Amendment is centered on punishment—the end product of the criminal justice process—its prohibition logically refers to unusual outcomes, not unusual motivations, and thus the appropriate focus is on disparate impact rather than discriminatory intent. Moreover, in an age of de jure race neutrality, where the de facto impact of the criminal justice system remains markedly unequal, a faithful interpretation of

---

26 See, e.g., M’Culloch v. Maryland, 17 U.S. (4 Wheat) 316, 407 (1819) (“[W]e must never forget that it is a constitution we are expounding.”); Paul Brest, The Intentions of the Adopters Are in the Eyes of the Beholder, in The Bill of Rights: Original Meaning and Current Understanding (Eugene W. Hickok, Jr., ed., 1991) (“[A]s Professor Paul Freund once remarked in a class, ‘We ought not read the Constitution like a last will and testament lest it become one.’”); id. at 23 (“[S]hould we not pay the authors the compliment of believing that they meant no more than they said? What they left unsaid, they left open for us to decide . . . . The Constitution has become something in its own right . . . . It has long ceased to be no more than what other men hoped they would do or intended them to do. The Constitution, together with the Court’s work, is not so much pushed by the plans of the past as pulled by hopes of the future. It is not stuffed, but pregnant with meaning.” (quoting Charles Curtis, Lions Under the Throne 3, 7 (1947))).
27 Balkin, supra note 25, at 7.
28 Id. at 75 (“Redemptive constitutionalism is a characteristic feature of the American constitutional tradition. In every generation people have seen injustice in their society and made claims in the name of the Constitution to remedy those injustices. The great political and social movements that secured the basic rights and liberties that Americans now take for granted are examples of redemptive constitutionalism at work.”).
29 Id.
these original constitutional principles in the modern context requires a focus on impact rather than intent.

Doctrinally, this interpretation would demand heightened judicial scrutiny into the cruelty of those punishments that bear a disparate impact upon minorities. This searching review would provide a far more robust counter-majoritarian check than either existing equal protection doctrine, which imposes discriminatory purpose as a gatekeeper to heightened scrutiny,30 or existing Eighth Amendment “gross disproportionality” review, which overtly defers to majoritarian legislative determinations of appropriate punishments.31 Critically, however, the reach of this doctrine would also be carefully cabined. Heightened scrutiny would be a targeted tool reserved for punishments imposed disparately upon minorities, thus preserving traditional deference to legislative determinations whenever possible, and leaving untouched the ordinary discriminatory purpose requirement of equal protection law.

I make my argument in four parts. First, I delve into the original principles behind the Eighth Amendment and the Reconstruction Amendments, and explain how these principles support an interpretation of the cruel and unusual punishments clause that provides robust counter-majoritarian and anti-discrimination protections against excessive punishment. Second, I make the case for an interpretation of the Eighth Amendment that focuses on disparate impact rather than discriminatory intent. I look to the changing historical circumstances and national ethos since Reconstruction and argue that, to remain relevant in the modern age, the Eighth Amendment must be responsive to the contemporary reality of a massive, punitive, formally race-neutral criminal justice system that metes out punishment disproportionately against minorities. I also argue that, as a matter of theory, any effective counter-majoritarian Eighth Amendment jurisprudence logically must reach disparate impact. Third, I critique existing doctrinal interpretations of the Eighth Amendment and equal protection clause as overly majoritarian and insufficiently attentive to the problem of disparate impact in the punishment context. Finally, I begin the conversation about a doctrinal path forward. I propose a two-step judicial inquiry into cruel and

31 See infra Part III.A; see also, e.g., Ewing v. California, 538 U.S. 11, 27-28 (2003) (plurality opinion).
unusual punishments: if a defendant can make a threshold showing that a particular punishment is “unusual”—that it is disproportionately imposed upon minorities—then the court should apply heightened scrutiny in assessing whether the punishment is “cruel.”

I. ORIGINAL PRINCIPLES

In this section, I examine the constitutional principles motivating the Eighth Amendment. I look first to the founding: to the counter-majoritarian and antidiscrimination principles permeating the Constitution, the Bill of Rights, and the Eighth Amendment in particular. I then turn to the principle of racial egalitarianism motivating the Reconstruction Amendments, and argue that a synthesis of these constitutional principles requires an interpretation of the Eighth Amendment that is robustly counter-majoritarian and concerned with discrimination against racial minorities.

A. The Founding

The need for a check on majoritarian rule is one of the core original principles infusing the structure and substance of the Constitution. Understanding the word “unusual” to reach punishments disproportionately imposed against minorities keeps faith with the counter-majoritarian promise of the Bill of Rights and of judicial review itself against the unchecked “tyranny of the majority.”

Scholars and judges have long been troubled by the “countermajoritarian difficulty” inherent when an arguably undemocratic judiciary exercises review over legislative action. As articulated by John Hart Ely, however, and as presaged by footnote four of Carolene Products, it is precisely this

---

34 See generally JOHN HART ELY, DEMOCRACY AND DISTRUST: A THEORY OF JUDICIAL REVIEW (1980).
controversial counter-majoritarian nature of the courts that preserves the rights of minorities from the vagaries of the political process. Unchecked majority rule would lead to entrenchment of certain despised and disadvantaged groups in positions of powerlessness, risking serious and irremediable threats to their individual liberties. Judges should embrace, rather than reject, their counter-majoritarian institutional role, and should interpret the Constitution—and, specifically, the Eighth Amendment—in light of their structural position within our system of government and in light of the counter-majoritarian purpose of the Bill of Rights.36

The founders were well aware of the dangers inherent in majority rule37 and infused our Constitution with strong counter-

---

processes ordinarily to be relied upon to protect minorities, and which may call for a correspondingly more searching judicial inquiry,” (citations omitted)).

36 See W. Va. State Bd. of Educ. v. Barnette, 319 U.S. 624, 638 (1943) (“The very purpose of a Bill of Rights was to withdraw certain subjects from the vicissitudes of political controversy, to place them beyond the reach of majorities and officials and to establish them as legal principles to be applied by the courts. One’s right to life, liberty, and property, to free speech, a free press, freedom of worship and assembly, and other fundamental rights may not be submitted to vote; they depend on the outcome of no elections.”); Consent and Consensus: Appeal for Amendments (speech by James Madison in the House of Representatives, June 8, 1789), in THE MIND OF THE FOUNDER: SOURCES OF THE POLITICAL THOUGHT OF JAMES MADISON 169 (rev. ed. 1973) (“The prescriptions in favor of liberty ought to be leveled against that quarter where the greatest danger lies, namely, that which possesses the highest prerogative of power. But this is not found in either the [E]xecutive or [L]egislative departments of Government, but in the body of the people, operating by the majority against the minority. It may be thought that all paper barriers against the power of the community are too weak to be worthy of attention . . . yet . . . it may be one means to control the majority from those acts to which they might be otherwise inclined.”); id. at 171-72 (“[If the Bill of Rights] are incorporated into the Constitution, independent tribunals of justice will consider themselves in a peculiar manner the guardians of those rights; they will be an impenetrable bulwark against every assumption of power in the Legislative or Executive; they will be naturally led to resist every encroachment upon rights expressly stipulated for in the Constitution by the declaration of rights.”); Julian N. Eule, Judicial Review of Direct Democracy, 99 YALE L.J. 1503, 1530 (1990) (“As James Madison noted, ‘the invasion of private rights is chiefly to be apprehended, not from acts of Government contrary to the sense of its constituents, but from acts in which the Government is the mere instrument of the major number of the Constituents.’ Although initially resistant to the inclusion of a bill of rights, Madison was later to press for congressional adoption of the first ten amendments, recognizing in it yet another device for filtering majoritarian preferences—for it afforded a role for the judiciary in curbing the more immediately responsive and accountable branches.”) (footnotes omitted)).

37 See, e.g., THE FEDERALIST NO. 10 (James Madison) (“When a majority is included in a faction, the form of popular government . . . enables it to sacrifice to its ruling passions or interest both the public good and the rights of other citizens.”); THE FEDERALIST NO. 51 (James Madison) (“It is of great importance in a republic, not only to guard the society against the oppression of its rulers, but to guard one part of the society against the injustice of the other part . . . . If a majority be united by a common interest, the rights of the minority will be insecure.”); see also, Erwin Chemerinsky, The Vanishing Constitution, 103 HARV. L. REV. 43, 75 n.147 (1989).
majoritarian features. Both in the substantive individual rights it guarantees and in the structural separation of powers it affords, the Constitution protects the rights of minorities against the excesses of majority rule, even as it empowers that majority through the establishment of the representative political branches. The framers recognized that a purely majoritarian democratic system creates winners and losers and risks cementing those winners and losers into their relative positions of power. They understood that if the majority has unchecked power to enforce its will, minorities’ individual rights would be vulnerable.

Although the wariness of majority rule runs throughout the Constitution and the Bill of Rights, both common sense and constitutional text identify a special danger when the majority has unlimited power in the criminal context. The Constitution recognizes that the potential for untrammeled abuses by the majority upon the minority is singularly acute in the context of criminal punishment—perhaps the most coercive of all actions a government takes upon its citizens. Punishment is a violent act, legitimized through the neutral trappings of the justice system. It is the ultimate embodiment of the state’s (and the majority’s)

38 These individual rights are contained not only in the Bill of Rights itself but also in the original Constitution of 1787—through the guarantee of habeas corpus, U.S. CONST. art. I, § 9, cl. 2; the prohibition against ex post facto laws and bills of attainder, U.S. CONST. art. I, § 9, cl. 3; and the limitations on treason, U.S. CONST. art. III, § 3, cl. 1. Alexander Hamilton argued that no Bill of Rights was necessary because “[t]he truth is . . . that the Constitution is itself, in every rational sense, and to every useful purpose, A BILL OF RIGHTS.” THE FEDERALIST NO. 84, at 475 (Alexander Hamilton).

39 See, e.g., THE FEDERALIST NO. 47 (James Madison) (“The accumulation of all powers, legislative, executive, and judiciary, in the same hands, whether of one, a few, or many, and whether hereditary, self-appointed, or elective, may justly be pronounced the very definition of tyranny.”).

40 See, e.g., Chemerinsky, supra note 37, at 74-75 (“[T]he framers openly and explicitly distrusted majority rule; virtually every government institution they created had strong anti-majoritarian features. Even more importantly, the Constitution exists primarily to shield some matters from easy change by political majorities. The body of the Constitution reflects a commitment to separation of powers and individual liberties (for example, no ex post facto laws or bills of attainder, no state impairment of the obligation of contracts, no congressional suspension of the writ of habeas corpus except in times of insurrection).” (footnote omitted)).

41 See generally THE FEDERALIST Nos. 10, 51 (James Madison); see also supra note 37.

42 See Mark D. Rosenbaum & Daniel P. Tokaji, Healing the Blind Goddess: Race and Criminal Justice, 98 MICH. L. REV. 1941, 1963 (2000) (“[T]here is textual, historical, and theoretical support for an approach that gives special attention to the command of equality in the area of criminal procedure. The rights of criminal defendants were, of course, a dominant concern of the Framers of the Constitution, provisions protecting these rights being embedded in the Fourth, Fifth, Sixth, and Eighth Amendments, not to mention the Bill of Attainder, Ex Post Facto, and Habeas Corpus Clauses.”).

absolute power over the individual—and the moment at which a counter-majoritarian guarantee is most critical.

The 1787 Constitution, which focuses more on structural protections against majority overreaching than substantive ones, nonetheless contains a number of specific guarantees against arbitrary, excessive, and illegal punishments: it protects against the suspension of habeas corpus,\textsuperscript{44} prohibits ex post facto laws and bills of attainder,\textsuperscript{45} and places important limitations on convictions for treason.\textsuperscript{46} The Bill of Rights strengthens these protections considerably. Four of the 10 constitutional amendments comprising the Bill of Rights relate to individuals’ rights when subjected to criminal prosecution. Three of these amendments—the Fourth, Fifth, and Sixth—deal with fair criminal procedures to which criminal defendants are entitled. The Eighth Amendment, uniquely, provides a substantive guarantee against excesses at the end point of the criminal process: punishment.

While this structural counter-majoritarian impulse and the special concern over majoritarian abuses of the criminal justice system are relatively clear, the legislative intent behind the text of the Eighth Amendment itself is notoriously opaque. The Eighth Amendment was drafted without extensive debate,\textsuperscript{47} and while deciphering original intent is always fraught with uncertainty,\textsuperscript{48} a definitive pronouncement of the original expected application of the Eighth Amendment is particularly precarious. But historical evidence suggests that the ban on “cruel and unusual punishments” was specifically motivated by the need to protect against discriminatory imposition of severe punishments.\textsuperscript{49} There is thus a direct link between the broad constitutional counter-majoritarian principle, the evident concern by the founders over the risks of coercive majoritarian

\textsuperscript{44} U.S. CONST. art. I, § 9, cl. 2.
\textsuperscript{45} U.S. CONST. art. I, § 9, cl. 3.
\textsuperscript{46} U.S. CONST. art. III, § 3, cl. 2.
\textsuperscript{47} See, e.g., Furman v. Georgia, 408 U.S. 238, 244 (1972) (Douglas, J., concurring) (quoting the paltry debate from the Annals of Congress).
\textsuperscript{49} See Furman, 408 U.S. at 242 (Douglas, J., concurring); see generally Laurence Claus, The Antidiscrimination Eighth Amendment, 28 HARV. J.L. & PUB. POL’Y 119 (2004). This reading is in striking contrast to the originalist interpretation advanced by Justices Scalia and Thomas that has focused on the prohibition of torturous methods of punishment that were considered barbarous or outdated at the time of the founding. See, e.g., Baze v. Rees, 553 U.S. 35, 97 (2008) (Thomas, J., concurring).
criminal punishment, and a distinct Eighth Amendment concern about discriminatory punishment.

The text of the Eighth Amendment derives from the English Bill of Rights of 1689. Evidence suggests that the English provision “was concerned primarily with selective or irregular application of harsh penalties and that its aim was to forbid arbitrary and discriminatory penalties of a severe nature.” The English guarantee was motivated at least in part by parliamentary outrage over discriminatory punishment “against those perceived to be political and religious opponents of the Stuart monarchy,” with particular attention to the extraordinary punishments imposed in the politically sensitive libel cases of Titus Oates and Samuel Johnson. Blackstone

50 Furman, 408 U.S. at 242 (Douglas, J., concurring); see also Claus, supra note 49, at 138, 141-42 (“The manner in which the two Houses of Parliament handled the [Titus] Oates and [Samuel] Johnson cases reveals an overarching concern about immorally discriminatory punishments that were harsher than the law allowed. It was the dual character of the Oates and Johnson punishments as cruel and illegal that caused the Parliament to act. The punishments were illegal, unusual and void because they departed from precedent for no morally—and therefore no legally—sufficient reason. The unusualness of the judgments was cited interchangeably with their illegality in Parliament’s deliberations on the judgments and in the Bill of Rights.”).

51 Claus, supra note 49, at 138.

52 Titus Oates, by all accounts, perjuriously accused dozens of people of engaging in a papist conspiracy against the King, with bloody results. IRVING BRANT, THE BILL OF RIGHTS: ITS ORIGIN AND MEANING 132-53 (1965). However, his trial became a symbol of governmental discriminatory excess. He was convicted of libel and perjury and sentenced with “bizarre savagery”; these sentences were set aside as “cruel and illegal” after King James was dethroned in the 1688 Revolution. Id. at 151; Claus, supra note 49, at 139. According to Claus,

Titus Oates’s punishment was unusual not because it successfully articulated a changed understanding of the common law of perjury, but because it was a departure from the common law of perjury. It was unusual because other perjurers were not subjected to it. When a legal system is built on custom, to impose a novel sentence is to impose an illegal sentence. Further, the punishments were “cruel, barbaric, inhuman, and unchristian” because they departed from precedent in the direction of greater severity.

Claus, supra note 49, at 142 (footnotes omitted).

53 “Samuel Johnson had been convicted on two counts of a misdemeanor called seditious libel.” Id. at 137 (footnotes omitted). Among other punishments, he was saddled with imprisonment until he paid a “prohibitively large” fine that he could not afford. Id.

A similar fate had befallen other political enemies of the Stuart kings . . . . Both fines and bail were susceptible of such discriminatory use to effect indefinite imprisonment, and the evil of that use was a primary consideration underlying the Bill of Rights’ condemnation of excessive bail and fines . . . . In 1680, a committee of the House of Commons condemned the King’s courts for an obvious pattern of setting excessive bail and imposing excessive fines. That pattern evidenced discrimination against those perceived to be political and religious opponents of the Stuart monarchy. The committee “resolved that in imposing fines the judges had acted ‘arbitrarily, illegally, and partially,’ and in favor of the Papists.” Seven members of that committee were later to sit on the committees that drafted the Bill of
shared this understanding that the English Bill of Rights of 1689 prohibited arbitrary and discriminatory punishments and fines, and the American founders were, in turn, intimately familiar with and heavily influenced by Blackstone. They were, moreover, well aware of the bloody English history of discriminatory punishment, including the notorious Bloody Assizes, in which hundreds of alleged political dissidents were executed after “pseudo trials.” This history “helped to place constitutional limitations on the crime of treason and to produce a bar against cruel and unusual punishments,” as, “[u]nless barred by fundamental law, the legal rulings that permitted th[ese] result[s] could easily be employed against any person whose political opinions challenged the party in power.” Justice Douglas in Furman recognized that:

Those who wrote the Eighth Amendment knew what price their forebears had paid for a system based, not on equal justice, but on discrimination. In those days the target was not the blacks or the poor, but the dissenters, those who opposed absolutism in government, who struggled for a parliamentary regime, and who opposed governments’ recurring efforts to foist a particular religion on the people. But the tool of capital punishment was used with vengeance against the opposition and those unpopular with the regime. One cannot read this history without realizing that the desire for equality was reflected in the ban against “cruel and unusual punishments” contained in the Eighth Amendment. . . . The

Rights. One of these complained during parliamentary debate in 1680 that “[m]en have been fined, not according to their Crimes, but their Principles: Sometimes because they have been Protestants.”

Id. at 138-39 (footnotes omitted).

54 See 4 WILLIAM BLACKSTONE, COMMENTARIES *371 (“[I]t is moreover one of the glories of our English law, that the nature, though not always the quantity or degree, of punishment is ascertained for every offense; and that it is not left in the breast of any judge, nor even of a jury, to alter that judgment, which the law has beforehand ordained, for every subject alike, without respect of persons. . . . [W]here an established penalty is annexed to crimes, the criminal may read their certain consequence in that law, which ought to be the unvaried rule, as it is the inflexible judge, of his actions.”); id. at *372 (“[H]owever unlimited the power of the court may seem, it is far from being wholly arbitrary; but it’s discretion is regulated by law. For the bill of rights has particularly declared, that excessive fines ought not to be imposed, nor cruel and unusual punishments inflicted: (which had a retrospect to some unprecedented proceedings is the court of king’s bench, in the reign of king James the second).”); see also Claus, supra note 49, at 144-46.


56 BRANT, supra note 52, at 154-58.

57 Id. at 155.
high service rendered by the “cruel and unusual” punishment clause of the Eighth Amendment is to require legislatures to write penal laws that are evenhanded, nonselective, and nonarbitrary, and to require judges to see to it that general laws are not applied sparsely, selectively, and spottily to unpopular groups.\(^{58}\)

In short, when seeking to understand the original principles behind the Eighth Amendment, we should consider, first, the pervasive structural counter-majoritarian imperative within the Constitution of 1787 and the Bill of Rights of 1791; second, the special constitutional solicitude toward individual rights in the criminal justice context; and third, the particular original concern with discriminatory imposition of punishments.

B. Reconstruction

Of course, there is a deep limitation to these original counter-majoritarian and anti-discrimination principles: the explicit protection of the institution of slavery in our founding documents. However, the story of our original constitutional principles does not end with the ratification of the Bill of Rights in 1791. To understand and remain faithful to the principles of the Constitution and the Eighth Amendment, we must next consider the effect of the Reconstruction Amendments—the “Second Founding”\(^{59}\)—upon the original document. The ratification of the Reconstruction Amendments fundamentally altered the structural and thematic imperatives of the Constitution\(^{60}\)—and, moreover, directly imported Fourteenth Amendment principles to the Eighth Amendment when applied against the states through incorporation.\(^{61}\)

Whereas a racial caste system was cemented in the


\(^{60}\) See BRUCE ACKERMAN, WE THE PEOPLE: FOUNDATIONS 82 (1991) (“The new amendments abolishing slavery, guaranteeing the ‘privileges or immunities of citizens of the United States,’ assuring ‘equal protection’ and ‘due process of law,’ safeguarding voting rights against racial discrimination—our modern disagreements about the precise meaning of these provisions should not blind us to the quantum leap the Republicans had made in nationalizing the protection of individual rights against state abridgment.”).

\(^{61}\) See, e.g., Jamal Greene, Fourteenth Amendment Originalism, 71 MD. L. REV. 978, 979 (2012) (“An originalist who believes that the Fourteenth Amendment incorporated against state governments some or all of the rights protected by the Bill of Rights should, in adjudicating cases under incorporated provisions, be concerned primarily (if not exclusively) with determining how the generation that ratified that amendment understood the scope and substance of the rights at issue.”); George C. Thomas III, When Constitutional Worlds Collide: Resurrecting the Framers’ Bill of Rights and Criminal Procedure, 100 MICH. L. REV. 145, 146-48 (2001).
Constitution of 1787 through the protection of slavery, the Thirteenth, Fourteenth, and Fifteenth Amendments signaled a profoundly different constitutional orientation. The Reconstruction Amendments served to enhance the counter-majoritarian structural and substantive imperatives of the 1787 Constitution and the Bill of Rights, and to link those imperatives directly to racial equality. To uncover how the meaning of the original Constitution should be applied today, we must engage in “principled synthesis” to reconcile these constitutional moments.

The Constitution of 1787 and the Bill of Rights were drafted in the aftermath of revolution. Although the Constitution created a democracy rather than a monarchy, the framers had a keen understanding of the potential for abuses of political power, whether held by elected representatives or authoritarian kings. The protections for individual rights in the Constitution and Bill of Rights—from the institution of habeas corpus to the protections of free speech and free exercise of religion—reflected a primary concern with overreaching by the majoritarian political authorities against minority political and religious dissidents. Counter-majoritarian protections for other types of minorities that we recognize today—such as racial, ethnic, and sexual minorities—were simply not part of the constitutional conversation.

The Reconstruction Amendments, by contrast, were drafted in the aftermath of civil war—a war that was waged, at least in part, over the legitimacy of race-based chattel slavery. After the Civil War, the dangers of racial majoritarian tyranny—rather than political majoritarian tyranny—came to the forefront. The Thirteenth, Fourteenth, and Fifteenth Amendments made a significant move from counter-majoritarian protections over the individual rights of political and religious dissidents to counter-majoritarian protections

62 See, e.g., Eric Foner, The Strange Career of the Reconstruction Amendments, 108 YALE L.J. 2003, 2006 (1999) (“Reconstruction represented less a fulfillment of the Revolution’s principles than a radical repudiation of the nation’s actual practice for the previous seven decades. Indeed, it was precisely for this reason that the era’s laws and constitutional amendments aroused such bitter opposition. The underlying principles—that the federal government possessed the power to define and protect citizens’ rights, and that blacks were equal members of the body politic—were striking departures in American law…. The Reconstruction amendments transformed the Constitution from a document primarily concerned with federal-state relations and the rights of property into a vehicle through which members of vulnerable minorities could stake a claim to substantive freedom and seek protection against misconduct by all levels of government.”).

63 ACKERMAN, supra note 60, at 94.
over the individual rights of racial minorities. By outlawing
slavery;\(^\text{64}\) guaranteeing citizenship rights, due process of law,
and equal protection of the laws;\(^\text{65}\) prohibiting race-based
disenfranchisement,\(^\text{66}\) and prioritizing federal power and
oversight over states’ rights,\(^\text{67}\) the Reconstruction Amendments
were designed to dismantle the legalized racial caste system
that existed in the American South. Within this broad effort to
overhaul the racial caste system, there is, moreover, significant
evidence that the drafters of the Reconstruction Amendments
were keenly aware of, and intended to remedy, discriminatory
use of the criminal justice system against African Americans.\(^\text{68}\)

\(^{64}\) U.S. CONST. amend. XIII.

\(^{65}\) U.S. CONST. amend. XIV, § 1.

\(^{66}\) U.S. CONST. amend. XV.

\(^{67}\) U.S. CONST. amend. XIII, § 2; U.S. CONST. amend. XIV, § 5; U.S. CONST.
amend. XV, § 2.

\(^{68}\) Dissenting in McCleskey v. Kemp, Justice Blackmun explained:

[The legislative history of the Fourteenth Amendment reminds us that
discriminatory enforcement of States’ criminal laws was a matter of great
concern for the drafters. In the introductory remarks to its Report to
Congress, the Joint Committee on Reconstruction, which reported out the
Joint Resolution proposing the Fourteenth Amendment, specifically noted:
“This deep-seated prejudice against color... leads to acts of cruelty,
oppression, and murder, which the local authorities are at no pains to
XVII (1866). Witnesses who testified before the Committee presented
accounts of criminal acts of violence against black persons that were not
prosecuted despite evidence as to the identity of the perpetrators.

The Court further stated:

(1866) (testimony of George Tucker, Virginia attorney) (“They have not any
idea of prosecuting white men for offenses against colored people; they do not
appreciate the idea”); id. at 209 (testimony of Dexter H. Clapp) (“Of the
thousand cases of murder, robbery, and maltreatment of freedmen that have
come before me, ... I have never yet known a single case in which the local
authorities or police or citizens made any attempt or exhibited any
inclination to redress any of these wrongs or to protect such persons”); id., at
213 (testimony of J.A. Campbell) (although identities of men suspected of
killing two blacks known, no arrest or trial had occurred); id., pt. III, p. 141
(testimony of Brev. Maj. Gen. Wager Swayne) (“I have not known, after six
months’ residence at the capital of the State, a single instance of a white man
being convicted and hung or sent to the penitentiary for crime against a
negro, while many cases of crime warranting such punishment have been
reported to me”); id., pt. IV, p. 75 (testimony of Maj. Gen. George A. Custer)
(“[I]t is of weekly, if not of daily, occurrence that freedmen are
murdered.... [S]ometimes it is not known who the perpetrators are; but
when that is known no action is taken against them. I believe a white man
has never been hung for murder in Texas, although it is the law”).

McCleskey v. Kemp, 481 U.S. 279, 346-47 & n.2 (1987) (Blackmun, J., dissenting); see
also Rosenbaum & Tokaji, supra note 42, at 1963 (“[C]oncerns regarding the equal
implementation of state criminal laws were a dominant concern at the time of the Civil
War Amendments. The Framers of the Fourteenth Amendment were especially wary of
These Amendments move away from the racial caste system previously in place; offer bolstered protections for individual rights against state infringement; and attempt to remedy the political process defects that permitted the perpetual enslavement of one group of people. Thus, a “one-two synthesis” of the Eighth Amendment and the later Reconstruction Amendments has the primary effect of expanding the original counter-majoritarian and antidiscrimination principles against irregular imposition of harsh punishments to encompass racial minorities.

II. THE CENTRALITY OF DISPARATE IMPACT

One might read all I have written above and be persuaded that the Eighth Amendment prohibits cruel punishments discriminatorily imposed upon minorities—but still remain unconvinced that this prohibition reaches disparate impact in the absence of discriminatory intent. In this section, I explain why a modern interpretation of the Eighth Amendment that remains faithful to original principles must read “unusual punishments” as those that bear a disparate impact on minorities, irrespective of invidious purpose.

Our interpretive project is not complete once we identify original principles. We must next engage in the process of constitutional construction by filling in the framework these principles outline. The phrase “cruel and unusual punishments” is paradigmatic of language that establishes a

southern states denying newly freed blacks through unequal administration of the criminal laws.”). But see Richard M. Re & Christopher M. Re, Voting and Vice: Criminal Disenfranchisement and the Reconstruction Amendments, 121 YALE L.J. 1584, 1628-29 (2012).

69 See ACKERMAN, supra note 60, at 94-95.

70 There is, however, a more specific reference to punishment in the Reconstruction Amendments that bears noting. The Thirteenth Amendment explicitly recognizes the connections between slavery—racial caste—and punishment. The Thirteenth Amendment states: “Neither slavery nor involuntary servitude, except as a punishment for crime whereof the party shall have been duly convicted, shall exist within the United States . . . .” U.S. CONST., amend. XIII, § 1 (emphasis added). Under one reading, the Thirteenth Amendment sanctions the perpetuation of racial caste system under the guise of criminal punishment. This reading, however, is improbable in light of the overall project of the Reconstruction Amendments to eradicate a racial caste system. More plausibly, the Thirteenth Amendment is an effort to de-link the illegitimate racial caste system from the legitimate system of punishment, and recognize that the work conditions involved in slavery could only be a punishment for crime: they could not be linked to racial subjugation. Under this reading, a criminal justice system that serves to perpetuate racial subjugation would be contrary to the spirit of the Thirteenth Amendment.

71 BALKIN, supra note 25, at 21-23.
constitutional principle, rather than a rule. It invites the infusion of interpretive content—the evolution of constitutional meaning over time. As John Hart Ely explained, “[T]he decision to use open-ended language can hardly have been inadvertent.”

The very ambiguity of the word “unusual” informs its interpretive scope. I argue that, both for historical reasons external to the Eighth Amendment and for theoretical reasons internal to it, the broad text and original principles should be read today to extend to disparate impact.

The original principles discussed above developed in an age when government was relatively limited in size and scope and when discrimination—whether against political, religious, or racial minorities—was overt. To remain faithful to the original antidiscrimination and counter-majoritarian principles of the Eighth Amendment and redeem them for the modern age, we must account for how the historical circumstances in which we now apply these principles have changed.

Below I will consider two such historical trends. First, we now live in a post-{{Brown}} world—an ostensibly race-neutral society. If anything, we hold more tightly to the Reconstruction ideal that eschews race as a legitimate marker of difference. Second, de facto racial inequities have persisted nonetheless. And, even as overt racism has been delegitimized, an entirely new field for racial inequality has emerged: mass incarceration and the prison-industrial complex. To fully realize the original counter-majoritarian and antidiscrimination principles of the Eighth Amendment in this modern context, the “unusualness” inquiry must center on disparate racial impact.

This conclusion is bolstered by a theoretical analysis of the counter-majoritarian principle itself. To meaningfully protect against cruel punishment, the Eighth Amendment must reach disparate impact. The dangers of majoritarian overreaching in the criminal context are not limited to those instances when legislators act with a discriminatory purpose. Rather, the structural risks of majoritarian rule arise because the

---

72 Id. at 7 (“[W]e should pay careful attention to the reasons why constitutional designers choose particular kinds of language. Adopters use fixed rules because they want to limit discretion; they use standards or principles because they want to channel politics through certain key concepts but delegate the details to future generations. When the Constitution uses vague standards or abstract principles, we must apply them to our own circumstances in our own time. When adopters use language that delegates constitutional construction to future generations, fidelity to the Constitution requires future generations to engage in constitutional construction. This is the essence of the method of text and principle.”).

73 ELY, supra note 34, at 97.
majority will always seek to externalize costs and internalize benefits. The threat of excessive punishment is accentuated when the majority places burdens disproportionately upon a minority—regardless of discriminatory intent.

A. The Ethos of Equality

Since the passage of the Reconstruction Amendments, our national commitment to the antidiscrimination principle has deepened. In today’s post-\textit{Brown}, post-Civil Rights Revolution America, we have fully incorporated the imperative of race neutrality into our “Higher Law.”

Initially, aside from abolishing the legal practice of slavery, the Reconstruction Amendments had little to no practical effect on racial inequality in the country, and race continued to operate as a constitutionally legitimate marker of differential status.

By the early 1880s the country largely had turned its back on the work of the Reconstruction Congress. Chattel slavery had ended, but in many places that was about it. . . . [A]s the country turned its back on the original commitments of the Civil War Amendments, the courts found altogether new meanings in the clauses of the Fourteenth Amendment, interpretations largely constructed to protect the interests of property holders and interstate businesses. These interpretations, arguably quite different from the original understanding of the Reconstruction Amendments, were dominant for almost half a century.\footnote{Friedman, \textit{supra} note 59, at 1205; \textit{see also} Foner, \textit{supra} note 62, at 2007.}

Jim Crow segregation persisted in the South for nearly 100 years after the end of the Civil War. It was only in the twentieth century that “renewed concern about civil rights led to a rebirth of attention to the original commitments of the Civil War Amendments.”\footnote{Friedman, \textit{supra} note 59, at 1206.} After World War II’s victory against racism abroad, the Warren Court and Civil Rights Movement ushered in an era of internal struggle over the role of racism in American public life. Through a combination of grassroots popular mobilization, judicial activism, and legislative initiative, America took significant strides toward combatting systemic, legalized racism and redeeming the discarded principles of the Reconstruction Amendments.\footnote{\textit{See, e.g.}, \textit{id.} at 1235-36 \textit{CAs the history of the Reconstruction Amendments demonstrates, the only real alternative is to adopt a synthetic understanding of the Constitution. One must holistically take account of the entire Constitution. And one must labor to read that document as it has changed over time. Reading in this way...}}
The social and legal developments of the Civil Rights Revolution fundamentally altered the way that Americans view race, from both constitutional and moral perspectives. To be sure, racism undeniably retains a strong hold on American life. Nevertheless, there has been a dramatic shift in cultural and doctrinal recognition that race should not be a determining factor in one’s chances in life. In Ackermanian terms, the Civil Rights Revolution was a “constitutional moment”—a cultural and political shift that took on the dimension of “higher lawmaking” by “We the People.”

Once sticky and divisive, the constitutional claim that race is an illegitimate marker for differential treatment now forms part of the bedrock of our understanding of the American system of law.

The constitutional assumptions of the New Deal and the civil rights revolution and the success of subsequent social movements for equality have become so thoroughly embedded in our contemporary understandings of the Constitution that they influence what we consider easy cases of constitutional equality and inequality. Constitutional politics has made constructions like Brown and Loving not only easy cases, but foundational to our understanding of the equal protection clause. Yet they were not always so central; at one point they would have been highly controversial or even clearly wrong constructions.

The legitimacy of overt racial discrimination is no longer a mainstream debate in our country. In the post-Civil Rights Revolution era, we have solidified our constitutional

---

77 See ACKERMAN, supra note 60, at 137 (“Brown became a symbol energizing a multiracial coalition of blacks and whites into an escalating political struggle against institutionalized racism. As the 1950’s moved on, this mobilized appeal for racial justice struck deepening chords amongst broadening sectors of the citizenry—enabling the Presidency and Congress of the mid-1960’s finally to transform the embattled judicial pronunciamentos of the mid-1950’s into the Civil Rights Acts of 1964 and 1968 and the Voting Rights Act of 1965. As a consequence, Brown came to possess the kind of numinous legal authority that is, I believe, uniquely associated with legal documents that express the considered judgments of We the People.”); see also Bruce Ackerman & Jennifer Nou, Canonizing the Civil Rights Revolution: The People and the Poll Tax, 63 NW. U. L. REV. 63, 97-110 (2009); Bruce Ackerman, Section Five and the On-Going Canonization of the Civil Rights Revolution, BALKINIZATION (June 22, 2009), http://balkin.blogspot.com/2009/06/section-five-and-on-going-canonization.html.

78 BALKIN, supra note 25, at 231; see also ACKERMAN, supra note 60, at 137; Ackerman & Nou, supra note 77.
commitment to the original principle that race is—or at least should be—irrelevant.

B. The Entrenchment of Inequality

Notwithstanding the growing intolerance for overt racism, post-Reconstruction American history demonstrates the practical difficulty of eradicating the entrenched effects of slavery and the ease with which “legitimate” criminal punishment emerged to fill the space previously occupied by illegitimate systems of legalized racial subjugation such as slavery, the Black Codes, and Jim Crow segregation.

After Reconstruction and the dismantling of slavery, white southerners began the process of “Redemption”—reinstating in practice the racial caste system that had been prohibited by law. One notable example was the rise of “convict leasing,” in which African Americans were arrested for arbitrary reasons, sentenced to fines they could not afford, and subjected to forced labor—virtual slavery—to pay off their debt.\(^\text{79}\) By 1900, in all areas of public life, Jim Crow—a pervasive legalized system of disenfranchisement, segregation, and discrimination—had firmly taken hold in the South.\(^\text{80}\)

After a long and bitter struggle, the Civil Rights Movement of the 1950s and 1960s, culminating in the passage of the Civil Rights Act of 1964 and the Voting Rights Act of 1965, brought an official end to Jim Crow.\(^\text{81}\) Yet once again, the end of legal racial inequality ushered in new and less overt forms of subjugation. Perhaps most notably, the rise of the drug war and the associated explosion of the prison population have marked a new era in America’s racial history: an era in which formal race neutrality coexists with a markedly unequal and unprecedentedly punitive system of mass incarceration.\(^\text{82}\)

Overall, the disparate impact of the criminal justice system today cannot be easily traced to the documented discriminatory intent of individual government officials. By

\(^{79}\) See generally ALEXANDER, supra note 2, at 31; DOUGLAS BLACKMON, SLAVERY BY ANOTHER NAME: THE RE-ENSLAVEMENT OF BLACK PEOPLE IN AMERICA FROM THE CIVIL WAR TO WORLD WAR II (2008); DAVID M. OSHINSKY, WORSE THAN SLAVERY: PARCHMAN FARM AND THE ORDEAL OF JIM CROW JUSTICE (1996).

\(^{80}\) See generally C. VANN WOODWARD, THE STRANGE CAREER OF JIM CROW (commemorative ed. 2002) (detailing the history of Jim Crow discrimination); see also, e.g., ALEXANDER, supra note 2, at 31.

\(^{81}\) ALEXANDER, supra note 2, at 37-38.

\(^{82}\) Id. at 49-57; see also Kenneth B. Nunn, Race, Crime and the Pool of Surplus Criminality: Or Why the “War on Drugs” Was A “War on Blacks”, 6 J. GENDER RACE & JUST. 381 (2002).
and large, the inequities are linked, instead, to a tough-on-crime political climate in which the costs of punitive measures are largely externalized to minority communities; and to the vast discretion placed in the hands of multiple layers of law enforcement officials and judicial decision-makers, many of whom are never required to justify their decisions. Although overt racism has been forced underground, the inequality of the system remains.

Numerous scholars have pointed to the parallels between Jim Crow and our present system of criminal justice, with its far-reaching and disproportionate impact on young African American men and its resulting disenfranchisement and disempowerment of large segments of the black population. State-sanctioned racism has been outlawed by the Constitution. But our punitive modern criminal justice system remains strikingly unequal. At the extreme, on penal plantations in many southern states, criminal punishment manifests as a latter-day reenactment of slavery.

We must not limit the contours of the Eighth Amendment according to the expected applications of framers

83 See, e.g., McCleskey v. Kemp, 481 U.S. 279, 332-33 (1987) (Brennan, J., dissenting) (“[A]mericans share a historical experience that has resulted in individuals within the culture ubiquitously attaching a significance to race that is irrational and often outside their awareness.’ . . . ‘We . . . cannot deny that, 114 years after the close of the War Between the States and nearly 100 years after Strauder, racial and other forms of discrimination still remain a fact of life, in the administration of justice as in our society as a whole. Perhaps today that discrimination takes a form more subtle than before. But it is not less real or pernicious.’ The ongoing influence of history is acknowledged, as the majority observes, by our “unceasing efforts to eradicate racial prejudice from our criminal justice system.’ These efforts, however, signify not the elimination of the problem but its persistence. Our cases reflect a realization of the myriad of opportunities for racial considerations to influence criminal proceedings: in the exercise of peremptory challenges, in the selection of the grand jury, in the selection of the petit jury, in the exercise of prosecutorial discretion, in the conduct of argument, and in the conscious or unconscious bias of jurors” (citations omitted) (some internal quotation marks omitted)); Angela J. Davis, Prosecution and Race: The Power and Privilege of Discretion, 67 FORDHAM L. REV. 13, 25-32 (1998).


85 See supra Introduction.

86 Andrea C. Armstrong, Slavery Revisited in Penal Plantation Labor, 35 SEATTLE U. L. REV. 869, 869-70 (2012) (“In states such as Arkansas, Florida, Louisiana, and Texas, inmates are forced to recreate a practice outlawed in 1865—slavery. For example, Louisiana State Penitentiary in Tunica, Louisiana was originally a slave plantation in the 1840s. It was—and is still—familiarly named ‘Angola,’ reportedly because the best slaves came from that African country. As recently as 1979, inmates were referred to as ‘hands’ in the fields, reminiscent of how masters referred to their slaves before the Civil War.” (footnotes omitted)).
who lived in a world entirely foreign to our own. The original constitutional counter-majoritarian and anti-discrimination principles arose in historical periods marked by overt discrimination against minorities. The Stuart monarchy’s discriminatory punishment of dissidents was plain to see, as was the subjugation of African Americans through slavery in antebellum America.

Times have changed. Our world is one that the founders never envisioned—one that is formally race neutral, but in which race disparity abounds; one in which our incarcerated population is roughly equal to the entire American population at the time of the Declaration of Independence.87

The systemic racial inequality in today’s sprawling criminal justice system is the successor to the discriminatory punishment of dissidents and the legalized differential punishment of minorities. To redeem the original constitutional principles—to make meaningful the counter-majoritarian imperative and the disavowal of racial caste in our society—we must account for the historical differences and consider how those principles apply in the modern context. Today, an interpretation of “unusual punishments” is woefully insufficient if limited to those that are the proven product of intentional discrimination by individual state actors. We must recognize, instead, that “unusual punishments” are those that bear a disparate impact on minorities.

C. The Externalization of Punishment

While our modern context demands an interpretation of the cruel and unusual punishments clause that reaches disparate impact, this interpretation also coheres with the text and original principles of the Eighth Amendment. As a matter of counter-majoritarian theory, it is appropriate—indeed, necessary—to consider the disparate impact of punishments upon minorities rather than discriminatory intent.

First, the primary concern of the Eighth Amendment is the effect of government action on a particular individual—the actual imposition of the punishment, rather than the motivations of the state actor. We all agree that the rack is an unconstitutional punishment, and this is so regardless of whether

87 The Census Bureau estimates that 2.5 million people lived America in July 1776. U.S. Census Bureau, Profile America: Facts for Features, http://www.census.gov/newsroom/releases/archives/facts_for_features_special_editions/cb10-ff12.html. By comparison, in 2011, the total number of people in American jails and prisons was more than 2.2 million. GLAZE & ERIKA PARKS, supra note 5, at 4, 8.
the intent is to harm the flesh or save the soul. While other constitutional protections regarding fairness in the criminal justice system are limited to procedural guarantees, the Eighth Amendment is also concerned with substantive outcomes—with impact.

Second, under a counter-majoritarian theory of rights, the reason for a punishment’s differential impact on minorities is of limited pertinence. Irrespective of discriminatory purpose, a legislator may predictably endorse a punishment that he would consider excessive if applied to people like him because he believes it will apply only to “the other.” In an unchecked majoritarian system, those in power may purposefully—out of hatred, spite, or determination to subjugate—pass legislation that externalizes costs upon a minority population. However, they may also do so unconsciously—because they fail to understand or internalize the costs imposed and lack any incentive to act other than in their own interest. In other words, from a counter-majoritarian perspective, the concern of the Eighth Amendment is not simply that the majority will intentionally target the minority with cruel punishment, but that the majority will tolerate cruelty when applied primarily against minorities. To the politically disadvantaged group of people, the motivation matters little.

The Eighth Amendment contains “a realization that in the context of imposing penalties . . . there is tremendous potential for the arbitrary or invidious infliction of ‘unusually’ severe punishments on persons of various classes other than ‘our own.’”’ If the criminal justice system is structured—whether intentionally or unintentionally—so as to concentrate punishment upon the “other,” the majoritarian political process will fail to protect against excessiveness in the setting or enforcement of those punishments. In fact, the majoritarian political process exacerbates the risk that the punishments set will be excessive because the majority internalizes the benefits of lengthy incarceration while externalizing the costs.

---

88 U.S. CONST. amends. IV, V, VI.
89 ELY, supra note 34, at 97.
90 These benefits to society include even marginal increases in deterrence, incapacitation, and retribution. They also include economic and political benefits that stem from the so-called prison-industrial complex. See, e.g., André Douglas Pond Cummings, “All Eyez on Me”: America’s War on Drugs and the Prison-Industrial Complex, 15 J. GENDER RACE & JUST. 417, 419-20 (2012). There is also a strong political incentive for legislators to appear tough on crime, irrespective of concrete benefits to the community. On the other hand, there are some societal costs to lengthy incarceration, including the expense in a time of fiscal austerity and budget shortfalls.
A simplistic majoritarian democratic theorist would claim that the political process should be sufficient to protect the rights of criminal defendants because criminal sanctions are universally applicable, their severity is set according to community values, and only punishments permitted by law are actually imposed. History, however, reveals the flaw in this logic. As explained by John Hart Ely:

Ely’s second alternative is not the exception, but the rule. It is far too easy—and entirely foreseeable—for the majority to structure its criminal justice system in such a way that people like those in power do not run a realistic likelihood of punishment. The disparate treatment in today’s criminal justice system confirms the risks of an unchecked majoritarian punishment regime. Our nation imposes harsh drug sentences—but few of the thousands of white teenagers and college students who use and distribute drugs are realistically at risk of doing time. The disparities in the operation of the criminal justice system in practice are symptomatic of the dangers inherent in a political process solution against cruel punishment. Irregularity in the imposition of punishments, even when laws are facially neutral, must give rise to constitutional concern.

Thus I argue that the prohibition on “cruel and unusual punishments” must reach disparate impact. In the modern American context, this view would have far-reaching

---

91 Assuming equal enforcement, excessive punishment of minorities can be avoided because the same majoritarian political branches that we elect to represent us will be subject to those same laws—as will we, their constituents. Why would we bind ourselves to the mast of cruel and disproportionate punishments of our own design?

92 ELY, supra note 34, at 173 (emphasis added).

93 For example, in Louisiana, the second marijuana offense is punishable up to five years in prison; the third is punishable up to 20 years in prison. LA. REV. STAT. ANN. §§ 40:966(E)(2)(a), (3) (2012). The fourth marijuana offense triggers a sentence of 20 year to life. LA. REV. STAT. ANN. § 15:529.1(A)(4)(a) (2012).

94 See ALEXANDER, supra note 2, at 7; MAUER & KING, supra note 7, at 2. Note that arrests even for marijuana possession (not distribution) are vastly disparate by race. “Nationally, Blacks are 3.73 times more likely than whites to be arrested for marijuana possession,” despite “roughly equal” rates of marijuana use. AMERICAN CIVIL LIBERTIES UNION, THE WAR ON MARIJUANA IN BLACK AND WHITE 47, 66 (2013), available at https://www.aclu.org/sites/default/files/assets/1114413-mj-report-rfs-rel1.pdf.
significance. Today’s punishment regime is historically unprecedented, both within our nation’s history and in the world today.\textsuperscript{95} The regime is characterized by a high rate of incarceration and entrenched racial disparities, despite a strong ethos of formal race neutrality. If the Eighth Amendment is to play any meaningful role in curbing the excesses of this system, a focus on disparate impact is critical. Yet that focus has been conspicuously absent from constitutional jurisprudence to date; we have, thus far, failed to redeem the Eighth Amendment for modern times. And, as a result, the Eighth Amendment’s relevance has diminished. While retaining importance in the death penalty and prison conditions contexts, the Eighth Amendment has been rendered largely obsolete in the context of prison sentences—which, in our age of mass incarceration, is precisely the area where we most urgently need a counter-majoritarian check.

III. CONSTITUTIONAL CONSTRUCTION: CRUEL AND INVISIBLE PUNISHMENT

In light of these principles and this history, an Eighth Amendment doctrine must have, at a minimum, three features if it is to redeem the Amendment’s original meaning in modern times. First, it must provide a robust counter-majoritarian protection against majoritarian overreaching. Second, in giving shape to that counter-majoritarian principle, it must focus on disparate impact upon minorities, rather than discriminatory intent. Third, it must provide some meaningful bulwark against cruelty—which, in the age of mass incarceration, must include draconian prison sentences.

The Supreme Court’s approach to “cruel and unusual punishments” has failed resoundingly to give substance to these features. The judiciary has paid insufficient attention to the counter-majoritarian and anti-discrimination principles behind the Eighth Amendment and has turned a blind eye to the disproportionate impact of mass incarceration upon poor, minority (particularly African American) communities.

In the 1980s and 1990s, as the wars on drugs and crime were building to a crescendo, a crisis was brewing over the legitimacy of the criminal justice system. During this time period, the Court heard a number of challenges to both the

\textsuperscript{95} The United States incarcerates the most people, at the highest rate, of any country in the world. See, e.g., Liptak, supra note 9.
endemic racial inequality in the criminal justice system and to the draconian sentencing practices introduced by the new tough-on-crime legislation. The Court refused to interpret either the Eighth or the Fourteenth Amendment to reach criminal legislation and sentencing practices that impose long prison terms—in many cases for victimless or nonviolent offenses, especially drug offenses—bearing a disparate impact on minorities.

A. Judicial Blindness to Draconian Prison Sentences

Some tried to challenge the modern phenomenon of mass incarceration through claims of excessive punishment, asserting that the imposition of lengthy prison sentences for relatively minor crimes was grossly disproportionate, in violation of the Eighth Amendment. With only one exception, however, the Supreme Court refused to use the prohibition on “cruel and unusual punishments” as a license to strike down tough-on-crime legislation, including three-strikes laws and harsh prison sentences for drug offenses.\(^{96}\)

Six cases over the past four decades have established the contours of modern Eighth Amendment proportionality review in the age of mass incarceration. In the early 1980s, the Supreme Court saw the first challenges to the constitutionality of the punitive laws that arose out of the “war on drugs” and the “war on crime.” Faced with the opportunity to curb the rising tide of punishment, however, the Court instead severely curtailed judicial scrutiny over harsh prison terms.

While the Supreme Court has taken a more active role in Eighth Amendment regulation of the imposition of the death penalty\(^ {97}\) and, in recent years, life without parole for juveniles,\(^ {98}\)


\(^{97}\) “The Court has... invoked proportionality to declare that capital punishment—though not unconstitutional per se—is categorically too harsh a penalty to apply to certain types of crimes and certain classes of offenders.” Graham v. Florida, 560 U.S. 48, 100-01 (2010) (citing cases).

\(^{98}\) Even in the death penalty context, however, the Court pays heavy deference to legislative determinations. See, e.g., McCleskey v. Kemp, 481 U.S. 279, 319 (1987) (“McCleskey’s arguments are best presented to the legislative bodies. It is not the responsibility—or indeed even the right—of this Court to determine the appropriate punishment for particular crimes. It is the legislatures, the elected representatives of the people, that are ‘constituted to respond to the will and consequently the moral values of the people.’ Legislatures also are better qualified to weigh and ‘evaluate the results of statistical studies in terms of their own local conditions and with a flexibility of approach...”\)
the Court has explicitly reserved decisions about length of incarceration to majoritarian legislatures in all but the most extreme circumstances. Over some objection, most vigorously by Justice Scalia,\(^99\) the Court continues to pay lip service to the notion that the Eighth Amendment provides a narrow substantive guarantee against punishments that are grossly disproportionate to the offense committed, irrespective of majority consensus.\(^100\) With the rarest exceptions, however, the Court has veered sharply away from the morass of legislative judgments about appropriate terms of incarceration and has refused to conduct any meaningful independent review of the proportionality of prison sentences to the crimes they punish.

First, in *Rummel v. Estelle*, the Court upheld a life sentence under Texas’s recidivist statute where the triggering felony was “obtaining $120.75 by false pretenses,” and the defendant’s previous convictions were for “fraudulent use of a

---


\(^{99}\) *Ewing*, 538 U.S. at 31 (Scalia, J., concurring) (“Out of respect for the principle of stare decisis, I might nonetheless accept . . . that the Eighth Amendment contains a narrow proportionality principle if I felt I could intelligently apply it. This case demonstrates why I cannot.”); *Harmelin*, 501 U.S. at 965 (opinion of Scalia, J.) (“We conclude from this examination that *Solem* was simply wrong; the Eighth Amendment contains no proportionality guarantee.”); see also *Solem* v. Helm, 463 U.S. 277, 310 (1983) (Burger, C.J., dissenting) (“In short, *Rummel* held that the length of a sentence of imprisonment is a matter of legislative discretion; this is so particularly for recidivist statutes. I simply cannot understand how the Court can square *Rummel* with its holding that ‘a criminal sentence must be proportionate to the crime for which the defendant has been convicted.’”).

\(^{100}\) *Lockyer*, 538 U.S. at 72 (“Through this thicket of Eighth Amendment jurisprudence, one governing legal principle emerges as ‘clearly established’ under § 2254(d)(1): A gross disproportionality principle is applicable to sentences for terms of years.”); see also *Harmelin*, 501 U.S. at 997 (Kennedy, J., concurring in part and concurring in the judgment) (“Our decisions recognize that the Cruel and Unusual Punishments Clause encompasses a narrow proportionality principle . . . . Since *Weems*, we have applied the principle in different Eighth Amendment contexts. Its most extensive application has been in death penalty cases . . . . The Eighth Amendment proportionality principle also applies to noncapital sentences.”).

Even Chief Justice Burger suggested that an extreme disproportion such as life imprisonment for failure to pay a parking meter would violate the Eighth Amendment, and that the Court might properly declare it so in the exercise of its own judgment, although he maintained that “[i]n all other cases, we should defer to the legislature’s line-drawing.” *Solem*, 463 U.S. at 311 n.3 (1983) (Burger, C.J., dissenting) (“*Both Rummel and Hutto* . . . leave open the possibility that in extraordinary cases—such as a life sentence for overtime parking—it might be permissible for a court to decide whether the sentence is grossly disproportionate to the crime. I agree that the Cruel and Unusual Punishments Clause might apply to those rare cases where reasonable men cannot differ as to the inappropriateness of a punishment. In all other cases, we should defer to the legislature’s line-drawing. However, the Court does not contend that this is such an extraordinary case that reasonable men could not differ about the appropriateness of this punishment.”).
credit card to obtain $80 worth of goods or services” and “passing a forged check in the amount of $28.36.”\textsuperscript{101} While acknowledging that “[t]his Court has on occasion stated that the Eighth Amendment prohibits imposition of a sentence that is grossly disproportionate to the severity of the crime,”\textsuperscript{102} the Court rejected the general availability of judicial review over the proportionality of felony prison sentences: “[O]ne could argue without fear of contradiction by any decision of this Court that for crimes concededly classified and classifiable as felonies, . . . the length of the sentence actually imposed is purely a matter of legislative prerogative.”\textsuperscript{103} Two years later, in \textit{Hutto v. Davis}, the Court denied habeas corpus relief to a defendant serving a 40-year prison sentence for possession with intent to distribute and distribution of less than nine ounces of marijuana.\textsuperscript{104} Refusing to engage in any factual inquiry into the proportionality of the punishment, the Court clarified that “\textit{Rummel} stands for the proposition that federal courts should be ’reluctan[t] to review legislatively mandated terms of imprisonment,’ and that ‘successful challenges to the proportionality of particular sentences’ should be ’exceedingly rare.’”\textsuperscript{105}

Only a year later, however, without overturning \textit{Rummel} or \textit{Hutto}, the Court made an about-face and offered a fleeting suggestion that it would engage in a reinvigorated constitutional inquiry into the excessiveness of prison sentences. In \textit{Solem v. Helm}, the Court reaffirmed an Eighth Amendment proportionality principle that was deeply rooted in the common law, firmly established at the time of the founding and acknowledged in Supreme Court precedent for nearly one hundred years,\textsuperscript{106} and explicitly rejected the “assertion that the general principle of proportionality does not apply to felony prison sentences.”\textsuperscript{107} Applying “objective criteria,”\textsuperscript{108} the Court reversed Helm’s life sentence without the possibility of parole for “uttering a ‘no account’ check for $100”—the seventh nonviolent felony in his record.\textsuperscript{109} While retaining a tether to

\textsuperscript{101} \textit{Rummel}, 445 U.S. at 265-66.
\textsuperscript{102} \textit{Id.} at 271.
\textsuperscript{103} \textit{Id.} at 274.
\textsuperscript{105} \textit{Id.} at 374 (citations omitted).
\textsuperscript{107} \textit{Id.} at 288.
\textsuperscript{108} These criteria included: “(i) the gravity of the offense and the harshness of the penalty; (ii) the sentences imposed on other criminals in the same jurisdiction; and (iii) the sentences imposed for commission of the same crime in other jurisdictions.” \textit{Id.} at 292.
\textsuperscript{109} \textit{Id.} at 279, 281.
majoritarian norms through its consideration of “objective criteria,” Solem presaged a far more robust role for the courts in counter-balancing growing punitiveness in the political sphere and a judicial grappling, in constitutional terms, with the apparent real-world justice problem of mass incarceration.

This promise proved short-lived. In a series of three decisions beginning in the early 1990s, the Court, without overruling Solem, effectively neutralized it by emphasizing the narrowness of the proportionality principle and by paying heavy deference to legislatures. A splintered Court in Harmelin v. Michigan upheld a life sentence without parole for possession of more than 650 grams of cocaine.\(^\text{110}\) Seven of the justices affirmed the continued relevance of at least a “narrow proportionality principle that has existed in our Eighth Amendment jurisprudence for 80 years.”\(^\text{111}\) Yet the tenor of the discussion had changed. Justice Kennedy, in concurrence, articulated four principles pertaining to proportionality review that, taken as a whole, emphasized deference to legislative judgments.\(^\text{112}\) The dissenters strenuously critiqued the erosion of the three-part test established in Solem and the narrowing of the proportionality inquiry, which, they asserted, brought it to the point of evisceration.\(^\text{113}\)

Ewing v. California\(^\text{114}\) and Lockyer v. Andrade,\(^\text{115}\) handed down on the same day in 2003, cemented the retraction from Solem. In Ewing, the Court, again without a majority opinion, upheld a sentence of 25 years to life under California’s three-strikes law for a recidivist convicted of stealing three golf clubs, each worth approximately $400.\(^\text{116}\) The plurality considered the factors articulated by Justice Kennedy in Rummel and, emphasizing the deference owed to the legislature, essentially

---


\(^{111}\) Id. at 996 (Kennedy, J., concurring in part and concurring in the judgment); id. at 1009-27 (White, J., dissenting); id. at 1027 (Marshall, J., dissenting).

\(^{112}\) Id. at 1001 (Kennedy, J., concurring in part and concurring in the judgment) (identifying four objective factors—“the primacy of the legislature, the variety of legitimate penological schemes, the nature of our federal system, and the requirement that proportionality review be guided by objective factors”—all of which “inform the final one: The Eighth Amendment does not require strict proportionality between crime and sentence. Rather, it forbids only extreme sentences that are ‘grossly disproportionate’ to the crime.”).

\(^{113}\) Id. at 1018 (White, J., dissenting) (“While Justice Scalia seeks to deliver a swift death sentence to Solem, Justice Kennedy prefers to eviscerate it, leaving only an empty shell. The analysis Justice Kennedy proffers is contradicted by the language of Solem itself and by our other cases interpreting the Eighth Amendment.”).


\(^{115}\) 538 U.S. 63 (2003).

\(^{116}\) Ewing, 538 U.S. at 18, 30-31.
articulated a rational or reasonable basis test for determining whether the severity of the sentence is constitutional. Justice O'Connor wrote, “We do not sit as a ‘superlegislature’ to second-guess [the state’s] policy choices. It is enough that the State of California has a reasonable basis for believing that dramatically enhanced sentences for habitual felons ‘advance[s] the goals of [its] criminal justice system in any substantial way.”

In Lockyer, the Court also denied relief, this time under the deferential AEDPA standard, to a recidivist serving “two consecutive sentences of 25 years to life” under California’s three-strikes law for “stealing approximately $150 in videotapes.” In so holding, the Court explained that “[t]he only relevant clearly established law... is the gross disproportionality principle, the precise contours of which are unclear, applicable only in the ‘exceedingly rare’ and ‘extreme’ case.”

These six cases constitute a—self-admittedly—muddled body of precedent on the scope of defendants’ rights against grossly disproportionate punishment. Solem, though still good law, is an outlier. The overall thrust of the doctrine emphasizes the narrowness of the proportionality principle, the rarity of its applicability, and the near-absolute deference owed to legislative determinations about the appropriateness of even the harshest sentences for relatively minor misconduct.

117 Id. at 28 (citations omitted).
120 Id. at 73 (citations omitted).
121 See id. at 72.
122 A more robust proportionality analysis may be suggested by the Court’s recent consideration of the constitutionality of sentences of life without parole for juveniles. See Miller v. Alabama, 132 S. Ct. 2455 (2012) (holding mandatory sentence of life without parole unconstitutional for juveniles convicted of homicide); Graham v. Florida, 560 U.S. 48 (2010) (holding life without parole unconstitutional for juveniles in non-homicide offenses). However, in these cases the Court consciously located itself at the intersection of the categorical Eighth Amendment death penalty cases and the case-specific “gross disproportionality” cases. See Miller, 132 S. Ct. at 2463-64 (explaining the “two strands of precedent reflecting our concern with proportionate punishment” and concluding that “the confluence of these two lines of precedent leads to the conclusion that mandatory life-without-parole sentences for juveniles violate the Eighth Amendment”); Graham, 560 U.S. 48 at 2021-23 (“Here, in addressing the question presented, the appropriate analysis is the one used in cases that involved the categorical approach, specifically Atkins, Roper, and Kennedy.”). In Miller, the Court went further and explicitly aligned juvenile life without parole cases with death penalty jurisprudence, capping Miller’s reach and limiting the tension it created with Harmelin. Miller, 132 S. Ct. at 2470 (“Harmelin had nothing to do with children and did not purport to apply its holding to the sentencing of juvenile offenders. We have by now held on multiple occasions that a sentencing rule permissible for adults may not be so for children... So if (as Harmelin recognized) ‘death is
Court has explicitly reserved determinations of the proportionality of prison sentences to the legislatures, and in large measure has extracted itself (and the Eighth Amendment) from that balancing inquiry. In practice, it is no longer a constitutional question whether a sentence is disproportionate to the offense; it is a majoritarian legislative judgment. As a result of this judicial reluctance to review majoritarian determinations of appropriate punishment, the modern “punishment wave”\textsuperscript{123} has swelled in a judicial void.

B. Judicial Blindness to the Counter-Majoritarian Principle

The Supreme Court’s laissez-faire approach to the growing punitiveness in the criminal justice system seems to be due in part to a pragmatic anxiety over how to make principled determinations of disproportionality.\textsuperscript{124} It is also, however, part and parcel of a larger, traditionally majoritarian understanding of the cruel and unusual punishments clause—one that abdicates the robust counter-majoritarian role that guarantee should serve.

The prevailing interpretation can be traced to Chief Justice Warren, who first looked to contemporary societal mores to identify “cruel and unusual punishments”—an overtly different, ’children are different too…. Our ruling thus neither overrules nor undermines nor conflicts with Harmelin.’). The Court’s apparently deliberate choice to locate Miller and Graham within or near to its capital punishment jurisprudence evinces its intent to limit these holdings to the juvenile context and not to change the substance of “gross disproportionality” doctrine. If this prediction does not bear out, these cases will certainly have an important impact in changing the landscape of proportionality review of prison sentences.

\textsuperscript{123} Haney, \textit{supra} note 1.

\textsuperscript{124} See, e.g., Harmelin v. Michigan, 501 U.S. 957, 986 (1991) (Scalia, J.) (“The real function of a constitutional proportionality principle, if it exists, is to enable judges to evaluate a penalty that some assemblage of men and women has considered proportionate—and to say that it is not. For that real-world enterprise, the standards seem so inadequate that the proportionality principle becomes an invitation to imposition of subjective values.”); Rummel v. Estelle, 445 U.S. 263, 275-76 (1980) (“But a more extensive intrusion into the basic line-drawing process that is pre-eminently the province of the legislature when it makes an act criminal would be difficult to square with the view expressed in Coker that the Court’s Eighth Amendment judgments should neither be nor appear to be merely the subjective views of individual Justices. . . . [T]o recognize that the State of Texas could have imprisoned Rummel for life if he had stolen $5,000, $50,000, or $500,000, rather than the $120.75 that a jury convicted him of stealing, is virtually to concede that the lines to be drawn are indeed ‘subjective,’ and therefore properly within the province of legislatures, not courts.”); \textit{see also} Solem v. Helm, 463 U.S. 277, 314-15 (1983) (Burger, C.J., dissenting).
majoritarian viewpoint. The notion of “evolving standards of decency” originated in Trop v. Dulles, a case in which the Court struck down as “cruel and unusual” the punishment of denationalization for military desertion. Trop did not explicitly locate the notion of “evolving standards of decency” in the word “unusual,” but rather read the words “cruel and unusual” together to conclude that “[t]he Amendment must draw its meaning from the evolving standards of decency that mark the progress of a maturing society.” The primary effect of this standard has been to link the Eighth Amendment’s protection to practices that are out of the ordinary, with majority norms providing the reference point. Even while recognizing that the Eighth Amendment contains a transcendent connection to the dignity of man, which outlasts the vagaries and fluctuations of contemporary societal norms, the Court has increasingly insisted upon heavy deference to legislative determinations about the appropriateness of punishments, in part because of its majoritarian focus on “evolving standards of decency.”

For example, the Court has prohibited the imposition of the death penalty against certain classes of people where such punishment was deemed inconsistent with “evolving standards of decency.” To identify these evolving standards, the Court

---

125 See, e.g., Roper v. Simmons, 543 U.S. 551, 560-61 (2005) (“[W]e have established the propriety and affirmed the necessity of referring to ‘the evolving standards of decency that mark the progress of a maturing society’ to determine which punishments are so disproportionate as to be cruel and unusual” (citing Trop v. Dulles, 356 U.S. 86, 100-01 (1958) (plurality opinion)).


127 Id. at 100-01 n.32 ("Whether the word ‘unusual’ has any qualitative meaning different from ‘cruel’ is not clear. On the few occasions this Court has had to consider the meaning of the phrase, precise distinctions between cruelty and unusualness do not seem to have been drawn . . . . These cases indicate that the Court simply examines the particular punishment involved in light of the basic prohibition against inhuman treatment, without regard to any subtleties of meaning that might be latent in the word ‘unusual.’ . . . If the word ‘unusual’ is to have any meaning apart from the word ‘cruel,’ however, the meaning should be the ordinary one, signifying something different from that which is generally done.") (citations omitted).

128 Id. at 100-01.

129 See, e.g., Gregg v. Georgia, 428 U.S. 153, 173 (1976) (plurality opinion) (“But our cases also make clear that public perceptions of standards of decency with respect to criminal sanctions are not conclusive. A penalty also must accord with ‘the dignity of man,’ which is the ‘basic concept underlying the Eighth Amendment.’ This means, at least, that the punishment not be ‘excessive.’”) (citations omitted).

130 Id. at 174-76.

has looked primarily to state legislative trends,\textsuperscript{132} as well as jury verdicts\textsuperscript{133} and, in some cases, international legal practices.\textsuperscript{134} In the death penalty and juvenile context, the Court has also brought its own judgment to bear in determining the constitutionality of the punishment,\textsuperscript{135} but has, as a general matter, been loath to assert its independent judgment without the backing of broader societal trends.\textsuperscript{136}

\textsuperscript{132} See, e.g., \textit{Roper}, 543 U.S. at 564 ("The beginning point is a review of objective indicia of consensus, as expressed in particular by the enactments of legislatures that have addressed the question."); \textit{Atkins}, 536 U.S. at 312 ("Proportionality review under those evolving standards should be informed by objective factors to the maximum possible extent . . . . We have pinpointed that the clearest and most reliable objective evidence of contemporary values is the legislation enacted by the country's legislatures." (internal citations omitted) (internal quotation marks omitted)).

\textsuperscript{133} \textit{Gregg}, 428 U.S. at 181 ("The jury also is a significant and reliable objective index of contemporary values because it is so directly involved . . . . The Court has said that 'one of the most important functions any jury can perform in making . . . a selection (between life imprisonment and death for a defendant convicted in a capital case) is to maintain a link between contemporary community values and the penal system.'" (citations omitted)).

\textsuperscript{134} See, e.g., \textit{Roper}, 543 U.S. at 575 ("Our determination that the death penalty is disproportionate punishment for offenders under 18 finds confirmation in the stark reality that the United States is the only country in the world that continues to give official sanction to the juvenile death penalty. This reality does not become controlling, for the task of interpreting the Eighth Amendment remains our responsibility. Yet at least from the time of the Court's decision in \textit{Trop}, the Court has referred to the laws of other countries and to international authorities as instructive for its interpretation of the Eighth Amendment's prohibition of "cruel and unusual punishments." (citation omitted)).

\textsuperscript{135} The Court acknowledges that "the objective evidence, though of great importance, did not 'wholly determine' the controversy, 'for the Constitution contemplates that in the end our own judgment will be brought to bear on the question of the acceptability of the death penalty under the Eighth Amendment.'" \textit{Atkins}, 536 U.S. at 312 (citing \textit{Coker v. Georgia}, 433 U.S. 584, 597 (1977)). In this way, review of disproportionality in the death penalty context has retained some limited measure of independent judicial review aside from majoritarian preferences.

\textsuperscript{136} There is some recent suggestion that the Court may be increasingly willing to conduct an independent review, even in the absence of compelling evidence of societal consensus. Unlike in \textit{Roper}, \textit{Atkins}, and \textit{Kennedy v. Louisiana}, 554 U.S. 407 (2008), in which the Court focused on "evolving standards of decency" to identify sentencing practices that the majority or a growing number of states had deemed unacceptable, the Court in \textit{Miller} struck down mandatory juvenile life without parole for homicide offenses notwithstanding evidence that there was neither nationwide consensus against nor a trend away from the practice. The Court justified its conclusion as a simple extension of existing precedent:

\begin{quote}
[O]ur decision flows straightforwardly from our precedents: specifically, the principle of \textit{Roper}, \textit{Graham}, and our individualized sentencing cases that youth matters for purposes of meting out the law's most serious punishments. When both of those circumstances have obtained in the past, we have not scrutinized or relied in the same way on legislative enactments.
\end{quote}

\textit{Miller v. Alabama}, 132 S. Ct. 2455, 2471 (2012). The Court, moreover, noted that several previous Eighth Amendment cases found similarly tentative evidence of national consensus to be sufficient. \textit{Id.} In his dissent, Justice Alito roundly criticized the majority and asserted that "our Eighth Amendment cases are no longer tied to any objective indicia of society's standards." \textit{Id.} at 2490 (Alito, J., dissenting).
Thus when the Court banned the execution of mentally retarded offenders in Atkins\(^{137}\) and of juveniles in Roper,\(^{138}\) after having reached precisely opposite conclusions only a few years earlier in Penry v. Lynaugh\(^{139}\) and Stanford v. Kentucky,\(^{140}\) the Court did not overrule those previous decisions. Rather, punishments that used to be constitutional given the majoritarian preferences of the day were no longer constitutional given the societal trends away from those punishments.\(^{141}\) Contemporary

Some have argued that, in any event, the Court’s consideration of “objective indicia” is a mere charade, and that the Court reaches the conclusion it wishes to reach, irrespective of the national data. See, e.g., Atkins, 536 U.S. at 348-49 (Scalia, J., dissenting) (“Beyond the empty talk of a ‘national consensus,’ the Court gives us a brief glimpse of what really underlies today’s decision: pretension to a power confined neither by the moral sentiments originally enshrined in the Eighth Amendment (its original meaning) nor even by the current moral sentiments of the American people. [T]he Constitution,’ the Court says, ‘contemplates that in the end our own judgment will be brought to bear on the question of the acceptability of the death penalty under the Eighth Amendment.’ . . . The arrogance of this assumption of power takes one’s breath away. And it explains, of course, why the Court can be so cavalier about the evidence of consensus. It is just a game, after all.” (citation omitted)); Roper, 543 U.S. at 611 (Scalia, J., dissenting) (“The attempt by the Court to turn its remarkable minority consensus into a faux majority . . . . is an act of nomological desperation.”); John F. Stinneford, The Original Meaning of “Unusual”: The Eighth Amendment As A Bar to Cruel Innovation, 102 N.W. U. L. REV. 1739, 1757 (2008) (“The differing outcomes in Stanford and Roper demonstrate the inherent instability and manipulability of the evolving standards of decency test . . . . Any change in societal attitudes between Stanford and Roper was incremental at best; in both cases societal attitudes about the acceptability of executing seventeen-year-olds were split nearly down the middle. The only real difference between these cases lies not in any ‘evolution’ of societal standards, but in an increased assertiveness of judicial will. The Roper majority wanted to strike down the death penalty for seventeen-year-olds, despite the fact that the evidence did not demonstrate that such executions violated any societal moral consensus, at least within the United States, and so it simply pretended that the evidence supported the desired result.”).

---

\(^{137}\) Atkins, 536 U.S. 304.

\(^{138}\) Roper, 543 U.S. 551.

\(^{139}\) 492 U.S. 302, 302-03 (1989).

\(^{140}\) 492 U.S. 361 (1989).

\(^{141}\) Roper, 543 U.S. at 563 (“Three Terms ago the subject [of execution of mentally retarded offenders] was reconsidered in Atkins. We held that standards of decency have evolved since Penry and now demonstrate that the execution of the mentally retarded is cruel and unusual punishment. The Court noted objective indicia of society’s standards, as expressed in legislative enactments and state practice with respect to executions of the mentally retarded. When Atkins was decided only a minority of States permitted the practice, and even in those States it was rare. On the basis of these indicia the Court determined that executing mentally retarded offenders ‘has become truly unusual, and it is fair to say that a national consensus has developed against it.’” (internal citations omitted)); see also Stinneford, supra note 136, at 1741 (“In Atkins v. Virginia and Roper v. Simmons, the Supreme Court appeared to agree that the imposition of the death penalty on the mentally retarded and on seventeen-year-olds respectively was not cruel and unusual punishment in 1989, when Penry v. Lynaugh and Stanford v. Kentucky were decided. Nonetheless, the Court held that such punishments are cruel and unusual today. As Justice Scalia stated in his Roper dissent, the decisions in Atkins and Roper are based on the proposition ‘that the meaning of our Constitution has changed over the past 15 years—not, mind you, that
majoritarian preferences changed the scope of the individual constitutional right.

Understood thus as a majoritarian standard, the constitutional protections for individuals under the Eighth Amendment are, perversely, most robust when society is predisposed against a particular punishment. The Court may intervene only when cruel punishments are moving or have already moved out of favor in society at large. It is precisely at the moment where a legislative solution seems plausible that a judicial remedy becomes available. The Court’s theory of the Eighth Amendment thus cedes control over the scope of a substantive individual right to the whims of the majority. Recognizing this perversity, numerous scholars have critiqued the concept of “evolving standards of decency” as insufficient to protect individuals against unconstitutional punishments.

Under a majoritarian Eighth Amendment, individual rights are protected only insofar as society’s standards of decency are in fact evolving rather than devolving—only insofar as we are in fact making “progress” as a “maturing society,” and not regressing to more vindictive, cruel, and even barbaric days of old. As Justice Scalia is wont to caution us: there is no such assurance of “progress.” Nor, as Chief Justice Roberts stated in Miller, is there any assurance that individuals agree that progress and decency correspond with leniency. Trop is written with Warren Court optimism that this Court’s decision 15 years ago was wrong, but that the Constitution has changed.”

142 Ely, supra note 34, at 69 (“[I]t makes no sense to employ the value judgments of the majority as the vehicle for protecting minorities from the value judgments of the majority.”); Stinneford, supra note 136, at 1754 n.81 (“A number of scholars have previously pointed out the cruel irony inherent in the fact that the evolving standards of decency test ties the rights of criminal defendants to the very same majority opinion from which the Eighth Amendment is supposed to protect them.” (citing scholarship)).
145 Miller v. Alabama, 132 S. Ct. 2455, 2478 (2012) (Roberts, C.J., dissenting) (“Mercy toward the guilty can be a form of decency, and a maturing society may abandon harsh punishments that it comes to view as unnecessary or unjust. But decency is not the same as leniency. A decent society protects the innocent from violence. A mature society may determine that this requires removing those guilty of the most heinous murders from its midst, both as protection for its other members and as a concrete expression of its standards of decency. As judges we have no basis for deciding that progress toward greater decency can move only in the direction of easing sanctions on the guilty.”).
society is, indeed, maturing, and that it is heading toward a state of progressive enlightenment rather than crotchety old age. Today’s mass incarceration, however, suggests a different trajectory. The nationwide trends toward harsh three-strikes laws and lengthy prison sentences for drug crimes indicate that we have grown more punitive, not less. If the scope of Eighth Amendment rights depends on society’s tolerance for cruelty, the protection afforded by these rights will always be precarious.

C. Judicial Blindness to Disparate Impact

As the Court has refused to intervene in the imposition of harsh prison sentences, it has simultaneously rebuffed challenges to the systemic disparate impact of the criminal justice system against minorities. The Court has required a showing of discriminatory intent to trigger heightened scrutiny in the equal protection context and a showing of individualized discrimination infecting the sentencing decision in the Eighth Amendment context. In so doing, the Court has subverted its own counter-majoritarian role and neutralized the efficacy of these constitutional guarantees against modern systemic inequality in the age of mass incarceration.146

_McCleskey v. Kemp_ is paradigmatic of the Court’s narrow judicial construction in this area.147 In _McCleskey_, the Court roundly rejected claims that the imposition of the death

---

146 In spite of the overall trend in the opposite direction, some Supreme Court precedent supports a counter-majoritarian vision of the Eighth Amendment that would advance racial equality. Even _McCleskey_ implicitly recognized that, if it could be shown that death sentences were imposed because of racial bias, they would be unconstitutional. See _McCleskey v. Kemp_, 481 U.S. 279, 313 (1987). And, although they have not won the day, more robust counter-majoritarian interpretations of the Eighth Amendment have emerged at times in the past—most notably, in Justice Douglas’s _Furman_ concurrence, in which he made a compelling case for a strong Eighth Amendment protection against arbitrary imposition of punishments against minorities. _Furman v. Georgia_, 408 U.S. 238, 240-57 (1972) (Douglas, J., concurring). Although he did not explicitly state that the relevant determination should be disparate impact rather than discriminatory intent, he suggested a primary concern with the former. For example, he wrote:

_We cannot say from facts disclosed in these records that these defendants were sentenced to death because they were black. Yet our task is not restricted to an effort to divine what motives impelled these death penalties. Rather, we deal with a system of law and of justice that leaves to the uncontrolled discretion of judges or juries the determination whether defendants committing these crimes should die or be imprisoned. Under these laws no standards govern the selection of the penalty. People live or die, dependent on the whim of one man or of 12._

_Id. at 253.

147 _481 U.S. 279 (1987)._
penalty was unconstitutional under the Eighth and Fourteenth Amendments, in the face of a comprehensive statistical study of the capital punishment system in Georgia demonstrating that the death penalty was imposed disproportionately based on the race of the victim and, to a lesser extent, the race of the accused.\textsuperscript{148} McCleskey’s Fourteenth Amendment equal protection claim failed because, notwithstanding this study, he could not establish discriminatory purpose—a requirement of traditional equal protection doctrine.\textsuperscript{149} The Court then rejected McCleskey’s Eighth Amendment argument that “the Georgia capital punishment system is arbitrary and capricious in application, and therefore his sentence is excessive, because racial considerations may influence capital sentencing decisions in Georgia.”\textsuperscript{150} Accepting that a certain degree of disparate impact in the imposition of punishment was inevitable,\textsuperscript{151} the Court refused to recognize racial bias as the cause of such inequity:

Where the discretion that is fundamental to our criminal process is involved, we decline to assume that what is unexplained is invidious. In light of the safeguards designed to minimize racial bias in the process, the fundamental value of jury trial in our criminal justice system, and the benefits that discretion provides to criminal defendants, we hold that the Baldus study does not demonstrate a constitutionally significant risk of racial bias affecting the Georgia capital sentencing process.\textsuperscript{152}

Thus, in both the Eighth and Fourteenth Amendment contexts, \textit{McCleskey} required individualized evidence of discriminatory purpose. Statistics showing that the capital punishment system had a disparate racial impact were insufficient to

\textsuperscript{148} Id.

\textsuperscript{149} Id. at 292 (“Our analysis begins with the basic principle that a defendant who alleges an equal protection violation has the burden of proving ‘the existence of purposeful discrimination.’ A corollary to this principle is that a criminal defendant must prove that the purposeful discrimination ‘had a discriminatory effect’ on him. Thus, to prevail under the Equal Protection Clause, McCleskey must prove that the decisionmakers in his case acted with discriminatory purpose.” (footnote omitted)). The McCleskey Court held that the statistical evidence presented on systemic inequality was insufficient to establish discriminatory purpose in his individual case, \textit{id.} at 297, and that the statistical information presented was also inadequate to prove that the legislature enacted or maintained the capital punishment system as a whole in order to further a racially discriminatory purpose, \textit{id.} at 298.

\textsuperscript{150} Id. at 308.

\textsuperscript{151} Id. at 312-13 (“At most, the Baldus study indicates a discrepancy that appears to correlate with race. Apparent disparities in sentencing are an inevitable part of our criminal justice system. The discrepancy indicated by the Baldus study is ‘a far cry from the major systemic defects identified in \textit{Furman}’ . . . .”) (footnotes omitted) (internal quotation marks omitted).

\textsuperscript{152} Id. at 313.
establish a constitutional violation, as long as other procedural protections over the criminal process remained in place.

Outside McCleskey, efforts to obtain a constitutional remedy against the disparate impact of the criminal justice system have met with similar obstacles. At every stage of the criminal process, there is evidence of disparate treatment of minorities—from policing practices to prosecutorial charging decisions to sentencing determinations. Yet, presented with evidence of unequal impact, the courts have time and again denied relief due to lack of individualized evidence of overt discriminatory purpose. For instance, courts rarely sustain selective prosecution and selective enforcement claims—again because defendants are unable to prove discriminatory intent. This evidentiary burden is particularly daunting given the inherent discretion afforded prosecutors and law enforcement officials and the additional...

---


156 Wayte v. United States, 470 U.S. 598, 608 (1985) (“It is appropriate to judge selective prosecution claims according to ordinary equal protection standards. Under our prior cases, these standards require petitioner to show both that the passive enforcement system had a discriminatory effect and that it was motivated by a discriminatory purpose.” (footnotes omitted) (citations omitted)).

157 See, e.g., United States v. Armstrong, 517 U.S. 456, 464-65 (1996) (“A selective-prosecution claim asks a court to exercise judicial power over a ‘special province’ of the Executive . . . . The Attorney General and United States Attorneys retain ‘broad discretion’ to enforce the Nation’s criminal laws . . . . They have this latitude because they are designated by statute as the President’s delegates to help him discharge his constitutional responsibility to ‘take Care that the Laws be faithfully executed.’ . . . As a result, ‘[t]he presumption of regularity supports’ their prosecutorial decisions and, ‘in the absence of clear evidence to the contrary, courts presume that they have properly discharged their official duties.’ . . . In the ordinary case, ‘so long as the prosecutor has probable cause to believe that the accused committed an offense defined by statute, the decision whether or not to prosecute, and what charge to file or bring before a grand jury, generally rests entirely in his discretion.’ . . . Of course, a prosecutor’s discretion is ‘subject to constitutional constraints.’ . . . One of these constraints, imposed by the equal protection component of the Due Process Clause of the Fifth Amendment, . . . is that the decision whether to prosecute may not be based on ‘an unjustifiable standard such as race, religion, or other arbitrary classification . . . . A defendant may demonstrate that the administration of a criminal law is ‘directed so exclusively against a particular class of persons . . . with a mind so
obstacles the Supreme Court has erected before defendants can even obtain discovery on selective prosecution practices.\(^{158}\) Likewise, requiring a discriminatory purpose, courts have denied relief in dramatic instances of racial inequality in punishment, including the 100:1 disparity under the Federal Sentencing Guidelines in punishment of crack and powder cocaine—drugs distinguishable primarily because of their differential usage by blacks and whites, respectively.\(^{159}\) Instead, the Supreme Court has merely held that judges *in their discretion* may take into account the disparate punishment for crack and cocaine in deciding whether to depart from the Federal Sentencing Guidelines, made advisory by *Booker.*\(^{160}\)

Because the Court has required cross-contextual uniformity in its equal protection doctrine, the traditional limiting principles on the equal protection clause apply in full to claims made in the criminal justice context.\(^{161}\) The universal discriminatory purpose requirement recognizes that much of government action, no matter how benign, will disparately affect individuals of different races, and if *all* legislation were subject to strict scrutiny merely on account of disparate impact,

\(^{158}\) Merely to obtain discovery in a case alleging selective prosecution, the defendant must make a "credible" threshold showing that the "government declined to prosecute similarly situated suspects of other races." *Id.* at 458. The Court explained that "[t]he justifications for a rigorous standard for the elements of a selective-prosecution claim . . . require a correspondingly rigorous standard for discovery in aid of such a claim." *Id.* at 468; *see also* *Davis,* supra note 83, at 18.

\(^{159}\) David A. Sklansky, *Cocaine, Race, and Equal Protection*, 47 STAN. L. REV. 1283, 1303 (1995) ("Black defendants have mounted equal protection challenges to the federal crack sentences in each of the regional federal courts of appeals. The precise forms of the challenges have varied. Some defendants have argued that Congress acted unconstitutionally in 1986, some have attacked the Sentencing Commission's extension of the 100:1 ratio adopted by Congress, and some have challenged Congress' and the Commission's failure to amend the ratio when presented with evidence of its overwhelmingly disproportionate impact on black defendants. The results, however, have been remarkably consistent: the defendants always have lost, and the opinions generally have been both unanimous and short."). (citing numerous cases).

\(^{160}\) Kimbrough v. United States, 552 U.S. 85, 91 (2007) ("A district judge must include the Guidelines range in the array of factors warranting consideration. The judge may determine, however, that, in the particular case, a within-Guidelines sentence is 'greater than necessary' to serve the objectives of sentencing. 18 U.S.C. § 3553(a) (2000 ed. and Supp. V). In making that determination, the judge may consider the disparity between the Guidelines' treatment of crack and powder cocaine offenses.").

\(^{161}\) *See* Sklansky, supra note 159, at 1284 ("For at least the past two decades the Supreme Court, along with many of its critics, has tended to assume that equal protection doctrine should remain relatively uniform regardless of factual context: the test for unconstitutional inequality in criminal sentencing, for example, should be the same as in civil service promotions.").
the legislature would be unable to legislate. As the Court explained in \textit{Washington v. Davis},

A rule that a statute designed to serve neutral ends is nevertheless invalid, absent compelling justification, if in practice it benefits or burdens one race more than another would be far-reaching and would raise serious questions about, and perhaps invalidate, a whole range of tax, welfare, public service, regulatory, and licensing statutes that may be more burdensome to the poor and to the average black than to the more affluent white.\footnote{Washington v. Davis, 426 U.S. 229, 248 (1976).}

Numerous scholars have explored the conceptual flaws with requiring discriminatory purpose to achieve heightened scrutiny under the equal protection clause.\footnote{See, e.g., Sheri Lynn Johnson, \textit{Unconscious Racism and the Criminal Law}, 73 CORNELL L. REV. 1016 (1988); Eric Schnapper, \textit{Two Categories of Discriminatory Intent}, 17 HARV. C.R.-C.L. L. REV. 31 (1982); Sklansky, supra note 159.} The intent requirement fails to adequately factor in the prevalence of unconscious racism in our society; it fails to respond to the savvy of modern-day racists who can easily avoid appearances of overt racial bias; and it fails to protect minorities who, at the end of the day, are trapped in an unequal system regardless of the intent of the state actors. It likewise fails to account for the historical shift from overt and legalized racist systems which imposed illegitimate punishments on African Americans (through slavery and Jim Crow), to covert racism that disproportionately imposes ostensibly legitimate punishments on minorities, aided by unconscious biases and discretionary decision-making by layers of state actors.\footnote{See, e.g., McCleskey v. Kemp, 481 U.S. 279, 332-33 (1987) (Brennan, J., dissenting).} Yet, in light of the breadth and uniformity of equal protection doctrine, any doctrinal change to account for these critiques and, specifically, to account for the inequity in the criminal justice system, would have far-reaching, even nuclear, implications—implications that the Court is unwilling to accept.\footnote{Of course, many have argued that Equal Protection doctrine should change—for example, to recognize that subconscious racism of legislatures and other state actors produces disparate racial impact far more, in modern times, than overtly discriminatory purpose. See, e.g., Charles R. Lawrence III, \textit{The Id, the Ego, and Equal Protection: Reckoning with Unconscious Racism}, 39 STAN. L. REV. 317 (1987); Richard Salgado, \textit{Dan the Xenophobe Rides the A-Train, or the Modern, Unconscious Racist in \textquotedblleft Enlightened America\textquotedblright}, 15 AM. U. J. GENDER SOC. POLICY & L. 69 (2006).}

Even if the Court were to understand the demands of equal protection differently in the criminal context,\footnote{For an extended discussion of how and why the Court should reject a uniform or universalist approach to the equal protection clause, see Sklansky, \textit{supra} note 159, at 1312-22.} the
pragmatic anxiety would remain that a hard look into racial inequality would unravel the entire criminal justice system as we know it. Although there may be disagreement about the cause of the inequality, it is indisputable that, in practice, the criminal justice system does operate disproportionately against African Americans and other minorities. Given that reality, the Court fears that acknowledging the racial inequity in constitutional terms could crumble the very foundation of the criminal justice system in America. In *McCleskey*, the Court stated with surprising honesty:

> McCleskey’s claim, taken to its logical conclusion, throws into serious question the principles that underlie our entire criminal justice system. The Eighth Amendment is not limited in application to capital punishment, but applies to all penalties . . . . Thus, if we accepted McCleskey’s claim that racial bias has impermissibly tainted the capital sentencing decision, we could soon be faced with similar claims as to other types of penalty. [n38]

n38 Studies already exist that allegedly demonstrate a racial disparity in the length of prison sentences.  

Justice Brennan, in dissent, criticized the Court for its apparent “fear of too much justice.”  

The Court recognizes that the criminal justice system bears a disparate racial impact—or, at least, that defendants might legitimately and predictably make such a claim. Indeed, it is because the imbalance is so marked, and so widespread, that a pragmatic Court finds itself unable to make any kind of pronouncement that the disparity is unconstitutional. The Court prefers not to remedy any of the inequality rather than risk invalidating the entire system. Though ostensibly limited to the death penalty context, *McCleskey* presented the Court with another nuclear option. The Court saw no limiting principle within the equal protection clause or the “arbitrary

---

168 *Id.* at 339 (Brennan, J., dissenting); *see also id.* at 365 (Blackmun, J., dissenting) (“One of the final concerns discussed by the Court may be the most disturbing aspect of its opinion. Granting relief to McCleskey in this case, it is said, could lead to further constitutional challenges. *Ante,* at 314-19. That, of course, is no reason to deny McCleskey his rights under the Equal Protection Clause. If a grant of relief to him were to lead to a closer examination of the effects of racial considerations throughout the criminal justice system, the system, and hence society, might benefit. Where no such factors come into play, the integrity of the system is enhanced. Where such considerations are shown to be significant, efforts can be made to eradicate their impermissible influence and to ensure an evenhanded application of criminal sanctions.”).
and capricious” doctrine that would cabin its reach and prevent the paralysis of the entire criminal justice system.\footnote{Arguably, a ruling in favor of McCleskey could have legitimately been limited to the death penalty context, under the oft-cited principle that death is different. \textit{See}, e.g., Woodson v. N. Carolina, 428 U.S. 280, 305 (1976).}

Thus the Court has resisted placing a robust counter-majoritarian check on draconian prison sentences, has more broadly interpreted the constitutional protection against excessive punishments in majoritarian terms, and has seen no constitutional infirmity when punishments are imposed differentially upon minorities than on the majority. These interpretations fail to adequately account for the original counter-majoritarian, anti-discrimination, and racially egalitarian principles enshrined in the Constitution, and fail to give those principles meaning in the modern world. The real-world ramifications have been stark. The Court has stood by as the majority has set in motion the modern American punishment machine and, predictably, placed the weightiest burdens of that punitive system disproportionately upon minorities.

IV. \textbf{CONSTITUTIONAL CONSTRUCTION: THE PATH FORWARD}

The Court has advanced an impoverished understanding of the Eighth Amendment that is blind to the original counter-majoritarian imperative and thus is unresponsive to the cruelty and inequality endemic to our age of mass incarceration. I advance a different interpretation—one that keeps faith with the amendment’s original counter-majoritarian and anti-discrimination principles and redeems those principles for our generation. The Eighth Amendment, I argue, is concerned with severe punishments that bear a disparate impact upon minorities.

We need a new Eighth Amendment doctrine that puts this theory into practice. This doctrine must contain, at a minimum, three core features. First, it must include a robust \textit{counter-majoritarian} dimension, not simply a consideration of majoritarian norms. Second, it must be responsive to punishments with a \textit{disparate impact} upon minorities, irrespective of discriminatory purpose. Third, it must be provide a substantive check on \textit{cruelty} in punishments, beyond the death penalty context.
Here I explore one doctrinal approach that achieves each of these three core features. I propose a two-step test for adjudicating Eighth Amendment claims about excessive punishment. First, a judge should determine whether a punishment for a particular crime is “unusual”—i.e., disproportionately meted out against minorities or a suspect class. If a punishment is “unusual,” the court should then assess whether the punishment is cruel, using tiered levels of heightened scrutiny that vary with the degree of unusualness. If the punishment is not “unusual,” but is rather imposed proportionately against the majority and minorities alike, traditional Eighth Amendment inquiry into the punishment’s constitutionality would apply—including a consideration of evolving standards of decency, arbitrariness, and the narrow gross disproportionality inquiry.

This approach offers important improvements over the Court’s existing set of doctrines. It takes seriously the counter-majoritarian imperative of the Eighth Amendment, and promises a meaningful engagement with injustices that existing case law fails to see.

A. Identifying “Unusual” Punishments

The first step in this test for “cruel and unusual punishments” is to determine whether a punishment is “unusual”: whether it has a disparate impact upon a minority that receives insufficient protection through the political process—or, in other words, upon an Eighth Amendment “suspect class.” This presents two more questions: First, what is a “suspect class” for the purposes of the Eighth Amendment? Second, what real-world measure of inequality is sufficient to establish disparate impact?

1. Eighth Amendment “Suspect Classes”

To understand whether a punishment bears an “unusual” impact upon a particular minority group, we must, of course, define which groups count as “minorities” or “suspect classes” for the purposes of the Eighth Amendment. As a starting point, judges can borrow from the approach of equal

---

170 It is not the only workable model; one can imagine other doctrinal innovations that might achieve similar ends, and the development of the ideal solution should be informed by practical application and real-world feedback.
protection law, where case law about suspect and quasi-suspect classifications is well developed.

The clearest examples of suspect classes are “discrete and insular minorities”\textsuperscript{171} that have faced historical discrimination or unequal treatment, and who that been significantly underrepresented in the political process. The Supreme Court recognizes a suspect classification when a group is “saddled with such disabilities, or subjected to such a history of purposeful unequal treatment, or relegated to such a position of political powerlessness as to command extraordinary protection from the majoritarian political process.”\textsuperscript{172} Racial minorities such as African Americans and Latinos would clearly meet that definition in the Eighth Amendment context, just as in the equal protection context.

Equal protection doctrine, however, should be a starting point and not the end point of the Eighth Amendment analysis, which must be tailored to the specific context of criminal punishment. “Suspect classes” for the purposes of Eighth Amendment analysis should specifically reflect a historical or heightened risk of disadvantage in the criminal justice system as well as a political marginalization in society at large. For this reason, equal protection suspect classes will be both under- and over-inclusive. For example, although gender-based classifications receive heightened scrutiny under equal protection law, the relative absence of a history of discrimination against women in the criminal justice system may counsel against gender imbalance being “unusual” for the purposes of the Eighth Amendment.\textsuperscript{173} By contrast, laws that disproportionately affect individuals on the basis of sexual orientation may well be “unusual,” given the nation’s history of criminalizing sodomy, transgender conduct, and homosexual status, even though the Supreme Court has yet to recognize sexual orientation as a suspect or quasi-suspect classification in the equal protection context.\textsuperscript{174} Other groups that have faced discrimination in the criminal justice system and underrepresentation in the political process—including poor people, disenfranchised felons, and illegal

\textsuperscript{171} United States v. Carolene Products Co., 304 U.S. 144, 153 n.4 (1938).
\textsuperscript{173} Some may reasonably disagree with this conclusion, and women could certainly litigate the historical question in the courts. A case could be made that women who have come into contact with the criminal justice system have been punished excessively relative to male counterparts; for example, prostitution may be punished more harshly than the purchasing of sex.
\textsuperscript{174} See, e.g., Windsor v. United States, 699 F.3d 169, 182 (2d Cir. 2012), aff’d on other grounds, 133 S. Ct. 267 (2013).
immigrants—present harder cases. Whether disparate punishment of these groups would be “unusual” is a complex question that goes beyond the scope of this article.\footnote{175}

2. Disparate Impact

Next, to determine whether punishment under a particular statute has a “disparate impact” upon members of an identifiable suspect class, a court should consider statistical evidence about the rate of punishment of different demographic groups for particular crimes.\footnote{176} The greater the statistical disproportion in conviction rates, the more “unusual” the punishment.\footnote{177} Evidence of discriminatory intent in the

\footnote{175} The Supreme Court has refused to consider classifications based on relative wealth to be “suspect” in the Equal Protection context. See Maher v. Roe, 432 U.S. 464, 471 (1977). However, the long and entrenched history of discrimination in the criminal justice system against poor people—as well as unique problems of providing adequate legal representation to the poor, notwithstanding the ostensible protections of Gideon v. Wainwright, 372 U.S. 335 (1963)—may counsel in favor of a different rule in the Eighth Amendment context. Not only are indigent people underrepresented in the political system, but they are also, as a whole, underrepresented in the legal system. See, e.g., American Bar Association Standing Committee on Legal Aid and Indigent Defendants, Gideon’s Broken Promise: America’s Continuing Quest for Equal Justice: A Report on the American Bar Association’s Hearings on the Right to Counsel in Criminal Proceedings iv (2004); Mary Sue Backus & Paul Marcus, The Right to Counsel in Criminal Cases, A National Crisis, 57 Hastings L.J. 1031, 1034 (2006). The recognition of suspect classification for indigent persons in the punishment context, even if not in the general legislative context, may be sensible and, in fact, would cohere with the Court’s existing recognition that there are special concerns raised when the poor receive disparate treatment in the criminal justice context. See Maher, 432 U.S. at 471 n.6.

Similarly, because disenfranchised felons and undocumented immigrants are barred from voting, they are perhaps the quintessential minorities at risk of majoritarian overreaching. There is a special concern that laws directly targeting recidivists and illegal immigrants may be excessive, even if it is perfectly legitimate to criminalize the underlying conduct, because the majoritarian legislators and people like them will experience only the benefits, and not the costs, of a harsh punishment regime. However, there are obvious differences between recidivism or undocumented status and the immutable traits such as race and gender that have given rise to traditional suspect and quasi-suspect classes.

In addition to considering the proportion of minorities convicted of a particular crime, it is arguable that judges considering the “unusualness” of a punishment should also take evidence on disparities in the relative harshness of sentences actually imposed upon members of the majority and minority groups when judicial sentencing discretion is available. However, because disparities in judicial sentencing practices bear a far more complex relationship to majoritarian excesses than do legislative and executive actions, a full doctrinal analysis of how judicial sentencing disparities should be considered under the counter-majoritarian Eighth Amendment goes beyond the scope of this article. This article focuses, instead, on legislative determinations of mandatory minimum sentences and executive patterns of law enforcement.

\footnote{177} The relevant population pool would be the jurisdiction setting the punishment. Thus, the “unusualness” of a statewide criminal statute should be evaluated according to the impact of that statute statewide—because the relevant
legislative history of a criminal statute or in the manner of its enforcement would be admissible and would contribute to the Eighth Amendment analysis by justifying a stronger assessment of the degree of “unusualness,” but evidence of such a purpose would not be required. The defendant would only need to prove a differential effect of a criminal statute upon a suspect class to satisfy the “unusual” prong.\footnote{A claim that a punishment is “unusual” should be available to all, not merely to members of the suspect class. The key to the counter-majoritarian concern here is that the incentives of the political process are skewed when a particular punishment disproportionately (but not exclusively) impacts minorities. Members of the majority who are subject to criminal sanctions approved under the skewed majoritarian process should likewise be able to remedy that procedural defect. Analogously, the courts have permitted \textit{Batson} challenges to the racially discriminatory use of peremptory strikes against potential jurors, irrespective of the race of the defendant. \textit{Holland v. Illinois}, 493 U.S. 474, 476 (1990).} \footnote{Much evidence suggests this is not merely an assumption but in fact reality. \textit{See, e.g., Substance Abuse & Mental Health Servs. Admin., U.S. Dept of Health & Human Servs., Results from the 2007 National Survey on Drug Use and Health: Detailed Tables tbl.1.24B (2012), available at http://www.samhsa.gov/data/NSDUH/}

Let us take an example. A state statute criminalizes theft under $500. Statewide, 75% of the people convicted of violating this statute are African American, although only 35% of the state population is African American. The criminal statute has a disparate impact upon a suspect class, and the “unusual” prong is satisfied.

Note that it is not clear from these statistics whether the disparate impact is a result of differential rates of crime commission or disparate patterns of law enforcement. In other words, it is unclear whether, per capita, more African Americans than whites in the state commit the crime of petty theft; or whether police officers and prosecutors simply enforce the law more strictly against African Americans. Any number of complicated reasons ranging from overt discrimination to unconscious bias at multiple points of discretionary decision-making could lead to a disparate racial impact. Under a counter-majoritarian theory of rights, however, the reason for the differential impact on minorities should not control the determination of whether the punishment is “unusual.”

Most would agree that there is something suspicious—something warranting a second look—when criminal laws are disproportionately enforced against minorities, despite relatively equal offense rates across the population. Thus, for example, let us assume\footnote{A claim that a punishment is “unusual” should be available to all, not merely to members of the suspect class. The key to the counter-majoritarian concern here is that the incentives of the political process are skewed when a particular punishment disproportionately (but not exclusively) impacts minorities. Members of the majority who are subject to criminal sanctions approved under the skewed majoritarian process should likewise be able to remedy that procedural defect. Analogously, the courts have permitted \textit{Batson} challenges to the racially discriminatory use of peremptory strikes against potential jurors, irrespective of the race of the defendant. \textit{Holland v. Illinois}, 493 U.S. 474, 476 (1990).} that whites and African Americans use
marijuana at similar rates, but that the laws prohibiting possession are disproportionately enforced against blacks. The end result is a defect in majoritarian fairness because the majority is not internalizing the costs of setting harsh sentences. Although the college-aged sons and daughters of the legislators are theoretically at risk of getting caught and punished for marijuana possession, that possibility is not internalized when the legislators set the sentencing range because the risk is not a realistic one. Legislators set penalties knowing that law enforcement has vast discretion in executing the law and that this discretion will generally be used to externalize the costs of punishment upon “the other.”

A similar, though less recognized, breakdown in majoritarian political protections against excessive punishment arises when criminal laws target conduct that is, ordinarily, committed more frequently by minorities. For instance, let us assume that, statistically, more blacks than whites use crack-cocaine, and thus disparate punishment rates for possession of crack are a result not of bias by law enforcement officials or prosecutors but of real-world behavioral differences. Similarly, let us assume that a disproportionate number of those convicted of the crime of illegal reentry in Texas are Latino, and that disproportionate numbers of Latinos relative to whites actually commit that crime in reality. At first glance, the disparate impact of these laws upon minorities may not be surprising or troubling, because it accurately reflects objective differences in criminal conduct. But even in these scenarios the danger remains that the majority will set cruel or excessive punishments, because the majority is able to externalize the burdens of punishment and internalize all the benefits of that criminal law. To attain even nominal advantages for society at large, the legislature may set harsh penalties for particular types of conduct that it would be unwilling to accept as proportionate if those punishments affected them. This may help explain why the crack/cocaine sentencing disparity arose.

180 See, e.g., ELY, supra note 34, at 173 (“A severe (or 'cruel') punishment to which any of us who transgresses is realistically subject is one thing; assuming an impartial enforcement regime, the political processes can be counted on to block beheading as the penalty for tax fraud. If, however, there are buffers, if the system is constructed so that “people like us” run no realistic risk of such punishment, some nonpolitical check on excessive severity is needed.”).
Perhaps in setting the punishment for crack possession, the majority was able to externalize all the burdens of punishment, but in setting the punishment for cocaine possession—a crime that was committed frequently by whites—the majority internalized the costs of that punishment and set a more moderate sentence when it affected “their own.”

Thus even when—and, perhaps, especially when—disparate impact is caused by disparate behavior, there is a necessary role for a court to play in scrutinizing the severity of punishments more carefully, given that the political process will not offer robust protections. When members of the majority (and their constituents) face little realistic possibility of being convicted under a criminal statute, the political process protections against excessive punishment fall apart.\footnote{Moreover, requiring proof of whether disparate impact is caused by disparate enforcement or disparate conduct would place an undue and frequently impossible burden upon the defendant to explain the causal significance of unequal effects.}

B. Identifying “Cruel” Punishments

The threshold determination of whether a particular punishment is “unusual”—imposed disproportionately upon minorities—informs the subsequent judicial scrutiny into the excessiveness or “cruelty” of that punishment.

When punishments for certain crimes are not “unusual”—when they are not imposed with disproportionate frequency or severity on suspect classes—traditional Eighth Amendment doctrine should apply. Here, courts should consider the cruelty of prison sentences with the traditional deference to majoritarian legislative judgments about appropriate punishments for crimes, taking into account all the legitimate penological purposes of retribution, deterrence, incapacitation, and rehabilitation.\footnote{See, e.g.,\value{footnote}.} The narrow “gross disproportionality” review of prison sentences that the Supreme Court has permitted to date would remain virtually unchanged, and in the death penalty and juvenile life without parole contexts, the Court would continue its more robust and independent inquiry into disproportionate punishment, referencing evolving standards of decency.

The more irregular or “unusual” the application of a particular punishment, however, the stricter the judicial scrutiny and the more compelling the governmental interests must be to justify a harsh punishment—including a lengthy
prison sentence. If a punishment has a disparate impact upon suspect classes of people, the courts should no longer pay heavy deference to the majoritarian legislative determinations of appropriate punishment. Rather, the burden should shift to the government to show that the punishment is not “cruel” or “excessive” under the circumstances. An extreme disparity in the punishment rates of a particular crime would lead to the strictest Eighth Amendment scrutiny, with the most rigorous testing of how compelling the government’s asserted interests are, how narrowly tailored the punishment scheme is to address the asserted governmental interest, and how reprehensible the conduct being punished is compared to the harshness of the penalty. A lesser, but still noticeable, disparity in application would lead to an intermediate level of scrutiny.

Under this two-pronged test, “unusual” application of a law against a minority group would not, in itself, be fatal to a particular punishment. A short prison sentence for a minor offense might not be “cruel” no matter how “unusual.” Conversely, the harsh sentence of life without parole or even the death penalty for a serious offense such as murder might not be “cruel,” even if disparately applied, because the court might conclude that it is justified by compelling government interests. However, heightened scrutiny would likely lead to the invalidation of other draconian prison sentences imposed for less serious offenses that bear a disparate impact on minorities—including severe penalties for simple or repeat drug possession, broadly defined three-strikes laws, and harsh mandatory minimums. In other words, this approach would pertain less to harsh punishments for murder, rape, and armed robbery, and much more to harsh punishments for nonviolent and victimless crimes.

In short, evidence of “unusual” effect would trigger a heightened judicial scrutiny into possible cruelty without heavy deference to majoritarian judgments about the appropriateness of the penalty—a marked divergence from the Court’s current approach under Rummel and its progeny. In today’s age of mass incarceration, this approach would have the critical effect of providing meaningful judicial review for draconian prison

183 Although, of course, the fact that a prison sentence is short is not conclusive that it is not cruel. As the Supreme Court has long recognized, “To be sure, imprisonment for ninety days is not, in the abstract, a punishment which is either cruel or unusual. But the question cannot be considered in the abstract. Even one day in prison would be a cruel and unusual punishment for the ‘crime’ of having a common cold.” Robinson v. California, 370 U.S. 660, 667 (1962).

sentences. Under rational basis scrutiny—which is essentially what the courts apply now—severe prison sentences will almost always survive, no matter how unequal their impact. With rare exceptions, there will always be some rational governmental interest in imposing a particular punishment. Even the paradigmatic example of life in prison for a parking violation does serve a legitimate government interest of deterring objectionable conduct. Heightened judicial scrutiny is needed to give meaning to the counter-majoritarian imperative of the Eighth Amendment, and will be of special consequence in today’s punitive criminal justice system.

When applying heightened scrutiny, courts should consider the type of penological interest alleged by the government alongside the severity of the sentence when a disparate impact can be established. The problem with majoritarian rule, and the reason why a counter-majoritarian check is necessary, is that the interests of the majority inevitably take precedence over the interests of the minority. An unchecked majority will seek to internalize benefits while externalizing costs. The internalized benefits may be entirely legitimate policy interests—such as minimizing crime, protecting property, or fighting drug addiction. But when the costs are externalized to a population other than the majority, the legislative cost-benefit analysis is skewed. The majority is able to place excessive burdens on a small population in order to obtain relatively minor benefits for itself. Were those burdens spread across the population equally, society may deem them unacceptable payment for the benefits sought.

The court should consider the majoritarian tendency toward cost-externalization when scrutinizing “unusual” punishments. The severity of the punishment must be such that the majority would accept it if the burdens were spread evenly across the population. Moreover, the court should place differing weight on the different types of legitimate penological interests according to whether the burdens and benefits associated with each asserted interest are tailored to the individual circumstances of the case or diffusely distributed across

---

185 See, e.g., Ewing v. California, 538 U.S. 11, 27-28 (2003) (plurality opinion) (“We do not sit as a ‘superlegislature’ to second-guess [the state’s] policy choices. It is enough that the State of California has a reasonable basis for believing that dramatically enhanced sentences for habitual felons ‘advance[s] the goals of [its] criminal justice system in any substantial way.’” (citations omitted)).

186 The challenges and importance of understanding what burdens the majority might accept to achieve a particular good finds some resonance in the Rawlsian concept of the veil of ignorance. See JOHN RAWLS, A THEORY OF JUSTICE 118 (1971).
society. Interests that are focused on the offender’s direct culpability—retribution and rehabilitation—should be primary in the determination of “cruelty.” Interests that are directed toward benefits for society as a whole—incapacitation and deterrence—should be accorded less weight, because there is a heightened risk that the majority will over-value any benefits it receives while under-valuing the associated costs imposed upon the minority.

Interestingly, in the wake of the budgetary crises and fiscal austerity measures of recent years, the majority has for the first time begun to internalize more of the costs of mass incarceration and the war on drugs. Although these financial costs are relatively minor in comparison with the extraordinary burdens that the war on drugs has placed upon minority communities, we are seeing the first real initiatives against mass incarceration in the context of the war on drugs. If the majority is unwilling to bear the financial cost of adequately funding prisons to secure the governmental interests asserted in favor of tough-on-crime legislation, we can safely assume that the majority would not be willing to endure the dramatic and severe criminal and community consequences that minorities have experienced if they were imposed equally against the majority.

C. Implementing the Doctrinal Framework

Imagine a 45-year-old African American man, Timothy Johnson, in my present home of New Orleans. Over the past 20 years, Mr. Johnson has had three prior felony convictions: one conviction for marijuana possession, and two for possession of

187 These burdens include, of course, the costs of disproportionate rates of incarceration. But additional burdens are imposed on the entire community—ranging from absent parents and spouses to eviction from public housing to reduced privacy in the home and neighborhood. See generally John A. Powell & Eileen B. Hershenov, Hostage to the Drug War: The National Purse, the Constitution and the Black Community, 24 U.C. Davis L. Rev. 557, 599-614 (1991).


coclaine.\textsuperscript{190} He has just been convicted again of possession of cocaine. Based on his history of recidivism, he was charged under the multiple bill statute as a “quadruple bill.”\textsuperscript{191} Although the sentencing range for simple possession of cocaine is up to five years in prison,\textsuperscript{192} when combined with the multiple bill statute,\textsuperscript{193} he is facing 20 years to life.

At his sentencing hearing, Mr. Johnson raises an Eighth Amendment challenge to the mandatory sentencing range. He presents statistical evidence about the rates of punishment under both the cocaine statute and the multiple bill statute, and makes the threshold showing of “unusual” punishment: that African Americans are disproportionately charged and convicted for violating these statutes.

This demonstration of disparate impact triggers a heightened judicial scrutiny into cruelty: into the proportionality of the punishment to the offense, without the traditional deference to the legislative determinations of appropriate sentencing practices. Mr. Johnson presents mitigating evidence about his own history and argues that a sentence within the given range would be inappropriate in his own case; he also argues more universally that 20 to life is an excessive punishment for cocaine possession, even when combined with a history of recidivism. Although the judge may consider societal interests in deterrence when assessing the proportionality of the punishment, she must place greater weight on individualized factors, including the culpability of the particular offender and the severity of the offense, taking into consideration whether the offense was nonviolent or victimless. The judge has the obligation, under the Eighth Amendment, to sentence Mr. Johnson to a lower term of years than the mandatory minimum, if she deems the sentencing range to be cruel or excessive without deferring to legislative determinations.

If Mr. Johnson loses his constitutional challenge in the trial court, he will be able to appeal the sentence actually imposed by the judge to the higher courts on Eighth Amendment grounds, and if unsuccessful in state court, may subsequently file a petition for habeas relief in a federal court.

\textsuperscript{190} Id. § 40:967(C)(2) (2012).
\textsuperscript{191} Under Louisiana’s Habitual Offender Law, criminal defendants may be charged as second-, third-, or fourth-time offenders (colloquially known as double, triple, or quadruple bills) and be subject to significantly harsher sentences. See Id. § 15:529.1.
\textsuperscript{192} Id. § 40:967(C)(2).
\textsuperscript{193} Id. § 15:529.1(A)(4)(a).
D. Cabining Judicial Anxiety

As a practical matter, the doctrinal approach proposed here has advantages over that rejected in *McCleskey*, in which existing equal protection and Eighth Amendment doctrines were leveraged to remedy systemic inequality in the criminal justice system. Specifically, this approach has the benefit of reducing judicial anxiety and, with it, judicial paralysis. If equal protection doctrine and Eighth Amendment doctrine are circles in a Venn diagram, a counter-majoritarian Eighth Amendment lies at the overlapping intersection of those circles—a narrower subset of the two. Equal protection doctrine covers all government action, not just criminal punishments; traditional Eighth Amendment doctrine covers all punishments, not just those disproportionately imposed against minorities.

My approach requires heightened judicial scrutiny into cruelty/proportionality where disparate impact (but not necessarily discriminatory intent) can be shown. This cabined set of circumstances neither threatens to dismantle *Washington v. Davis* (which requires discriminatory intent in addition to discriminatory effect for heightened scrutiny under the equal protection clause) nor forces a court to engage in unbounded review of “gross disproportionality” (rejected in *Harmelin*). The narrower context of criminal punishment—and, even more specifically, cruel punishment—makes it feasible to tackle disparate effect without threatening to invalidate all governmental action, no matter how benign, that
affects individuals of different races differently. While in the equal protection context, the courts have refused to employ strict scrutiny on a showing of disparate impact alone, the combination of severity or harshness and disparate impact has a focused and contained constitutional significance.

McCleskey and Harmelin demonstrate acute judicial anxiety over the radical consequences of unbounded disparate impact or proportionality review. By recognizing that the Eighth Amendment places a distinctly counter-majoritarian limitation on excessive sentences, we can take these nuclear options off the table and use judicial scrutiny to target a still massive, but more manageable, problem: the imposition of punitive criminal sanctions disproportionately against minorities. The counter-majoritarian Eighth Amendment does not prohibit all disparate impact (or “unusual” impact) in the criminal justice system. But where there is disparate impact, and the punishments are particularly cruel or harsh, then the Eighth Amendment raises a red flag that majoritarian protections have been insufficient and may in fact have contributed to improper cost-externalization upon a minority group.194

CONCLUSION

My reading of the Eighth Amendment as prohibiting cruel (particularly severe) and unusual (disparately imposed) punishments redeems its original counter-majoritarian and anti-discrimination constitutional principles for the modern age. By mandating heightened judicial scrutiny into “cruelty” when “unusualness” can be shown, the doctrinal framework I propose would empower the judiciary as a robust counter-majoritarian check against excessiveness when the effects of the criminal justice system fall disproportionately on minorities—when the political process protections against cruelty fall short. But it would do so without dismantling decades of equal protection jurisprudence or requiring the court to second-guess all legislative determinations about appropriate punishment. This doctrinal model would focus the courts’ attention on the most constitutionally suspect cases: those in which the punishment is harsh, the severity of the underlying conduct is moderate, and the disparate impact is apparent.

194 One could always argue that harshness of punishment is indicative of discriminatory intent; however, there are so many other penological reasons for harsh punishment that this argument will rarely if ever succeed.
For years, the courts have stood by as America has waged a punitive war on drugs and crime that has exacted a heavy and disproportionate toll on minorities. The current state of affairs is a paradigm of majoritarian excess. The Constitution can—and must—be read to limit this inequality. My approach would illuminate and begin to remedy pervasive inequities in the American system of punishment that have, to date, been constitutionally invisible.
“Good Fences Make Good Neighbors”

AN ENVIRONMENTAL JUSTICE FRAMEWORK TO PROTECT PROHIBITION BEYOND RESERVATION BORDERS

Sean J. Wright†

INTRODUCTION

“*My apple trees will never get across
And eat the cones under his pines, I tell him.
He only says, ‘Good fences make good neighbors.’”¹

- Robert Frost

In Whiteclay, Nebraska, a desolate town of 10 people, four rickety shacks line the main road. On average, 13,000 cans of beer and bottles of malt liquor are sold per day from these shacks.² The closest sizeable city is two hours north, but just 240 yards across the state line into South Dakota is the expansive Pine Ridge Indian

† J.D., The Ohio State University Moritz College of Law, 2013; M.A. in Public Policy and Management, The Ohio State University John Glenn School of Public Affairs, 2013; B.A. in Political Science and History, Miami University, 2009. I would like to thank Dean Kathy Northern for encouraging me to develop the independent study that led to this piece and Professor Cinnamon Carlarne for her helpful comments and continued support. I must thank my family for always supporting my dreams—the title is in your honor. Finally, a special thanks is due to my wonderful fiancée, Dania Korkor, for reviewing numerous drafts and inspiring me every day.

¹ Robert Frost, *Mending Wall*, in NORTH OF BOSTON 12 (1914).

Reservation.\textsuperscript{3} There, alcohol consumption has been prohibited since 1834.\textsuperscript{4} Nearly all the alcohol sold in Whiteclay “winds up on Pine Ridge or is consumed by its residents, tribal officials say.”\textsuperscript{5} Pine Ridge is home to the Oglala Sioux Tribe and according to 2010 census data is one of the poorest places in the country.\textsuperscript{6} Whiteclay’s singular purpose is to make alcohol available for consumption across the border. This effectively undermines prohibition and negates the decision of the Oglala Sioux to remain dry.

In 2010, the tribal police made 20,000 alcohol-related arrests.\textsuperscript{7} According to the tribal president, 90% of the criminal cases brought in the tribal courts and a similar number of reservation illnesses were caused by alcohol—“the vast majority of which, he said, was brought illegally from Whiteclay.”\textsuperscript{8} Nationally, excessive alcohol consumption is the leading cause of preventable death among American Indians; affecting that population at twice the rate of the national average.\textsuperscript{9} Alcoholism among American Indian populations is well known; thus, the decision to remain dry is based on a legitimate public health concern.

The convoluted history of regulating alcohol in Indian country,\textsuperscript{10} the historical harms alcohol has had on indigenous

\begin{itemize}
  \item[5] Williams, \textit{supra} note 3.
  \item[6] \textit{Id}.
  \item[7] \textit{Id}.
  \item[8] \textit{Id}.
  \item[10] Generally Indian country is defined at 18 U.S.C. § 1151 (2012) as,
    \begin{itemize}
      \item[(a)] all land within the limits of any Indian reservation under the jurisdiction of the United States Government, notwithstanding the issuance of any patent, and, including rights-of-way running through the reservation,
      \item[(b)] all dependent Indian communities within the borders of the United States whether within the original or subsequently acquired territory thereof, and whether within or without the limits of a state, and
      \item[(c)] all Indian allotments, the Indian titles to which have not been extinguished, including rights-of-way running through the same.
    \end{itemize}
tribes, and the scope of tribal authority to control activity within reservation borders are well known.\textsuperscript{11} What is unclear, however, is the extent to which tribes are able to assert their rights beyond the border of the reservation.\textsuperscript{12} Wading into further turbidity, tribal authority to regulate the sale of alcohol in Indian country is not an inherent sovereign power of Indian tribes; rather, Congress has delegated powers of the federal government to “regulate behavior that would otherwise be beyond the reach of their power.”\textsuperscript{13}

To unpack how and when tribal governments can assert authority outside the boundaries of their reservations, this essay suggests framing the regulation of alcohol flowing into Indian country as an environmental justice concern. In turn, utilizing environmental justice tools and strategies can illuminate how tribes should approach limiting the harmful effects the flow of alcohol has upon their communities. This can be framed as an environmental justice issue because the problem that tribes face is that their decision to remain dry is undermined, which produces an inequitable distribution of environmental hazards. Because tribal governments are limited jurisdictionally and are thus unable to directly limit transboundary activity, tribal communities continue to be exposed to a public health risk.\textsuperscript{14}

The purpose of environmental justice is to minimize environmental inequality.\textsuperscript{15} Framed as an environmental justice concern, Indian tribes who decide to remain dry should not be undermined by extraterritorial threats. In the context of Indian

\footnotesize Federal prohibition policy, however, has narrowed to the definition in 18 U.S.C. § 1154(c) which states:

The term ‘Indian country’ as used in this section does not include fee-patented lands in non-Indian communities or rights-of-way through Indian reservations, and this section does not apply to such lands or rights-of-way in the absence of a treaty or statute extending the Indian liquor laws thereto.

\textsuperscript{11} See Peggy Anderson, \textit{Yakama’s Reservationwide Alcohol Ban Upheld}, SEATTLE TIMES (Dec. 16, 2000, 12:00 AM), http://community.seattletimes.nwsource.com/archive/?date=20001216&slug=TT972LsGL.


\textsuperscript{13} Mark T. Baker, Note, \textit{The Hollow Promise of Tribal Promise of Tribal Power to Control the Flow of Alcohol into Indian Country}, 88 VA. L. REV. 685, 686–87 (2002); see United States v. Mazurie, 419 U.S. 544, 557 (1975) (“These same cases, in addition, make clear that when Congress delegated its authority to control the introduction of alcoholic beverages into Indian country, it did so to entities which possess a certain degree of independent authority over matters that affect the internal and social relations of tribal life. Clearly the distribution and use of intoxicants is just such a matter.”).

\textsuperscript{14} See Centers for Disease Control & Prevention, supra note 9, at 938–39.

country, “the concept of environmental justice is not very useful unless it is broader than just the intersection of civil rights and environmental law.”16 Instead, “in Indian country a vision of environmental justice must also include the tribal right of self-government.”17 This promotion of tribal self-government cannot occur without an ability to control and improve their reservation.18 Without such authority, an injustice occurs.19 After 179 years of prohibition, the Oglala Sioux are still fighting to stop the flow of alcohol into their reservation. Their inability to do so effectively is an injustice.

Framed as an environmental justice concern, not only does the continued flow of alcohol into Indian country represent an inequitable distribution of environmental risk, but the siting of alcohol distribution centers near reservations is an example of locally undesirable land uses (LULUs).20 As the environmental justice movement has made clear, “[s]everal major studies have found that hazardous waste sites, solid waste dumps, polluting factories, and other locally undesirable land uses are located in areas that contain, on average, a higher percentage of racial minorities and are poorer than nonhost communities.”21 As of now, the (dry) Pine Ridge Reservation is one of the poorest areas of the country and is exposed to over 4.3

17 Id.
18 See Sarah Krakoff, Tribal Sovereignty and Environmental Justice, in JUSTICE AND NATURAL RESOURCES: CONCEPTS, STRATEGIES, AND APPLICATIONS 161, 163 (Kathryn M. Mutz et al. eds., 2002) (“[E]nvironmental justice for tribes must be consistent with the promotion of tribal self-governance.”).
20 For a general discussion of LULUs see Vicki Been, Locally Undesirable Land Uses in Minority Neighborhoods: Disproportionate Siting or Market Dynamics? 103 YALE L.J. 1383, 1384 (1994). Essentially, the Environmental Justice movement purports that people of color and other minorities are exposed to greater environmental harms than wealthier individuals. Importantly, the siting of “waste dumps, polluting factories, and other locally undesirable land uses (LULUs) have been racist and classist.” Id. Moreover,

Examples of LULUs are many. Some are well-known and a source of frequent objections, such as adult use establishments. Other LULUs may include uses that are widely used and needed and are often well-established in their locations, but whose well-known emissions of odor or noise—such as airports, landfills, or asphalt plants—become problematic as new neighbors move toward these uses.

million cans of beer and malt liquor a year.\textsuperscript{22} Thus, analyzing the application of environmental justice responses to LULUs can elucidate how tribes can and should respond.

The benefits of framing the problem within the environmental justice framework are twofold. First, this essay presents a novel approach to conceptualizing environmental justice. As scholars debate the future of environmental poverty lawyering, it is helpful to take stock of the successes the movement has had in addressing the inequitable siting of environmental harm. As this essay will establish, environmental justice strategies can provide a roadmap to address inequity beyond the traditional areas of concern.\textsuperscript{23} Second, framing alcohol prohibition within the context of environmental justice will inform the Environmental Protection Agency (EPA) as it develops Plan EJ 2014.\textsuperscript{24} As the EPA attempts to integrate environmental justice into agency programs, policies, and activities, this framing will establish alcohol prohibition as a necessary cross-agency focus area.\textsuperscript{25}

This essay makes an important and innovative contribution to the scholarly literature on tribal environmental justice,\textsuperscript{26} the “good neighbor” principle,\textsuperscript{27} and enforcing policy beyond borders. Ultimately, tribal governments must be enabled to ensure their policy decisions are enforceable. This requires enforcement beyond the borders of the reservation. By conceptualizing prohibition under an environmental justice paradigm, tribes have a direct proxy of when and how to enforce their policy decisions outside their border—environmental statutes.\textsuperscript{28} Additionally, a significant bulk of scholarly attention has been paid to when and how tribes can exercise jurisdiction over non-members who venture inside

\begin{thebibliography}{9}
\bibitem{22} Williams, supra note 3.
\bibitem{23} \textit{See infra} Part IV.
\bibitem{25} \textit{Id.}
\bibitem{26} Krakoff, supra note 18.
\bibitem{28} “[E]nvironmental justice advocates have sought to use the National Environmental Policy Act (NEPA), the National Historic Protection Act (NHPA), the Fourteenth Amendment of the U.S. Constitution, the Civil Rights Act, the Clean Water Act (CWA) and the Clean Air Act (CAA) to pursue citizen suits to remedy environmental injustice.” Melissa O’Connor, \textit{A Failure to Protect: After 13 Years Environmental Justice Never Materializes}, 35 S.U. L. REV. 119, 123 (2007) (citations omitted). Also, grassroots advocacy remains a constant reform strategy. \textit{See} Luke W. Cole, \textit{Empowerment as the Key to Environmental Protection: The Need for Environmental Poverty Law}, 19 ECOLOGY L.Q. 619, 637 (1992).
\end{thebibliography}
their borders, and the Supreme Court has restricted this right in significant ways. But the literature has only begun to discuss extraterritorial authority. This essay explores the ability of tribes to reach beyond their borders and limit the flow of pollution under the various environmental statutes—particularly the Clean Air Act (CAA) and Clean Water Act (CWA)—and suggests that environmental law offers a model approach for reenvisioning/remodeling prohibition policy.

Part I of this essay frames prohibition as an environmental justice concern requiring transboundary enforcement, which is frequently observed in international environmental law. Part II reviews the history of federal Indian alcohol policy—in particular how the Pine Ridge Reservation has struggled to maintain prohibition—finding that legislative and judicial outcomes have limited tribal powers delegated to them by Congress. Part III argues that, analytically, tribes have the same authority to reach beyond their borders to protect public health in much the same way that tribes implement environmental laws—through cooperative-federalism. Finally, Part IV proposes three possible approaches based upon environmental justice tools and strategies to effectuate tribal decisions to remain dry.

I. Framing Tribal Prohibition as an Environmental Justice Concern

Environmental justice seeks to address the disparate distribution of environmental harms throughout society. Understood in this context, the environment “include[s] the

ecological, physical, social, political, aesthetic, and economic environments.” Alcohol is a public health threat in Indian country. Framing the continued flow of alcohol into dry Indian country as an environmental justice concern provides a structure to explore various new approaches to this longstanding issue. When facing public health threats from industrialization, pollution, and contamination, several tools and strategies undertaken by the environmental justice movement can shed light on how to respond to targeted siting of environmental harms. However, tribal environmental justice is different from the broader environmental justice movement. Based upon a unique history of subjugation and exclusion, tribal justice requires not only addressing environmental harm but also supporting tribal self-determination. Additionally, some mechanism for tribes to enforce public health concerns beyond the border must be implemented. Looking to international environmental law is illustrative for these purposes.

A. Tribal Environmental Justice

Unfortunately, “whether by conscious design or institutional neglect, communities of color in urban ghettos, in rural ‘poverty pockets,’ or on economically impoverished Native-American reservations face some of the worst environmental devastation.” The Environmental Justice movement has addressed these concerns in a variety of ways. These approaches provide a lodestar toward addressing the harmful effects of alcohol distribution throughout Indian reservations.

Environmental justice is based upon the premise that the underprivileged and people of color bear a disproportionate share of society’s environmental burdens. Over the past two

---

33 Yang, supra note 15, at 19.
34 See Centers for Disease Control & Prevention, supra note 9, at 938–39; see also A. Mercedes Nails et al., American Indian Youth’s Perception of Their Environment and Their Reports of Depressive Symptoms and Alcohol/Marijuana Use, 44 ADOLESCENCE 965, 968 (2009) (noting that American Indian youths begin drinking earlier than non-Indian peers).
37 See MANASTER, supra note 35, at 246–315.
38 ROBERT D. BULLARD, DUMPING IN DIXIE: RACE, CLASS, AND ENVIRONMENTAL QUALITY 1 (3d ed. 2000) (noting that “[a]n abundance of documentation shows blacks, lower-
decades, the movement has sought to address these inequities. The environmental justice movement has established a variety of important findings. First, there are distributional impacts of environmental rules. This means that environmental regulations may have a demonstrable—and harmful—impact on minority communities. Frequently, industrialization takes on a “not in my backyard” (NIMBY) tone. Thus, “because noxious sites are unwanted . . . and because industries tend to take the path of least resistance, communities with little political clout are often targeted for such facilities.” Second, lacking political influence, minority communities are unable to organize as effectively as majority communities and this leads to underrepresentation in governing bodies, which in turn leads to further limited access to policymakers.

Environmental justice claims by Native Americans have typically fallen into two camps: “claims for regulatory control over reservation lands,” and “claims by indigenous peoples that they have unique interests and ought to be represented as ‘rights-holders’ in national or international decision-making that impacts their communities.” These claims are intrinsically linked to the fundamental difference between the environmental justice movement generally and tribal environmental justice—tribal sovereignty. While Indian country has been significantly harmed by environmental hazards, tribal leaders advance the belief that unlike other communities of color, Indian communities primarily suffered from the federal government’s devaluing of tribal sovereignty and resulting paternalistic management policies. As part of advancing tribal self-government to redress

income groups, and working-class persons are subjected to a disproportionately large amount of pollution and other environmental stressors”).

39 Torres, supra note 32, at 840.
41 See id.
42 Tsosie, supra note 19, at 1627.
43 Id. at 1627-28.
44 Id. at 1631-32; see also Judith V. Royster, Native American Law, in THE LAW OF ENVIRONMENTAL JUSTICE: THEORIES AND PROCEDURES TO ADDRESS DISPROPORTIONATE RISKS 200 (Michael B. Gerrard & Sheila R. Foster eds., 2d ed. 2008).
46 Tsosie, supra note 19, at 1632; Royster, supra note 44, at 199 (“Indian tribes connect to their lands not only on economic and emotional levels, but also on the levels of culture, religion, and sovereignty.”).
environmental injustice, Professor Sarah Krakoff believes that environmental justice can be coextensive with recognition of tribal regulatory authority.\textsuperscript{47} The federal government’s acknowledgment of tribes’ treatment as state (TAS) status reinforces self-governance.\textsuperscript{48} In fact, “the active exercise of tribal regulatory authority over the reservation environment is seen as an antidote to the perceived victimization of reservation communities by exploitive and environmentally hazardous industries.”\textsuperscript{49} Thus, resolving any environmental justice dispute—such as the flow of alcohol—through a tribal regulatory framework is the panacea to existing injustice.

Collectively, the environmental justice movement has advanced in fits and starts. The high-point of the Environmental Justice movement came when President Clinton signed Executive Order 12,898. The President required that “each Federal agency . . . make achieving environmental justice part of its mission by identifying and addressing, as appropriate, disproportionately high and adverse human health or environmental effects of its programs, policies, and activities on minority populations, and low-income populations.”\textsuperscript{50} Since that time, the federal government has become more aware of how decision-making processes impact minority communities. In fact, the EPA has an Advisory Council on Environmental Justice. Part of this council is an Indigenous Peoples Subcommittee tasked with ensuring that Native peoples have a role in environmental decision-making. These advances, however, have not translated to successful legal challenges to unjust siting of environmental harms.

Outside of the major environmental statutes, Title VI of the Civil Rights Act of 1964\textsuperscript{51} was long seen as the viable vehicle for litigating environmental justice claims. But in 2001, the Supreme Court limited the likely success of these claims. In \textit{Alexander v. Sandoval}, the Court ruled that there is no private right of action to enforce disparate impact regulations under Title VI.\textsuperscript{52} This ruling

\textsuperscript{47} Krakoff, supra note 18, at 163.

\textsuperscript{48} See infra Part III.

\textsuperscript{49} Tsosie, supra note 19, at 1632.

\textsuperscript{50} See \textit{Exec. Order No. 12,898, 3 C.F.R. § 6-609} (1995); see also Bradford C. Mank, \textit{Executive Order 12,898, in THE LAW OF ENVIRONMENTAL JUSTICE: THEORIES AND PROCEDURES TO ADDRESS DISPROPORTIONATE RISKS, supra note 44, at 142 (“The Order and the accompanying memorandum have already had a major impact on how agencies integrate environmental justice issues into their activities.”).}

\textsuperscript{51} 42 U.S.C. § 2000d (2012) (“No person in the United States shall, on the ground of race, color, or national origin, be excluded from participation in, be denied the benefits of, or be subjected to discrimination under any program or activity receiving Federal financial assistance.”).

\textsuperscript{52} Alexander v. Sandoval, 532 U.S. 275, 293 (2001).
was further narrowed by the Third Circuit in *South Camden Citizens in Action v. New Jersey Department of Environmental Protection*. The Third Circuit found that not only did the plaintiffs lack standing to bring claims of discrimination under Title VI, but that Title VI did not create freestanding rights enforceable under 42 U.S.C. § 1983. To a large extent, the limitations of disparate impact litigation are in stark contrast with successful challenges raised through environmental statutes. A key takeaway from the nearly two decades of environmental justice litigation is that alternative avenues for remedy are necessary.

**B. Preventing Transboundary Harms: The International Environmental Law Framework**

Extending a tribe’s reach beyond its territorial boundaries is akin to the prerogative of nation-state sovereigns within international environmental law. In fact, prohibiting activity beyond one’s border is a cornerstone of this legal regime. The origins of this principle stem from an arbitration between the United States and Canada. The award from the Trail Smelter Arbitration remains one of the modern touchstones of international environmental law because it established the principle that transboundary harms were to be prevented. As a basic overview, “[t]he case involved air pollution that originated in a smelter in Trail, British Columbia, and caused damage to farmlands located south of the Canada–United States border, in Washington State.”

The United States and Canada twice submitted the dispute for arbitration. However, it was the final decision in 1941 that set forth a cornerstone of international environmental law.

In key part, the arbiters found that “no state has the right to use or permit the use of its territory in such a manner as to

---

54 Id. at 790.
cause injury by fumes in or to the territory of another or the
properties or persons therein, when the case is of serious
consequence and the injury is established by clear and convincing
evidence.\(^{59}\) This “no harm” rule remains a common conceptual
underpinning governing state activity that has been reaffirmed in
various international agreements\(^{60}\) and by the International Court
of Justice (ICJ).\(^{61}\)

There is a strong lesson to be learned from these
transboundary environmental conflicts that could have a bearing
on Indian prohibition. The general idea that “nations have a
responsibility to not allow their territory to be used in ways that
cause environmental harm to, or within, the territory of other
nations”\(^{62}\) can be applied to the flow of alcohol. The Oglala Sioux
have remained dry since 1834 but have consistently been in
conflict with bootleggers and liquor distributors to stop the flow of
alcohol into Indian country. Based on the well-known harmful
effect alcohol has on Indian communities\(^{63}\) and the struggles the
Oglala Sioux have with illegal alcohol,\(^{64}\) the town of Whiteclay
is violating the “no harm” principle. By setting up four liquor
stores in a town of barely 10 residents meant to serve a dry
community, the nearly 4.3 million cans of beer and malt liquor
sold a year directly harm the territory of another nation. Tribes
need to be empowered to prevent their tribal decisions from
being undermined by external forces.

II. FIREWATER AND SOVEREIGNTY: AN OVERVIEW OF INDIAN
ALCOHOL PROHIBITION

Before addressing the specific parameters of federal tribal
prohibition, it is helpful to establish the rights and prerogatives of
Indian tribes to dictate activities on reservations and over non-
members. Thus, this part will begin by discussing the nature of
tribal sovereignty generally and move specifically through the

---

59 Trail Smelter Case, 35 AM. J. INT’L L. at 716 (responding to Question 2).
60 See, e.g., United Nations Conference on the Human Environment,
Declaration of the United Nations Conference on the Human Environment, princ. 21,
(1972); United Nations Conference on Environment and Development, Rio Declaration
61 Legality of the Threat or Use of Nuclear Weapons, Advisory Opinion, 1996
I.C.J. 226, ¶ 29 (July 8).
62 AARON SCHWARZ, INTERNATIONAL ENVIRONMENTAL DISPUTES: A
63 See Centers for Disease Control & Prevention, supra note 9, at 938-39.
64 See Williams, supra note 3.
changing landscape of tribal authority over the flow of alcohol into Indian country.

A. An Overview of Tribal Sovereignty

Tribal sovereignty is a varied and fluid concept. Attempting to discern the developing nature of this sovereignty, Felix Cohen chronicled three changes or principles that mark the modern day conception of tribal sovereignty. First, “prior to European contact, a tribe possessed ‘all the powers of any sovereign state’.”  

Second, following European conquest, tribes were “subject to the legislative power of the United States” and were stripped of their external sovereign powers. Third, “tribes retain internal sovereignty ‘subject to qualification by treaties and by express legislation of Congress.’” Thus, tribal powers generally are not ‘delegated powers granted by express acts of Congress,’ but instead are ‘inherent powers of a limited sovereignty which have never been extinguished.” However, there are instances where Congress delegates authority to tribal governments. In doing so, the tribe receives Congressional authorization to act in a similar manner as states in the federal system. Contemporary prohibition of alcohol is one such example.

B. The Historical Development of Indian Liquor Laws

The prohibition of alcohol on tribal lands has been part of the American experience since the founding. Following the American Revolution, the newly formed government began negotiating directly with Indian tribes to regulate trade and commerce. During these negotiations, tribal leaders implored the federal government to restrict the flow of alcohol to tribes through traders. President Thomas Jefferson is said to have been inspired to limit the flow of alcohol to tribes after hearing from a leading Indian chief about the devastating effects alcohol had on his people. President Jefferson was so moved by this letter that he

65 FELIX S. COHEN, HANDBOOK OF FEDERAL INDIAN LAW 123 (1941).
66 Philip P. Frickey, A Common Law for Our Age of Colonialism: The Judicial Divestiture of Indian Tribal Authority over Nonmembers, 109 YALE L.J. 1, 8 (1999).
67 Id.
70 Id. In his letter, Chief Little Turtle said:
called upon Congress to end the flow of alcohol into Indian country.\footnote{71}{FELIX S. COHEN, HANDBOOK OF FEDERAL INDIAN LAW 352 (Robert L. Bennett & Frederick M. Hart eds., Five Rings Press 1986) (1942).} Beseeching Congress, the President said, “These people (the Indians) are becoming very sensible of the baneful effects produced on their morals, their health, and existence, by the abuse of ardent spirits: and some of them earnestly desire a prohibition of that article from being carried among them. The Legislature will consider whether\footnote{72}{7 AMERICAN STATE PAPERS, INDIAN AFFAIRS 653, 655 (Walter Lowrie & Matthew St. Claire Clark eds., 1832).} legislative action can be taken.

Congressional response was quick. Under the Act of 1802, the President was authorized to “take such measures, from time to time, as to him may appear expedient to prevent or restrain the vending or distributing of spirituous liquors among all or any of the said Indian tribes.”\footnote{73}{Act of Mar. 30, 1802, ch. 13, § 21, 2 Stat. 139, 146; Miller & Hazlett, supra note 68, at 240–41.} This act was the first in a series of enactments meant at curbing the flow of alcohol to Indians. The negotiations between the federal government and Indian tribes did not cease after the Act of 1802; rather, a series of Trade and Intercourse Acts were passed to address various issues in Indian country.\footnote{74}{FRANCIS PAUL PRUCHA, AMERICAN INDIAN POLICY IN THE FORMATIVE YEARS: THE INDIAN TRADE AND INTERCOURSE ACTS 1790–1834 at 1–3, 43–50 (1962).} Collectively, “[t]he Acts outlined the authority of the states and national government over Indian affairs, rules regarding land distribution in response to intrusions on Indian land by non-Indian settlers, and the regulation of contact between Indians and non-Indians on the frontier.”\footnote{75}{Baker, supra note 13, at 691.}

The most important Trade and Intercourse Act—for our present purposes—was the Act of 1834. The Act prohibited the introduction of and attempts to introduce alcohol into Indian country. The only exemption was that alcohol could be provided to officers and troops of the United States. Moreover, the Act

But father, nothing can be done to advantage unless the great council of the Sixteen Fires, now assembled, will prohibit any person from selling any spirituous liquors among their red brothers . . . . Father: Your children are not wanting in industry; but it is the introduction of this fatal poison which keeps them poor. Your children have not that command over themselves, which you have, therefore, before anything can be done to advantage, this evil must be remedied. Father: When our white brothers came to this land, our forefathers were numerous and happy; but, since their intercourse with the white people, and owing to the introduction of this fatal poison, we have become less numerous and happy.

7 AMERICAN STATE PAPERS, INDIAN AFFAIRS 653, 655 (Walter Lowrie & Matthew St. Claire Clark eds., 1832).
clarified the definition of Indian country and established penalties for violators of the Act.\textsuperscript{76} While the Act sought to bring clarity to the issue of alcohol in Indian country, it caused significant confusion in some areas. For example, an 1854 amendment to the Act, which is no longer in effect, exempted Indians from prosecution—especially when an individual had already been prosecuted under tribal law.\textsuperscript{77}

C. Congressional Delegation to Tribes: Contemporary Policy

Contemporary law regulating the flow of alcohol into Indian country was codified during an era in which Congress attempted to abolish tribal sovereign power and assimilate tribal members as full, taxpaying citizens of the United States. Referred to as the Termination Era,\textsuperscript{78} Congress began in the 1950s to end the special relationship between the federal government and over 100 Indian tribes.\textsuperscript{79} These challenges resulted in a variety of important demographic changes in Indian country. Termination Era policy (1) changed land ownership patterns in and around Indian country; (2) ended tribal sovereignty and the trust obligation owed to tribes by the United States after revoking the special relationship; (3) granted states jurisdiction over Indian country; (4) revoked state taxation exemptions; and (5) curtailed tribes’ access to federal programs for tribes.\textsuperscript{80}

In keeping with the general trend of terminating tribal recognition and privilege, Congress began to fundamentally reshape federal Indian prohibition. Beginning in 1948, the absolute prohibition of introducing alcohol into Indian country, first adopted by the Act of 1834, was recodified. Under 18 U.S.C. §§ 1154 and 1156, the sale and distribution of alcohol in Indian country was considered criminal activity.\textsuperscript{81} Originally, this prohibition applied throughout Indian country regardless of who sold or distributed the alcohol.\textsuperscript{82} However, a year later, Congress amended the definition of Indian country to exclude “non-Indian communities or rights-of-

\textsuperscript{76} Miller & Hazlett, supra note 68, at 242–44.
\textsuperscript{77} Id. at 244. This caused significant conflict between the legislature and the courts. The Act’s impact was unclear; traditionally Indians had the sovereign authority to punish their own members, but this needed to be codified 20 years later. Thus, as one author has described, “the Act did more to reduce tribal authority than it did to curb the problems caused by rampant alcohol abuse in Indian country.” Baker, supra note 13, at n.32.
\textsuperscript{78} Baker, supra note 13, at 693.
\textsuperscript{79} Miller & Hazlett, supra note 68, at 262.
\textsuperscript{81} Baker, supra note 13, at 693.
\textsuperscript{82} Id.
way through Indian reservations” unless expressly extended by a treaty or statute extending prohibition to these now exempt areas. Additionally, Congress left intact § 3113, which includes broad language permitting any Indian to confiscate alcohol brought into Indian country and provides the steps federal officials can take to enforce prohibition.

This severe treatment was reversed by the mid-1950s. By 1953, Congress began to reject the paternalistic nature of Indian prohibition—a belief that Indian communities were unable to cope with alcohol. This longstanding belief, termed the “drunken Indian myth,” had persisted from the Act of 1802. In changing the policy toward prohibition, Congress turned over alcohol regulation in Indian country to the tribes. Overturning nearly 120 years of strict alcohol prohibition, Congress delegated the power to determine and apply policy during an era of significant deterioration in tribal sovereignty. In codifying § 1161, Congress “restored some of the sovereign powers that the [drunken Indian] myth had taken away, while on the other hand, it returned a shell too long eviscerated by prejudice to retain much power.” Congress did limit the sale of alcohol in Indian country to the boundaries of state law, but otherwise, the scheme relies upon tribal determination of alcohol policy. If a tribe legalized the flow of alcohol into Indian country, consistent with requisite state laws, federal prohibition would subside. Tribes could also choose to remain dry. This policy change is representative of the federal government’s current approach toward Indian tribes—self-determination. The current policy has been to turn control back over to tribes. The prohibition on alcohol is one of many areas in which this occurred.

The tribes’ response was mixed. After the first 18 months of passing § 1161, 22 tribes legalized alcohol on their reservations, and by 1974 that number had grown to 115. However, incrementally, tribes began to return to prohibition. By the 1990s, “approximately sixty-nine percent of the nation’s 293 reservations either immediately passed, or have since instituted, tribal ordinances ensuring that the blanket federal

---

85 Miller & Hazlett, supra note 68, at 225, 263.
86 Id. at 225.
88 Miller & Hazlett, supra note 68, at 225.
89 Id. at 267.
prohibition stays in place.”91 The most recent national survey reveals that this trend has dipped, but only slightly. As of 2008, 36% of tribes nationally (121 of 334) remain dry.92 As mentioned by one author, alcohol bans span the country, from the Yakama of Washington, to the Navajo of the Southwest, and the Oglala Sioux in the Dakotas.93 The effectiveness of this decision faces serious obstacles. Building upon the changes made during the Termination Era, much of the land within the exterior boundaries of the reservation belong to non-tribal members, and sham towns have developed along the borders to provide alcohol a mere feet outside the prohibited zone.

**D. The Controversy in Pine Ridge**

The Oglala Sioux made the decision to prohibit alcohol on tribal land. But what happens when alcohol distributers establish stores just outside the border of Indian country? For the Oglala Sioux, this has been a lingering question that has vexed the tribe for over a decade.94 This issue also carries significant national concern. Life expectancy on the reservation is 48 years for men and 52 for women while the national average is 78.95

Not only does the illegal flow of alcohol harm public health, it also places a significant burden on the community and law enforcement. Such has been the case on the Pine Ridge Reservation. The County Sheriff’s office, responsible for patrolling and monitoring Whiteclay, is 19 miles away.96 The Sheriff has five deputies. The tribal police department, which lacks jurisdiction in Whiteclay, has 38 officers—down from 101 six years ago.97 The Pine Ridge Reservation is roughly the size of Connecticut and has a population of approximately 45,000 people including non-tribal

---

91 Baker, supra note 13, at 694–95 (citing Study Says Tribal Alcohol Ban Increases Other Risks of Death, MINNEAPOLIS STAR-TRIB., Mar. 11, 1992, at 7A).
95 Id.; FastStats: Life Expectancy, CENTERS FOR DISEASE CONTROL & PREVENTION, http://www.cdc.gov/nchs/fastats/lifexp.htm (last visited Apr. 21, 2013) (noting that the national life expectancy is 78.7 years).
97 THE BATTLE FOR WHITECLAY, supra note 94.
members. Beyond that, in the Pine Ridge Reservation, two-thirds of the population resides below the poverty line, and the unemployment rate on the reservation is at 75%.

This subpart will chronicle the history of prohibition in the Pine Ridge Reservation by the Oglala Sioux people to illustrate the difficulties many tribes face in administering prohibition policy. First, the historical challenges in keeping the flow of alcohol out of the Pine Ridge area will establish the long-standing conflict between tribal authorities and distributors of alcohol located just beyond tribal borders. Next, the recent grassroots movements against extra-territorial distribution will demonstrate the difficulty tribes face in enforcing the policy when faced with external threats. Finally, the Oglala Sioux’s recent failure in federal district court will establish that the Congressional delegation over alcohol policy is a false promise of greater tribal authority and unenforceable, especially when undermined by external threats to public health—by both legislative and judicial means. Something else is needed to provide redress.

1. The Historical Harms of Alcohol in the Pine Ridge Reservation

Whiteclay is not a recent development. The sham town of at most 10 people “is a successor to the so-called whiskey ranches set up in the 1880s to move alcohol onto what was then called Pine Ridge Agency.”\(^{98}\) Concerned that bootleggers were ushering alcohol to the Sioux and at the urging of the U.S. Indian Agent and Oglala leaders, President Chester A. Arthur established a 50-square mile buffer zone in Nebraska, south of the Pine Ridge Reservation.\(^\text{99}\) By 1889, the Oglala Sioux received their own reservation,\(^\text{100}\) establishing the Pine Ridge Agency. The Agency occupied entirely within the borders of modern day South Dakota with the 50-mile extension into Nebraska. In both 1889 and 1890, Congress enacted legislation to continue the incorporation of the buffer zone, called the White Clay Extension, into the boundaries of the reservation,

---


\(^{100}\) Act of Mar. 2, 1889, 25 Stat. 888.
but only for “so long as it may be needed for the use and protection of the Indians . . . at the Pine Ridge Agency.”

In 1904, President Theodore Roosevelt ceded 49 of the 50 square miles into the public domain by executive orders, despite failing to establish that the need for the buffer zone had ceased to exist. In fact, these executive orders were carried out over the protests of Oglala leaders and the Indian Agent. This caused a land grab by white settlers who used the land to produce bootlegged alcohol, though alcohol consumption was still illegal nationally under the Eighteenth Amendment. Because prohibition still applied on the reservations, skirting prohibition did not cease either. Rather, “[t]he bootleggers who first supplied the liquor were replaced during the second half of the 20th century by bars and later by retail stores that were—and still are—licensed to operate by the state of Nebraska.” First, in the mid-1950s, the State of Nebraska licensed two bars in Whiteclay for on-site consumption of alcohol. After two decades, the owners converted their licenses for off-site drinking only. However, two more additional off-site stores were added. Thus, since the 1970s, four permanent off-site (beer only) liquor stores have served the barely 10-person town of Whiteclay, Nebraska.

These bars have gained notoriety. They commonly sell beer to “minors and intoxicated persons, knowingly sell[] to bootleggers who resell the beer on the reservation, permit[] on-premise consumption of beer in violation of restrictions placed on off-sale-only licenses, and exchange[] beer for sexual favors.”

Absurdly enough, the vast majority of those who purchase beer in Whiteclay have no legal place to consume the beer. The store’s permits prohibit on-site consumption of beer and it is illegal to possess or consume alcohol throughout the entire reservation.

2. Grassroots Efforts to Effectuate Prohibition

Over the years, the Oglala Sioux have tried to spotlight the tragedy occurring in Whiteclay and encourage the Nebraska legislature to take action. Largely, they have been unsuccessful. The grassroots movement to prohibit Whiteclay

101 Id. at § 1.
103 Mohr, supra note 99, at 2.
104 Woodard, supra note 98.
105 Mohr, supra note 99, at 2.
from selling alcohol, which undermines the Oglala’s ban on alcohol, has faced significant resistance from the Nebraskan government. Over the years, the “Nebraska authorities have repeatedly said it’s too ‘complicated’ to address the issue of Whiteclay and the misery it causes.”106 Though a series of proposed bills have been introduced in the Nebraska legislature, they have either failed, been gutted, or made little impact.

In 2002, State Senator Donald G. Priester introduced a law to prohibit issuing any future liquor license within the original White Clay Extension area, or within five miles of Indian country where the tribal council had banned the sale and consumption of alcohol.107 However, the bill grandfathered in existing liquor licenses, and so the four existing license holders would be able to continue to sell beer unimpaired.108 Senator Priester’s bill did not even make it out of committee. Nor did his identical proposal the next year.109 The Nebraska legislature only went so far as to pass a bill proposed by Senator Ray Jannsen, which empowered the Nebraska Liquor Control Commission to hesitate in granting new licenses in areas saturated with existing licenses.110 This bill passed in a modified form and was signed into law in 2006.111

During the 2009 session, the Nebraska legislature commissioned an interim study of Whiteclay. This included numerous public hearings before the General Affairs and Judiciary Committees. From this study, two separate bills were proposed to redress the harmful impact selling alcohol in Whiteclay was having on the Pine Ridge Reservation. First, Senator LeRoy Louden introduced a bill that earmarked alcohol sales tax revenue for economic development and healthcare or law enforcement assistance.112 The bill was quickly gutted. The amended bill made no mention of earmarking alcohol tax revenue, but rather, would establish a meager general assistance fund for Whiteclay. The final version of the bill was signed into law with $25,000 set aside for Whiteclay.113 Second, a bill introduced by Senator Russ Karpisek would have created the

106 Woodard, supra note 98.
107 Legis. B. 1306, 97th Leg., 2d Sess. (Neb. 2002).
108 Id.
111 Legis. B. 845, 99th Leg., 2d Sess. (Neb. 2006) (codified at NEB. REV. STAT. § 53-132) (provides that the Nebraska Liquor Commission has leeway to withhold additional licenses based upon new factors).
113 Legis. B. 1002 Slip Law Copy, 101st Leg., 2d Sess. (Neb. 2010).
Substance Abuse Treatment Grant Program under the Native American Public Health Act. Unfortunately, this bill had a worse fate than Louden’s and was indefinitely postponed.\textsuperscript{114} Senator Louden submitted another proposed bill. As introduced, the bill was designed to empower the Liquor Commission to grant “alcohol impact zone” status on localities to discourage public inebriation.\textsuperscript{115} Once an area was designated as an “alcohol impact zone,” the Liquor Commission was authorized to “promulgate rules and regulations” of alcohol sales to effectuate this policy.\textsuperscript{116} This bill also failed to become law.\textsuperscript{117}

Against this backdrop of Nebraskan legislative impotence, the Oglala Sioux tribe has continued to advocate for change. This has included blockages, marches, and public meetings. Specifically, “[a] New Year’s Eve protest of beer sales in Whiteclay ended peacefully with three of the Nebraska town’s four alcohol stores closing early and making very few late-night sales to Natives.”\textsuperscript{118} Recently, a new alcohol checkpoint has been set up at the border to try to stop the flow of alcohol into the reservation.\textsuperscript{119} Still, all told, the illegal flow of alcohol into Indian country remains a constant environmental problem.

3. The Federal Judiciary: A Dead End

After failing to produce a grassroots legislative resolution to the on-going problem in Whiteclay, the Oglala turned to the federal judiciary. Unfortunately, this too proved to be a dead end. Tribal leaders brought a lawsuit targeting “major beer manufacturers, including Coors, Miller and Anheuser Busch, distributors, and local stores that sell beer.”\textsuperscript{120} The lawsuit sought over $500 million to help cover the costs of healthcare, law enforcement, and social services resulting from over-consumption


\textsuperscript{115} Id. § 5.

\textsuperscript{116} See \textit{id.} at 1 (creating “alcohol impact zones”); Williams, \textit{supra} note 2 (noting the substantial sum paid by the “alcohol lobby” as campaign contributions to members of the General Affairs Committee who defeated the alcohol impact zone bill).


\textsuperscript{118} Id.

\textsuperscript{119} South Dakota Tribe Goes Up Against Big Brewers, NPR (Feb. 24, 2012), http://www.npr.org/2012/02/24/147348297/south-dakota-tribe-goes-up-against-big-brewers.
of alcohol within the community. Moreover, the tribes sought to restrict the amount of beer sold in Whiteclay—not to permanently prevent the operation of liquor in the area.

The case, *Oglala Sioux Tribe v. Schwarting*, raised three distinct causes of action: conspiracy to distribute alcohol that cannot be legally consumed on the reservation, conspiracy to violate 18 U.S.C. § 1161, and a refusal by the state of Nebraska to enforce its own statutes. The legal argument was premised on the belief that the brewers and store owners knew the alcohol sold would go to individuals with no legal place to consume it, and thus knew that individuals were purchasing the alcohol to smuggle it into the reservation for illegal use or resale.

The federal district court in Nebraska dismissed the suit, without prejudice, for lack of jurisdiction. Essentially, the court held that it lacked jurisdiction because “whatever merit the Tribe’s claims may have, they are not claims that depend on the resolution of a substantial question of federal law—and therefore, they are not claims that can be decided in federal court.” The lawsuit presented a number of jurisdictional challenges on review. At the onset, “it [was] clear that an Indian tribe is not a citizen of any state and cannot sue or be sued in federal court under diversity jurisdiction. Nor [were] Indian tribes foreign states.” Thus, as diversity jurisdiction was unavailable, the claim was dependent on a finding of federal question jurisdiction.

While the Oglala cited 18 U.S.C. § 1362, which confers the court with “original jurisdiction of all civil actions brought by any Indian tribe or band,” they turned first to the more general federal-question jurisdictional hook—§ 1331. The court moved forward hesitantly, noting that the statutes cited by the tribe failed to create a private right of action, and that their claims’ reliance upon “public policy” concerns did not meet the standard by which an issue arising under a federal statute can provide jurisdiction. For the court, “[t]he ultimate import of concluding that there is no federal private cause of action is that it would flout congressional intent to provide a private federal remedy for the violation of the federal statutes.”

Next, the court turned to the actual jurisdictional hook cited by the Oglala, § 1362. This too was unavailing. Under

---

122 Id. at 1205.
123 Id. at 1198–99 (internal citations omitted).
124 Id. at 1199 (quoting 28 U.S.C. § 1362).
125 Id. at 1200–01.
126 Id. at 1201.
Supreme Court precedent, the district court found that the common thread through cases that extended jurisdictional reach under § 1362 is that “they all involved possessory rights of the tribes to tribal lands.” The claims alleged by the Oglala Sioux related to conspiracy to flout the prohibition of alcohol, not tribal land. Thus, while the “United States could criminally prosecute anyone suspected of violating §§ 1154 and 1156, there [was] no basis . . . for concluding that the United States could, as a trustee, establish standing to allege the claims that the Tribe [sought] to allege in this case.” As a result, § 1362 did not extend jurisdiction. The Oglala Sioux were again stymied from enforcing prohibition within the Pine Ridge Reservation.

III. THE COOPERATIVE-FEDERALISM APPROACH TO TRIBAL PROHIBITION

“I am authorized by the great council of the United States to prohibit [sales of liquor]. I will sincerely cooperate with your wise men in any proper measures for this purpose.”

- President Thomas Jefferson

The ramification of the district court’s finding that it lacked jurisdiction to review the Oglala Sioux’s challenge, “[regardless of the] merit the Tribe’s claims may have,” is dissatisfying. Congress has delegated the choice of remaining dry to the tribes but they have been prevented from enforcing prohibition by transboundary activity. Moreover, neither legislative nor judicial action has stopped the flow of alcohol into Indian country. The federal system has given rise to numerous transboundary environmental disputes that have been litigated in the United State courts. In fact, “[i]n these cases, the Supreme Court is acting as an arbiter between sovereign states, not unlike an international court or arbitration panel” as in the Trail Smelter Case. Domestic history of transboundary environmental disputes has involved the harmful effects of

---

127 Id. at 1204.
128 Id.
130 Schwarding, 894 F. Supp. 2d at 1205.
smelter pollution settling upon a down-wind state, and sewage disposal traveling downstream to another urban setting.

In the context of our modern domestic environmental regulations, states are given significant authority to establish the methods of achieving nationally determined public health goals. This has been termed a “cooperative-federalism model” of regulation. Part of this cooperative approach is federally allowed state implementation of various quality standards. Recently, Congress delegated similar authority to Indian tribes. This has advanced tribal self-government and addressed environmental justice concerns in Indian country. There are a variety of well-known examples, but this essay will specifically review Congressional delegation under the CAA and CWA. In turn, understanding these processes by which tribes can enforce Congressional delegation will elucidate the cooperative-federalism nature of federal Indian prohibition.

A. The Clean Air Act

The purpose of the Clean Air Act is “to protect and enhance the quality of the Nation’s air resources so as to promote the public health and welfare and the productive capacity of its population.” The Clean Air Act covers various forms of air emissions, applying to both aging and decrepit power plants as well as ones to be built in the future. To fulfill this broad regulatory mandate, the Clean

---

132 Georgia v. Tenn. Copper Co. (Georgia II), 237 U.S. 474 (1915); Georgia v. Tenn. Copper Co. (Georgia I), 206 U.S. 230 (1907) (Georgia sought to enjoin Tennessee copper plants from polluting property located in Georgia).

133 Missouri v. Illinois (Missouri II), 200 U.S. 496 (1906); Missouri v. Illinois (Missouri I), 180 U.S. 208 (1901) (City of Chicago disposing sewage into the Mississippi that impacted St. Louis).

134 Jonathan H. Adler, When Is Two a Crowd? The Impact of Federal Action on State Environmental Regulation, 31 HARV. ENVTL. L. REV. 67, 87 (2007); see also New York v. United States, 505 U.S. 144, 167–68 (1992) (“Where Congress has the authority to regulate private activity under the Commerce Clause, we have recognized Congress’ power to offer States the choice of regulating that activity according to federal standards or having state law pre-empted by federal regulation . . . . This arrangement . . . has been termed ‘a program of cooperative federalism.’” (internal citations omitted)).


Air Act attempts to preserve air quality through interlaced health-based and technology-based standards. The National Ambient Air Quality Standard (NAAQS), defined under the Clean Air Act, establishes a cap of specified levels of emission for six criteria of air pollutants. The NAAQS emission caps are designed to reflect an adequate degree of safety necessary to protect public health. The areas of the country that do not meet the NAAQS for one or more of the six criteria pollutants are characterized as “nonattainment” zones for that pollutant because they have not “attained” the proper levels of each criteria pollutant. Likewise, those areas meeting NAAQS for criteria pollutants are considered “attainment” zones. While major stationary sources in both attainment and nonattainment zones must obtain a permit before new development or modification of the source, designation as nonattainment could significantly impact economic development. The distinction also obligates pollution-emitting entities to meet more aggressive regulatory obligations.

Moreover, the NAAQS regime operates on the basis of a “cooperative-federalism model.” Part of this cooperative approach is federally allowed state implementation of air quality standards. The Clean Air Act encourages states to submit State Implementation Plans (SIPs) to ensure compliance with the NAAQS. A state’s SIP must include a number of specific pollution control measures. Failure to submit an adequate SIP by the appropriate deadline subjects the state to various federal sanctions, such as the loss of federal highway funds or the imposition of an EPA-enforced Federal Implementation Plan (FIP).

The process of ensuring that new sources, or modifications made to existing sources, meet statutory obligations has become known as New Source Review (NSR). NSR is an essential component of the efforts to maintain air quality. At the same

139 42 U.S.C. § 7409(b)(1).
140 Id. at §§ 7501(2), 7407(d)(1)(A)(i).
141 Id. at § 7503(a).
143 See supra note 134 and accompanying text.
145 Id. § 7410(c)(1)(B).
time, the NSR program has been a cause of contention between the EPA, state air quality agencies, and existing facilities.\textsuperscript{146} From its inception, the CAA has required new sources to install advanced pollution control technology.\textsuperscript{147} This policy decision established two different regulatory regimes—those for new and old sources. While many old sources were “grandfathered” in such that they were able to continue operating subject to few restrictions, new sources that produce air pollution “which may reasonably be anticipated to endanger public health or welfare” are required to install federally established new source performance standards (NSPS).\textsuperscript{148} The NSR program thus adds another layer of regulation to facilities that may already be subject to NSPS standards.

Permitting under NSR is divided into three programs. First, the prevention of serious deterioration (PSD) program applies to new major sources or sources making major modifications to existing sources in attainment zones.\textsuperscript{149} Second, the Nonattainment NSR (NNSR) program is designed to promote air quality in areas that are out of compliance with one or more of the NAAQS.\textsuperscript{150} “NSR requires the most stringent emission limits and also requires sources to offset increased emissions by reducing emissions elsewhere at the facility, or by obtaining Emission Reduction Credits (ERCs) from nearby facilities.”\textsuperscript{151} Finally, the Minor Source NSR program provides permits for sources not covered by PSD or NNSR. “The purpose of minor NSR permits is to prevent the construction of sources that would interfere with attainment or maintenance of an [NAAQS] or violate the control strategy in nonattainment areas.”\textsuperscript{152} Together

\begin{footnotesize}
\begin{enumerate}
\item The friction between the EPA and state air quality agencies will be highlighted \textit{infra}.
\item See 42 U.S.C. § 7411(b).
\item \textit{Id.} § 7411(b)(1)(A) (requiring the Agency to set emission performance standards for stationary sources that “cause[] or contribute[] significantly to, air pollution which may reasonably be anticipated to endanger public health or welfare”).
\item 40 C.F.R. § 52.21 (2013).
\item 40 C.F.R. pt. 51, App. S(D).
\end{enumerate}
\end{footnotesize}
these permitting requirements clarify the law governing new sources of modification to sources throughout the United States.

The last 40 years in federal Indian law have been marked as a time of tribal self-determination. Throughout this period, the EPA has reinforced this self-determination model. In 1984, the EPA became the first federal agency to formally adopt an “Indian policy” governing its interactions with tribes. The policy noted that tribal governments are “sovereign entities with primary authority and responsibility for the reservation populace.” Accordingly, the policy assumes tribal regulatory responsibility but is premised upon a default model where the EPA retains responsibility until a tribe assumes responsibility. Further, the EPA, with tribal assistance, secured amendments inserting general “treatment as state” (TAS) provisions in most of the major environmental statutes. Specifically, for the purposes of administering air programs, tribes have been granted TAS status under the CAA.

Next, to implement CAA’s “treatment as states” provision, the EPA promulgated the Tribal Authority Rule (TAR) in 1998. Under 42 U.S.C. § 7410, tribes can develop TIPs, the tribal equivalent of State Implementation Plan (SIPs), to regulate air quality on reservations. TIPs are

---


155 Id. (emphasis added).


157 James M. Grijalva & Daniel E. Gogal, The Evolving Path Toward Achieving Environmental Justice for Native America, 40 ENVTL. L. REP. 10905 n.46 (2010) ("[TAS] eligibility criteria vary among the statutory programs but require generally that the tribe be federally recognized by the U.S. Department of the Interior, have a governing body carrying out substantial duties and powers, and demonstrate technical capability and legal authority to manage and protect the Indian country environment.").


159 Clean Air Act of 1990, 42 U.S.C. § 7601(d)(1)(B) (2012); see id. § 7474(c) (stating that re-designation of reservation air quality standards can only be done by “the appropriate Indian governing body”).


161 42 U.S.C. § 7410(o) (“Indian tribes. If an Indian tribe submits an implementation plan to the Administrator pursuant to section 7601(d) of this title, the
created to implement a regime to regulate air quality to meet the air quality standards required under the NAAQS, and TIPs must neither interfere with PSD nor hinder the air quality of neighboring states or tribal areas. Unlike SIPs, TIPs can allow joint tribal and EPA management. Finally, Indian tribes have the same authority as states to petition the EPA to enforce Clean Air Act requirements on surrounding states or tribes.

This regulatory framework has provided tribes authority to control air quality management decisions affecting their jurisdiction. Since the 1990 amendments, 32 tribes have received TAS under the TAR. Three separate tribes have successfully petitioned the EPA for approval of TIPs to implement and enforce tribally-designed air quality standards. Finally, “one tribe has received a delegation (under Clean Air Act Part 71) to implement a Title V operating permit program for their reservation.” Such steps have furthered the model articulated through the EPA’s Indian policy. This program is “based on initial federal implementation where feasible, with aspirations for later program assumption by Indian tribal governments.” To date, this pattern has largely been followed. “[T]ribes have demonstrated increasing interest in developing and administering their own air programs. As one illustration, the number of tribes receiving federal grants to initiate or operate air programs has

plan shall be reviewed in accordance with the provisions for review set forth in this section for State plans, except as otherwise provided by regulation promulgated pursuant to section 7601(d)(2) of this title. When such plan becomes effective in accordance with the regulations promulgated under section 7601(d) of this title, the plan shall become applicable to all areas (except as expressly provided otherwise in the plan) located within the exterior boundaries of the reservation, notwithstanding the issuance of any patent and including rights-of-way running through the reservation.”).  


See 42 U.S.C. § 7426 (explaining state’s authority to petition the EPA to exercise its enforcement powers).


Grijalva & Gogal, supra note 166.

Grijalva & Gogal, supra note 157, at 10905.
grown from about 20 in 1995 to more than 120 in 2002."\textsuperscript{170} Tribes have exercised the power to develop permitting programs for new or modified stationary source polluters,\textsuperscript{171} craft and implement CAA air quality standards,\textsuperscript{172} and potentially influence neighboring states’ air policies.\textsuperscript{173}

Air pollution has a distinctly extraterritorial nature. So too, it seems, does the flow of alcohol. Both programs are permitting tribes additional authority that Congress has delegated to address specific public health concerns.

\textbf{B. The Clean Water Act}

Like the concerns that inspired the passage of CAA, Congress has recognized the threat that unclean water posed to public health and welfare. In response, Congress enacted the Federal Water Pollution Control Act (FWPCA) in 1948 to “establish a national policy for the prevention, control and abatement of water pollution."\textsuperscript{174} The FWPCA was amended and eventually codified as the Clean Water Act (CWA) that included the objective of “restor[ing] and maintain[ing] the chemical, physical, and biological integrity of the nation’s waters.”\textsuperscript{175} The CWA also operates under a cooperative-federalism regime. Thus, “[w]hile the EPA is responsible for setting minimum WQSs [Water Quality Standards] for certain pollutants that all states must meet, states are free to set standards that are more stringent than the EPA requires.”\textsuperscript{176}

Like the CAA, the CWA also treats tribes like states for specific regulation. The “‘Treatment as a State’ or ‘TAS’ provision was a ‘prequalification’ requirement that, once satisfied, allowed the qualifying tribe to become eligible to apply for these grants and program approvals.”\textsuperscript{177} Importantly, tribes are permitted to address up-stream pollution.\textsuperscript{178} To do so,

\begin{itemize}
  \item \textsuperscript{170} Milford, supra note 156, at 213–14.
  \item \textsuperscript{171} 42 U.S.C. § 7502(c)(5) (2012).
  \item \textsuperscript{172} Id. § 7410.
  \item \textsuperscript{173} Vanessa Baehr-Jones & Christina Cheung, An Exercise of Sovereignty: Attaining Attainment for Indian Tribes Under the Clean Air Act, 34 ENVIRONS ENVTL. L. & POLY J. 189, 191 (2011).
  \item \textsuperscript{175} 33 U.S.C. § 1377(e) (2012).
  \item \textsuperscript{176} Marren Sanders, Clean Water in Indian Country: The Risks (and Rewards) of Being Treated in the Same Manner as a State, 36 WM. MITCHELL L. REV. 533, 536 (2010).
  \item \textsuperscript{177} Id. at 537; see also Royster, supra note 44, at 206.
  \item \textsuperscript{178} Royster, supra note 44, at 206 (“[A]ny [National Pollution Discharge Elimination System] permit issued for a point source upstream of the tribe’s territory, whether the permit issued by the EPA or the state pursuant to a delegated NPDES
they must petition the EPA—who may grant approval—which is then used by the tribe to limit pollution that can impact their local water quality.

This ability to enforce water quality beyond Indian country is similar to the needs of tribes to enforce prohibition beyond the reaches of Indian country. Because both of these models fit within a cooperative-federalism scheme, this process could serve as a model for reform.

IV. ENVIRONMENTAL JUSTICE RESPONSES TO TRIBAL PROHIBITION

A. Identifying a Private Right of Action

While the Oglala Sioux were foiled in the district court, this does not need to be the end of their legal challenges. To begin, the case was dismissed without prejudice. There is a chance to replay this challenge. Nationally, tribes concerned about similar transboundary harms should consider raising these challenges in different federal courts. This is much like federal agencies defending their statutory interpretation in different circuit courts.

Additionally, raising a statutory claim is a classic strategy in the environmental justice toolbox. As Luke Cole has noted, environmental law claims are preferred over all other strategies. However, before challenges are raised, additional thinking about how to articulate a private right of action is necessary. As the district court expressed in Schwarting, “none of the statutes cited create a private right of action” that can be alleged by the tribe. Without a right of action, there is no jurisdiction, and without jurisdiction, the claims will be dismissed. Stating that Congress implied a right of action is in many ways a long-shot. Generally, a private citizen cannot file suit under a federal statute unless Congress has

program, must include conditions to ensure compliance with the downstream tribe’s [Water Quality Standards]."

179 Schwarting, 894 F. Supp. 2d at 1205.
180 Samuel Estreicher & Richard L. Revesz, Nonacquiescence by Federal Administrative Agencies, 98 YALE L.J. 679, 735-36 (1989) ("Given the lack of intercircuit stare decisis, and the reasons underlying our system of intercircuit dialogue, an agency’s ability to engage in intercircuit nonacquiescence should not be constrained.").
181 Cole, supra note 55, at 526 (discussing the litigation hierarchy).
182 Schwarting, 894 F. Supp. 2d at 1200.
183 An implied right of action allows private individuals to file lawsuits under a federal statute that does not explicitly provide for such a right.
manifested its intent for such suit. In cases involving implied rights of action, federal courts look at the statutory language and legislative intent (solely to clarify the text).

By delegating the power to tribal governments to regulate alcohol throughout their reservations, the tribes are in fact acting as agents of the federal government. From this perspective, the language of both §§ 3113 and 1161 suggest that a right of action does exist to enforce prohibition beyond the reservation’s border. First, under § 3113, the language of the statute indicates congressional intent for tribes to enforce prohibition. The section specifically provides that “any Indian may take and destroy any ardent spirits or wine found in the Indian country.” Next, the section permits any Indian to seize and destroy alcohol found on fee land—land belonging to the Reservation. Thus, “the grant to ‘any Indian’ would be construed as also authorizing the organized efforts of the tribal government to make effective use of this power.” This empowered tribal members to take actions in furtherance of prohibition against non-members. Undercutting this argument, however, is that, following the reauthorization of § 3113, Congress narrowed the authority of tribes in §§ 1154 and 1156 to exclude non-members and pass-through roads from regulation. A court may view these subsequent enactments as constricting broad power of tribes to enforce prohibition.

Second, under § 1161, Congress gave Indian tribes the authority to maintain prohibition in Indian country. Few federal courts have reviewed the full extent of this delegation. In City of Timber Lake v. Cheyenne River Sioux Tribe, the Eighth Circuit held that § 1161 provided tribes the authority to regulate alcohol throughout land within the boundaries of the reservation—including non-member communities. Importantly, the court

---

184 See e.g., Cannon v. Univ. of Chi., 441 U.S. 677, 732-34 (1979) (Powell, J., dissenting) (reviewing history of Supreme Court decisions implying or prohibiting a private right of action).


186 Baker, supra note 13, at 714.


188 Id.

189 Baker, supra note 13, at 716.

190 10 F.3d 554, 558 (8th Cir. 1993).
viewed tribal jurisdiction over the flow of alcohol as spanning the “reservation’s four corners.”\textsuperscript{191}

Asserting enforcement \textit{beyond} the four corners of the reservation is a leap. The Supreme Court has been cautious when permitting extraterritorial enforcement. Historically, however, “the United States can, through treaties with Indian tribes, preempt state regulations affecting off-reservation treaty rights.”\textsuperscript{192} For example, in \textit{United States v. Forty-Three Gallons of Whiskey}, the Court held that the United States and the Red Lake and Pembina Band of Chippewa Indians could agree by treaty to extend federal prohibition to lands ceded by the tribe in treaty agreements.\textsuperscript{193} Like Whiteclay, Nebraska, the ceded lands were organized into a county in Minnesota but the Court held that this was within Congress’s treaty power. Also, in \textit{United States v. Holliday},\textsuperscript{194} the Court held that Congress could enact a statute preempting state law to prohibit the sale of liquor to individual Indians in a county where there was no Indian reservation. Finally, the Court has more recently described the limits of congressional power. In \textit{Perrin v. United States},\textsuperscript{195} the Court held that “the power [to control the flow of alcohol] is incident only to the presence of the Indians and their status as wards of the Government, [and thus] it must be conceded that it does not go beyond what is reasonably essential to their protection.”\textsuperscript{196} Moreover, “to be effective, its exercise must not be purely arbitrary, but founded upon some reasonable basis.”\textsuperscript{197} Ultimately, the test settles upon wide discretion for Congress to act.\textsuperscript{198}

While it helps to view the prohibition delegation to tribes as akin to delegation to tribes to enforce environmental regulations, there are clear distinctions. Unlike \textit{Forty-Three Gallons of Whiskey}, enforcement of prohibition in Pine Ridge is not based upon treaty obligations. In fact, it should be remembered that President Roosevelt specifically sold off the buffer zone created around the reservation. Additionally, unlike the CAA and CWA, Congress did not provide a concrete mechanism to enforce these regulations.

\textsuperscript{191} Id.
\textsuperscript{192} Skibine, \textit{supra} note 12, at 1031.
\textsuperscript{193} 93 U.S. 188, 189 (1876).
\textsuperscript{194} 70 U.S. (3 Wall.) 407, 419 (1865).
\textsuperscript{195} 232 U.S. 478 (1914).
\textsuperscript{196} Id. at 482, 486.
\textsuperscript{197} Id. at 486.
\textsuperscript{198} Id. (”[I]t must also be conceded that, in determining what is reasonably essential to the protection of the Indians, Congress is invested with a wide discretion, and its action, unless purely arbitrary, must be accepted and given full effect by the courts.”).
beyond the border. In fact, many scholars have even found the City of Timber Lake’s hold to be an overly broad application of the statute and Supreme Court precedent. However, while the history of alcohol prohibition has trended toward restricting the ability of tribes to make decisions concerning alcohol—based upon the drunken Indian myth—the best indication of congressional intent is the delegation of authority to the tribes. If this delegation is to mean anything, tribes must be able to assert prohibition beyond their borders.

B. File a Public Nuisance Suit

If the door to the courts remains closed to statutory challenges, tribes should seek to file a public nuisance suit. Importantly, “public nuisance law provides the broadest potential for raising environmental justice claims.” Public nuisance claims are premised on the notion that someone’s actions are “unreasonable given the circumstances and could cause injury to someone exercising a common, societal right.” Other types of public nuisance include interfering with public health and safety.

Admittedly, public nuisance suits face an uphill battle in court, and the likelihood of obtaining a positive ruling in a public nuisance suit is attenuated at best. To begin, the special injury rule/different-in-kind doctrine limits the type of plaintiff who can bring this tort action. In particular, “[o]nly those who can first prove some injury that is ‘special,’ ‘particular,’ or ‘peculiar,’ defined as ‘different-in-kind’ and not just ‘different-in-degree’ from the general public who might also be affected by the nuisance, be it a house of prostitution or a polluting factory, may bring an action.” Few state courts have shifted away from the traditional different-in-kind standard. This makes it difficult to raise environmental justice claims through public nuisance.

---

199 See, e.g., Baker, supra note 13, at 719-21.
202 See Friends of the Sakonnet v. Dutra, 738 F. Supp. 623, 635-36 (D.R.I. 1990). Examples include storing explosives within the city, interfering with reasonable noise levels at night, or interfering with breathable air, such as through emitting noxious odors into the public domain. See RESTATEMENT (SECOND) OF TORTS § 821B cmt. b (1979).
203 Denise E. Antolini, Modernizing Public Nuisance: Solving the Paradox of the Special Injury Rule, 28 ECOLOGY L.Q. 755, 761 (2001); see also Garrells, supra note 200, at 163–64.
In the context of climate change litigation, advocates for climate justice raising tort claims have also been stymied in the federal courts. First, in *American Electric Power Co. v. Connecticut (AEP)*, the Supreme Court held that the Clean Air Act (CAA) displaced any federal common law claims against carbon-dioxide emissions. Second, in *Native Village of Kivalina v. ExxonMobil Corp.*, the Ninth Circuit held that the CAA displaced all federal public nuisance claims. These decisions effectively close the door to the use of tort-based claims to address climate change. However, tort litigation in state courts may provide an area for successful litigation.

Notwithstanding the high bar of proving negligence, duty, and causality, courts are hesitant for three additional reasons. First, courts might be hesitant to regulate conduct of an industry because regulation is not a core competency of the judiciary. Second, judges may fear criticism for stretching the boundaries of tort law to cover conduct that is attenuated to the plaintiff’s injury. Finally, there is a concern about developing “slippery slope” precedent.

As a result of these constrictions on public nuisance claims, while tribes like the Oglala Sioux could utilize public nuisance theory to effectuate prohibition, the bar will be high. They will need to establish that the conduct is the type covered by nuisance theory. Factually, they will be required to establish a linkage between the distribution of alcohol outside their reservation and the significant health effects that result. Doing so will require development of a detailed record and could be quite costly. This limits the overall effectiveness of public nuisance suits as an environmental justice tool.

---


207 Id. at 2537 (“[T]he Clean Air Act and the EPA actions it authorizes displace any federal common law right to seek abatement of carbon-dioxide emissions from fossil-fuel fired power plants.”).


209 Id. at 858 (extending the rule established in *AEP*).


C. **Amend Title 18 to Empower the Tribes to Effectuate Prohibition**

If the preceding challenges are unsuccessful, tribal members should strongly consider lobbying Congress to provide tribes the ability to truly enforce prohibition. Delegating this responsibility to tribes is only a half-measure. Tribes must also be given the ability to enforce prohibition beyond reservation borders. In this respect, prohibition policy is like environmental policy. In both the CAA and CWA, tribes are treated like states and are provided a method to enforce these statutes outside their jurisdiction. To prevent transboundary harms, Congress should expand tribal power beyond reservation boundaries. Specifically, amending Title 18 to mirror congressional delegations under the CAA and CWA will empower tribes to actually enforce prohibition—a choice they are permitted to make within the cooperative-federalism scheme.

A possible amendment would establish a petition authority permitting tribes to petition the EPA to find that a site endangers public health beyond the requisite threshold and would have serious, harmful impacts upon the nearby states. As applied, tribes would retain the ability to enforce prohibition and would be permitted to petition the EPA to gain approval to enforce this policy in federal court. This approach is similar to that codified in the CAA and CWA which permits tribes to extraterritorially enforce CWA after receiving EPA approval. While both processes require federal approval—thus undercutting tribal prerogatives—this is consistent with the conceptualization of prohibition as cooperative-federalism. As Sarah Krakoff has noted, tribal environmental justice utilizes regulatory authority as a means of furthering self-government. Because tribes do not retain absolute sovereignty in this context, unlike many others where tribal authority is significant, this scheme is understandable. Just as states are forced to seek federal approval of myriad programs, tribes would be too. This scheme would permit tribes to fulfill their decision to remain dry by garnering federal support of

---


213 See 42 U.S.C. § 7426(b) (2012) (explaining state’s authority to petition the EPA to exercise its enforcement powers).

214 Tribes should petition the EPA over other agencies within this proposed reconfiguration of prohibition policy because alcohol is to be understood as an environmental harm.

215 See supra note 18.
extraterritorial enforcement of prohibition. This would promote environmental justice in Indian country.

* * * *

The story of the Oglala Sioux’s fight against the flow of alcohol on the Pine Ridge reservation has undergone a recent, significant change. Bowing to growing pressure to legalize the sale of alcohol, hopeful for the opportunity to regulate its use, and frankly exhausted from decades of conflict, the tribe voted to end prohibition. In a close election—1,871 for legalization and 1,679 against it—the lasting sentiment of those supporting the end of prohibition was that the history of the Pine Ridge reservation showed that “prohibition didn’t work.” Under the new law, “the tribe will own and operate stores on the reservation, and profits will be used for education and detoxification and treatment centers, for which there is currently little to no funding.” Opponents of the law fear that the change will, at best, begin reducing dependence on alcohol, but that alcohol-related crime and violence will spike. At bottom, “[b]oth sides in the debate do agree something must be done to limit the scourge of alcohol on the Lakota people. They also share a goal of putting out of business the current main suppliers of booze—four stores in Whiteclay, Neb[raska].”

CONCLUSION

The flow of alcohol into Indian country is a serious public health concern. New ideas are needed to address its harmful consequences. This essay has framed prohibition policy in an environmental justice context. This illustrates how sham towns, such as Whiteclay, are located near poor communities just as polluting facilities are located in or around poor communities. Both establish inequitable distributions of environmental hazards. Moreover, the siting of alcohol distribution centers can be considered LULUs just as local smelting plants are. But tribal governments have limited jurisdiction to assert their rights. This raises the question, does “tribal sovereignty... end at the reservation border[?]” This essay has suggested a variety of new

217 Id.
218 Id.
219 Id.
220 Id.
221 Skibine, supra note 12, 1004.
approaches based upon concrete environmental justice strategies to address this problem. Conceptualizing prohibition as an environmental justice concern—for the first time—highlights the breadth of tools available to the environmental poverty lawyer. Tribal governments can reassert statutory authorization of a private cause of action, bring a public nuisance suit, or lobby Congress to amend Title 18 to further empower tribes. Any of these strategies, if successful, would be ideal. Regardless, until change occurs, tribal communities will continue to struggle against an increasing threat from the flow of alcohol. This is an injustice.
The Securities Act of 1933
A JURISDICTIONAL PUZZLE

INTRODUCTION

On February 1, 2012, the social network giant Facebook, Inc. filed for an initial public offering (IPO).1 Although going public raised $16 billion for the company,2 Facebook’s first day of trading was plagued with problems and signaled the trouble that lay ahead.3 “[T]echnical glitches on [NASDAQ] created confusion” as its systems were unprepared for the “massive volume of the highest-profile IPO of the year.”4 Despite the issues, demand for the shares was unprecedented. Within the first 30 seconds of trading, 80 million shares had changed hands.5 By the end of the day, the stock reached a trading volume of 567 million shares, smashing the previous volume record of around 450 million by General Motors.6

However, the trading issues were not the company’s only problems. Just days after the IPO, class-action lawsuits began to pour into courts across the country.7 Some plaintiffs

5 See Pepitone, supra note 3.
6 See id.
7 See, e.g., Complaint at 15, Brian Roffe Profit Sharing Plan v. Facebook, Inc., No. 12 CV 4081 (S.D.N.Y. May 23, 2012); Complaint at 16, Spatz v. Facebook, Inc., No. 12 CV 2262 (N.D. Cal. May 23, 2012).
asserted federal claims, others brought class actions in state courts. To consolidate some of these actions, Facebook removed the class actions from the Superior Court of California, leaving the District Court for the Northern District of California to face an issue that has evaded an answer for many years: whether class actions alleging claims arising solely under the Securities Act of 1933 (Securities Act) can be removed from state to federal court.

The landscape of federal securities law was created nearly 80 years ago. Franklin D. Roosevelt took office in 1933 while the country was still in the depths of the Great Depression. With 13 million Americans unemployed, and the stocks listed on the New York Stock Exchange decimated, many pointed to corporate and stock market abuse as the cause of the crash. President Roosevelt made clear that his New Deal would include comprehensive securities reform. President Roosevelt turned to Felix Frankfurter, “a legendary Harvard law professor” and one of his most trusted advisors, to draft the securities laws.

---

8 They brought claims under sections 11, 12, and 15 of the Securities Act of 1933. Securities Act of 1933, 15 U.S.C. §§ 77k, 77l, 77o (2012). See Complaint at ¶ 2, Brian Roffe Profit Sharing Plan, No. 12 CV 4081. The plaintiffs asserted that Facebook, Inc.’s registration statement and accompanying prospectus “contained untrue statements of material facts, omitted to state other facts necessary to make the statements made not misleading and were not prepared in accordance with the rules and regulations governing their preparation.” Id. ¶ 20.


13 See id. (quoting JOEL SELIGMAN, THE TRANSFORMATION OF WALL STREET 19 (1982)). In his inaugural address, the president “denounce[d] the ‘unscrupulous money changers who stand indicted in the court of public opinion, rejected in the hearts and minds of men.’” Poser, supra note 12, at 1 (quoting KENNETH S. DAVIS, FDR: THE NEW DEAL YEARS 1933–1937, 30 (1979)).

14 Poser, supra note 12, at 2. Six years later, Professor Felix Frankfurter was appointed to the Supreme Court. Id. The actual drafting of the statutes was accomplished by two of Frankfurter’s protégés, Benjamin Cohen was “a shy, soft-spoken religious idealist with a tough, practical lawyer’s mind”; and Thomas “Tommy the Cork” Corcoran was an “exuberant Irishman of great personal charm, who liked to sing ballads, accompanying himself on the accordion.” Id. Despite their conflicting personalities, both “shared a sophisticated approach to socio-economic problems and a capacity for hard, prolonged intellectual effort.” Id. (citing SELIGMAN supra note 14, at 62-63 (1982)).
In the following years, two statutes were enacted: the Securities Act of 1933 (Securities Act), which regulated the issuance and distribution of securities; and the Securities Exchange Act of 1934 (Exchange Act), which “regulated the . . . trading markets—meaning the New York Stock Exchange—and which outlawed market manipulation.”

The Securities Act was designed to protect investors and increase investor confidence in the market through a system of “full and fair disclosure of securities sold in interstate and foreign markets.” For further protection, the Securities Act created a private right of action and gave concurrent jurisdiction in both state and federal courts. The statute also prevented the removal of these claims brought in state court.

---

18 See Poser, supra note 12, at 2. The Pecora hearings before Congress helped “galvanize[] broad public support for the securities laws.” See id. The hearings revealed that National City Bank (predecessor to Citicorp) “had aggressively pushed the sale of Peruvian bonds to the public,” despite the bank’s own representatives telling the bank that Peru was not likely to repay the interest or principal on the debt. See id. (citing JOHN KENNETH GALBRAITH, THE GREAT CRASH: 1929 171 (1955); SELIGMAN, supra note 14, at 27-28 (1982)).

The hearings revealed manipulation in the offering of securities. The public was informed that “J.P. Morgan & Co. . . . had a ‘preferred list’ of influential individuals who received stock in securities distributions at a low price shortly before they went public at a much higher figure,” enabling these individuals to sell the securities for “a sure profit.” See id. at 2-3 (citing SELIGMAN, supra note 14, at 34-35 (1982)).

The Pecora hearings also disclosed that “the market abuses of the 1920s were continuing in full force during the opening months of the New Deal.” See id. at 3 (citing KENNETH S. DAVIS, FDR: THE NEW DEAL YEARS 1933–1937 362 (1979)). Operators, including members of prestigious investment banks, company officers, specialists on the New York Stock Exchange, and even the father of president John Kennedy, Joseph Kennedy, created a “false effect of great activity and widespread buying that played on the gullibility and greed of the public.” See id. at 3.

In light of these abuses, the legislation garnered large public support, and with the help of Sam Rayburn, then Chairman of the House Commerce Committee and later the Speaker of the House, the two laws passed through Congress. See id.

19 See Denise Mazzeo, Securities Class Actions, CAFA, and a Nationwide Crisis: A Call for Clarity and Consistency, 78 FORDHAM L. REV. 1433, 1441 (2009).
20 Michael Sorota, (Mis)Interpreting SLUSA: Closing the Jurisdictional Loophole in Federal Securities Class Actions, 7 BERKELEY BUS. L.J. 162, 164 (2010). This was accomplished through the “filing of a registration statement with the Securities and Exchange Commission and providing prospective investors with detailed information” concerning the securities. Id.
21 Securities Act of 1933, 15 U.S.C. § 77v(a) (1998) (“The district courts of the United States and United States courts of any Territory, shall have jurisdiction . . . concurrent with State and Territorial courts, of all suits . . . brought to enforce any liability or duty created by this subchapter.”).
22 Id. (”[N]o case arising under this subchapter and brought in any State court of competent jurisdiction shall be removed to any court of the United States.”).
However, investor protection “seems to ebb and flow with the market.” In times of economic troubles, like the Great Depression, investor protection increases, because investor confidence is a key to shifting a bear market to bull. And until the 1990s, Section 22 of the Securities Act, the section concerning jurisdiction, remained largely untouched. But in times of prosperity, “Congress seems to be less concerned with protecting investors, and more so with deregulating and preventing litigious abuses of the system.” And during the economic boom of the 1990s, Congress whittled away the broad investor protections built up by the Securities Act.

In 1995, Congress “attempted to limit the number of securities class actions” by passing the Private Securities Litigation Reform Act of 1995 (PSLRA). The law purported to fight meritless class actions alleging fraud in the sale of securities (strike suits). However, plaintiffs began to strategically circumvent the PSLRA by filing class actions in state court, where federal law did not apply and where the Securities Act enjoyed concurrent and non-removable jurisdiction.

To prevent the evasion of the protections the PSLRA provides against abusive litigation, Congress enacted the Securities Litigation Reform Act of 1998 (SLUSA). Originally, Section 22(a) of the Securities Act barred removal of Securities Act claims brought in a state court of competent jurisdiction. SLUSA amended this provision by allowing the removal of certain covered class actions. But federal courts have “struggled with [SLUSA’s] application to class action removal for claims arising [solely] under the [Securities Act].”

23 See Mazzeo, supra note 19, at 1436.
24 Id. at 1436-37. The traditional definition of a “bear market” is a decline in the average price of stocks by 20% or more from the most recent high point, while the traditional definition of a “bull market” is a 20% increase from the most recent low point. See Tom Lauricella, Is This Bull Cyclical or Secular?, WALL ST. J. (June 15, 2009), http://online.wsj.com/article/SB124501817200213499.html.
25 See Mazzeo, supra note 19, at 1444.
26 Id. at 1436.
27 See id. at 1437.
28 Serota, supra note 20, at 164.
30 Serota, supra note 20, at 164.
31 Id. at 165.
35 Serota, supra note 20, at 165.
Some district courts have construed the amendment narrowly, finding that the plain meaning of its text only allows for removal of state law claims. Thus, Securities Act claims, asserted without any corresponding state claims (pure Securities Act claims), are still barred from removal by Section 22(a). Other courts have construed it broadly, looking toward legislative intent to find that pure Securities Act claims are removable. The issue is further complicated by a lack of binding authority. For example, “[p]rocedural rules make it . . . difficult for federal appellate courts to review . . . [a district court’s] decision[] to remand [Securities Act] class action claims [back to state court],” and other district courts only provide persuasive authority.\(^{36}\)

Some courts, however, are beginning to look past SLUSA’s amendment to the anti-removal provision of the Securities Act;\(^{37}\) instead, they focus on SLUSA’s amendment to the concurrent jurisdiction provision of the Securities Act.\(^{38}\) These courts conclude that the amendment to the concurrent jurisdiction provision completely removed state court jurisdiction over pure Securities Act claims. The anti-removal provision only applies to claims brought in a “State court of competent jurisdiction.”\(^{39}\) Therefore it simply does not bar pure Securities Act claims from removal in the first place. Because this approach addresses the shortcomings of other interpretations, courts should apply this analysis when faced with the issue of removal of a class action asserting Securities Act claims.

The note proceeds in three parts. Part I provides background on the Securities Act and the two subsequent statutes amending its provisions: the PSLRA and SLUSA. Part II focuses on two of SLUSA’s amendments to the Securities Act and the judicial confusion about whether pure Securities Act claims are removable from state to federal court. This part will explore the three approaches courts use in interpreting the effect that SLUSA and the PSLRA have had on the anti-removal and concurrent jurisdiction provisions of the Securities Act. Part III argues that an interpretation that allows for the removal of pure Securities Act claims not only serves the goals of the PSLRA and SLUSA, but also furthers the goals of the Securities Act. This Part concludes that innocent investors should be protected from the damage caused by meritless class actions.

\(^{36}\) See id. at 166.


\(^{38}\) See id.

I. STATUTORY BACKGROUND

A. General Removal Jurisdiction

To understand the jurisdiction of federal securities law, one must begin with an understanding of basic federal jurisdiction.\(^{40}\) As a general rule, “state and federal courts have concurrent jurisdiction to hear most cases that fall within Article III.”\(^{41}\) This rule gives plaintiffs the initial choice of forum, and a plaintiff may choose to bring a federal claim in state court.\(^{42}\) While plaintiffs get the first choice, defendants are not without power of their own. Removal jurisdiction is the mechanism by which defendants can transfer a federal claim from state court to federal court.\(^{43}\) But the power of removal is not absolute. Under the general removal statute, 28 U.S.C. 1441(a), a defendant may remove claims “[e]xcept as otherwise expressly provided by Act of Congress.”\(^{44}\) As noted by one court, this “‘except’ provision is ‘clearly a reference to statutes such as the [Securities Act].’”\(^{45}\)

B. Securities Act of 1933

Congress expressly provided for such an exception to removal jurisdiction when it passed the Securities Act\(^{46}\) in response to the Crash of 1929.\(^{47}\) Securities fraud plagued the market in the 1920s and, despite many state securities

\(^{40}\) Federal jurisdiction “extend[s] to all Cases, in Law and Equity, arising under this Constitution, the Laws of the United States, and Treaties made, or which shall be made, under their Authority . . . .” U.S. CONST. art. III, § 2.

\(^{41}\) Ernest A. Young, Dual Federalism, Concurrent Jurisdiction, and the Foreign Affairs Exception, 69 GEO. WASH. L. REV. 139, 145 (2001). The model of concurrent regulatory authority provide for removal statutes, which “create a right for some parties to be in federal court if they want to be, and most areas of concurrent jurisdiction exist at Congress’s sufferance.” Id.


\(^{43}\) 28 U.S.C. § 1441(a) (2012) (“Except as otherwise expressly provided by Act of Congress, any civil action brought in a State court of which the district courts of the United States have original jurisdiction, may be removed by the defendant or the defendants, to the district court of the United States for the district and division embracing the place where such action is pending.”).

\(^{44}\) See id.; see also Costa, supra note 42, at 1198 (“Congress has reserved the right to circumscribe removal in all cases in which they chose to do so in . . . 28 U.S.C. § 1441.”).


\(^{47}\) Serota, supra note 20, at 164.
statutes, “the public sustained severe losses at the hands of securities dealers and corporations.”\textsuperscript{48} With the clear need to restore investor confidence in the securities market in mind, President Roosevelt pushed for massive securities reform and the creation of the Securities Act.\textsuperscript{49} In a comprehensive effort to regulate securities markets, the objective of the Securities Act was to require companies to provide “full and fair disclosure . . . by filing a registration statement with the Securities and Exchange Commission.”\textsuperscript{50} The Securities Act also created an express private right of action to “hold[ ] those who file the statements liable for any misstatements or omissions.”\textsuperscript{51}

Within this private right of action, Congress chose to enhance investor protection by giving class action plaintiffs the ultimate power to choose the forum.\textsuperscript{52} Section 22(a) of the Securities Act contains two mechanisms to accomplish that goal.\textsuperscript{53} The section’s concurrent jurisdiction provision provides for concurrent state and federal jurisdiction,\textsuperscript{54} and its anti-removal provision eliminates the defendant’s ability to remove certain actions brought in state court.\textsuperscript{55} The effect of these provisions is generally recognized as being pro-plaintiff.\textsuperscript{56}

But the Securities Act is limited to regulating the initial distribution of securities.\textsuperscript{57} One year after its passage, Congress enacted the Exchange Act to address a different type of harm: intentional misconduct in the manipulation of stock prices and trading of securities.\textsuperscript{58} Noting the concurrent jurisdiction

\textsuperscript{48} Mazzeo, supra note 19, at 1440-42. \\
\textsuperscript{49} Id. at 1441. \\
\textsuperscript{50} Serota, supra note 20, at 164. The belief behind the Securities Act was that, “if companies and their underwriters were required to disclose all information to investors, shady deals would be impossible.” Poser, supra note 12, at 3. \\
\textsuperscript{51} Mazzeo, supra note 19, at 1442. \\
\textsuperscript{52} Serota, supra note 20, at 164. \\
\textsuperscript{54} See Mitchell A. Lowenthal & Timothy M. Haggerty, Jurisdictional Struggle Continues over 1933 Act Class Suits, N.Y. L.J., June 14, 2010, at S4, available at http://www.newyorklawjournal.com/PubArticleNY.jsp?id=1202462634140&Jurisdictional_Struggle_Con tinues_Over_1933_Act_Class_Suits&sreturn=20120825180339 [hereinafter Lowenthal & Haggerty, Jurisdictional Struggle]; see also 15 U.S.C. § 77v(a) (“The district courts of the United States and United States courts of any Territory, shall have jurisdiction . . . concurrent with State and Territorial courts . . . of all suits . . . brought to enforce any liability or duty created by this subchapter.”). \\
\textsuperscript{55} See Lowenthal & Haggerty, Jurisdictional Struggle, supra note 54, at S4 see also 15 U.S.C. § 77v(a) (“[N]o case arising under this subchapter and brought in any State court of competent jurisdiction shall be removed to any court of the United States.”). \\
\textsuperscript{56} See Mazzeo, supra note 19, at 1444. \\
\textsuperscript{57} See Snyder, Jr., supra note 45, at 673. \\
\textsuperscript{58} See Mazzeo, supra note 19, at 1443 n.67 (citing Ernst & Ernst v. Hochfelder, 425 U.S. 185, 194-95 (1976)); see also Snyder, Jr., supra note 45, at 673 (“The Exchange Act allows for the regulation of the secondary market and regulates all
provided by the Securities Act, Congress granted federal courts exclusive federal jurisdiction to hear claims arising under the Exchange Act.\(^59\) Congress also considered amending the Securities Act to provide exclusive federal jurisdiction, “but expressly declined to do so”\(^60\) for reasons that are still unclear.\(^61\)

C. Private Securities Litigation Reform Act of 1995 (PSLRA)

The anti-removal provisions of the Securities Act would prove to be particularly important after the passage of the PSLRA. Congress enacted the PSLRA in the face of a growing number of “‘strike suits[,]’ the ‘meritless class actions that allege fraud in the sale of securities.’”\(^62\) Title I of the PSLRA, titled “Reduction of Abusive Litigation,” added Section 27 to the Securities Act\(^63\) and it contained “some of the most sweeping amendments since the inception of federal securities law.”\(^64\) Most notably, Section 27 includes an automatic stay of discovery upon the filing of a motion to dismiss, as well as heightening the pleading standards for plaintiffs.\(^65\) The goal of the PSLRA was to


\(^60\) See Mazzeo, supra note 19, at 1444 (citing 78 CONG. REC. 8571 (1934) (statement of Sen. Byrnes) (noting that the Senate’s version of the bill provided for concurrent jurisdiction, while the House version of the bill granted exclusive federal jurisdiction)).

\(^61\) See Snyder, Jr., supra note 45, at 673.

\(^62\) Mitchell A. Lowenthal & Timothy M. Haggerty, SLUSA's Elimination of State and Court Jurisdiction over Securities Class Actions, CLEARY GOTTLIEB STEEN & HAMILTON LLP LITIGATION & ARBITRATION REPORT 20-21 (Dec. 2006) [hereinafter Lowenthal & Haggerty, SLUSA's Elimination].


\(^65\) Id. at 335 (footnotes omitted). The heightened pleading standard requires a plaintiff to “state with particularity facts giving rise to a strong inference that the defendant acted with the required state of mind.” 15 U.S.C. § 78u-4(b)(2) (2012) (emphasis added). The automatic stay of discovery provides:

In any private action arising under this subchapter, all discovery and other proceedings shall be stayed during the pendency of any motion to dismiss, unless the court finds, upon the motion of any party, that particularized discovery is necessary to preserve evidence or to prevent undue prejudice to that party.

\(^{Id.} \) § 77z-1(b)(1).

The PSLRA also includes several other important reforms to federal securities laws. A “lead plaintiff provision” rejects the race-to-the-courthouse method of
“provide protection . . . for investors, issuers, and all who are associated with the American capital markets.”\footnote{66} With the PSLRA, Congress reacted to concerns in the business community that class-action plaintiffs were abusing the anti-fraud provisions of securities laws.\footnote{67} “[W]henever a company’s stock price declined, plaintiffs rushed to file class actions under the federal securities laws, even though there was, in fact, no evidence of fraud at the time of the suit.”\footnote{68} Complaints generally charged deep-pocket defendants with cookie-cutter violations, such as allegedly issuing misleading public documents, thereby causing the price of the security to artificially climb or decline.\footnote{69} A cost-benefit analysis often forced the defendants to “settle even non-meritorious actions because the settlement amount would cost the defendant less than litigation expenses associated with discovery requests.”\footnote{70} The PSLRA was enacted under the belief that “both investors and the national economy suffer when innocent parties are forced to pay exorbitant settlements in meritless lawsuits.”\footnote{71} The PSLRA sought to curtail these strike suits through a series of procedural and jurisdictional reforms, measures intended to make it more difficult to bring private securities-fraud actions under federal securities laws.\footnote{72} Congress hoped that the increased cost and difficulty in bringing these actions would “weed out non-meritorious actions at the pleading stage, thereby discouraging strike suits.”\footnote{73} But the PSLRA had identifying the plaintiff, and instead presumes that the largest shareholder is the proper plaintiff, and is “entitled to control the private class action (and therefore appoint counsel) . . . .” O’Hare, supra note 64, at 336; see also 15 U.S.C. § 77z-1(a)(3) (2012). The court is also required to undertake an inquiry at the conclusion of each case to determine and impose mandatory sanctions on counsel for violations of Rule 11. See id. § 77z-1(e).
\footnote{66} Costa, supra note 42, at 1201 (citing H.R. REP. No. 104-369, at 32 (1995)) (internal quotation marks omitted).
\footnote{67} O’Hare, supra note 64, at 334.
\footnote{68} Id.
\footnote{69} Id. at 335; see also Mazzeo, supra note 19, at 1446 (“The systematic ‘abusive’ practices . . . embodied several characteristics: (1) routinely filing frivolous suits alleging cookie-cutter violations of the federal securities laws whenever there was a significant change in stock price, (2) targeting deep-pocketed defendants, and (3) abusing discovery practices in the hopes that the defendant would make a quick and sizeable settlement in order to avoid the expense of litigation.”).\footnote{70} O’Hare, supra note 64, at 335.
\footnote{71} Mazzeo, supra note 19, at 1446 (internal quotation marks omitted).
\footnote{72} O’Hare, supra note 64, at 335.
\footnote{73} Id. at 336. Additionally, “Congress hoped . . . [to] encourage[] companies to make projections, forecasts, and other kinds of forward-looking statements.” Id. Companies could expect to be hit by a lawsuit alleging fraud whenever a forward-looking projection or statement failed to materialize. Id. Accordingly, businesses were discouraged from making such statements. Id. In addition to discouraging strike suits
unintended consequences that would play a large role in shaping the field of federal securities laws. Most significantly, there was a large increase in the number of securities class actions filed in state court.\textsuperscript{74} After all, “[t]he restrictions added by [the] PSLRA . . . appl\{ied\} only to claims brought in federal court . . . .”\textsuperscript{75} Due to the concurrent jurisdiction of the Securities Act, plaintiffs were able to strategically “exploit a jurisdictional loophole” by bringing suit under state law and in state court.\textsuperscript{76} In attempts to circumvent the more stringent procedural requirements imposed by the PSLRA, such lawsuits were routinely filed contemporaneously with claims filed in federal court.\textsuperscript{77}

The Exchange Act did not fare much better at avoiding strike suits than the Securities Act. While the Exchange Act’s exclusive federal jurisdiction prevented plaintiffs from filing Exchange Act claims in state court, many states’ securities laws, as well as common law fraud, provided remedies similar to those under the Exchange Act.\textsuperscript{78} Thus, by filing state law claims in state court, plaintiffs simply avoided federal securities laws entirely.

\textsuperscript{74} Mazzeo, supra note 19, at 1448 (citing Office of the Gen. Counsel, United States Sec. & Exch. Comm’n, Report to the President and the Congress on the First Year of Practice Under the Private Securities Litigation Reform Act of 1995 (1997), available at http://www.sec.gov/news/studies/rereform.txt.). The SEC Report was conducted at the request of President Clinton, “in an effort to measure the level of success the [PSLRA] had achieved in attaining its aforementioned goals.” Costa, supra note 42, at 1201. The PSLRA did achieve its goal of decreasing the volume of securities class actions in federal court. Id. at 1202. There was also an “increased delay between the release of adverse information . . . and the filing of the action,” evidencing “greater research and investigation” . . . and the filing of the action, the “conclusion that . . . meritorious claims were still developed and brought, [while] frivolous claims were not.” Id. at 1202-03.

\textsuperscript{75} Snyder, Jr., supra note 45, at 675.

\textsuperscript{76} Serota, supra note 20, at 168. Under the Securities Act of 1933, class-action plaintiffs could generally avoid the heightened requirements of the PSLRA in two ways: first, plaintiffs could file claims arising under the Securities Act of 1933 directly in state court, where they enjoyed concurrent and non-removable jurisdiction; second, plaintiffs could file claims in state court based on a state law theory of securities or common law fraud. See Snyder, Jr., supra note 45, at 676.

\textsuperscript{77} See Mazzeo, supra note 19, at 1449. In particular, plaintiffs could take advantage of more lenient discovery in state court, and use those facts to withstand a motion to dismiss. Id. Traditionally, state courts did not provide remedies as broad as federal remedies for securities fraud, but “other advantages [include] nonunanimous jury verdicts, punitive damages, and aiding and abetting liability.” Id.

\textsuperscript{78} See Snyder, Jr., supra note 45, at 676 (citing 2 Thomas Lee Hazen, Treatise on the Law of Securities Regulation § 7.17[2] (5th ed. 2005)).
D. Securities Litigation Uniform Standards Act of 1998 (SLUSA)

Congress concluded that the concurrent jurisdiction and anti-removal provisions of the Securities Act prevented the PSLRA from accomplishing its goals. 79 In 1998, Congress passed SLUSA. 80 "SLUSA sought to cure this infirmity by enacting ‘national standards’ for securities class actions involving ‘nationally traded securities.’" 81 SLUSA amended the Securities Act in order to make federal court the "primary venue for securities fraud class actions." 82 The ultimate result of the SLUSA amendment was to allow a defendant to remove certain preempted claims sounding in state and common law to federal court, where the claim would be dismissed.

A preempted claim under SLUSA is a "covered class action based upon the statutory or common law of any State" alleging a misrepresentation or use of a manipulative or deceptive device in connection with the purchase or sale of a covered security. 83 Preemption only applies to "[c]overed securities," securities "listed, or authorized for listing" on a national exchange such as the New York Stock Exchange or NASDAQ. 84 A preempted claim must also be a "covered class action," which is of a form similar but "not identical to a class action brought under Rule 23 of the Federal Rules of Civil Procedure." 85 If the claim is preempted, the class action cannot "be maintained in any State or Federal court by any private party." 86

79 See Lowenthal & Haggerty, Jurisdictional Struggle, supra note 54, at S4.
81 Mazzeo, supra note 19, at 1450.
82 See Snyder, Jr., supra note 45, at 677.
84 See 15 U.S.C. § 77r(b) (defining “covered security”); see also Mazzeo, supra note 19, at 1450. (“Thus, securities traded over-the-counter on the Nasdaq Small Capital Market are not covered securities, and anti-fraud actions for these securities are not preempted.”); O’Hare, supra note 64, at 339-40.
85 O’Hare, supra note 64, at 340. A covered class action includes: “(1) actions brought on behalf of more than 50 persons, (2) actions brought on a representative basis, and (3) a group of joined or consolidated actions.” Id.; see also 15 U.S.C. § 77p(0)(2); 15 U.S.C. § 78bb(f)(3)(B).
86 See 15 U.S.C. §§ 77p(b), 78bb(f)(1); see also Costa, supra note 42, at 1205.
In order to facilitate bringing preempted claims to federal court for dismissal, SLUSA added corresponding removal provisions.\(^87\) SLUSA amended Section 22(a) of the Securities Act to add for an exception to the removal bar, “as provided in [Section 16(c)].”\(^88\) Section 16(c) provides that “[a]ny covered class action brought in any State court involving a covered security, as set forth in subsection (b) of this section, shall be removable to the Federal district court . . . and shall be subject to subsection (b) of this section.”\(^89\) SLUSA also amended the concurrent jurisdiction provision of the Securities Act so that state court jurisdiction is still concurrent with federal jurisdiction “except with respect to . . . ‘covered class actions.’”\(^90\)

Supporters of this legislation argued that it was necessary for two reasons. First, preemption of state law-based securities fraud class actions, for certain nationally traded securities, was necessary to ensure the effectiveness of the PSLRA. Second, continued circumvention of the PSLRA could potentially be “exacerbated by a projected race-to-the-bottom, in which one or more states enact laws decidedly more favorable to plaintiffs than federal law.”\(^91\) The race-to-the-bottom would be further exaggerated because “companies with publicly traded securities cannot control where their securities are traded after an initial public offering; thus, issuers . . . cannot choose to avoid jurisdictions that present unreasonable litigation costs.”\(^92\)

II. DISTRICT COURT INTERPRETATIONS

“[T]he SLUSA amendment is hardly a model of clarity.”\(^93\) There has been considerable judicial confusion about the removability of pure Securities Act claims filed in state court. Despite SLUSA’s amendment and the attempt to create a uniform federal standard, “plaintiffs have continued to bring class claims under the Securities Act in state courts.”\(^94\) These plaintiffs argue that “SLUSA [has] failed to capture all securities class actions” because, by its plain language, SLUSA

---

90 15 U.S.C. § 77v(a); see also Schacter & Gearns, supra note 83.
91 See Costa, supra note 42, at 1204-05 (internal quotation marks omitted).
92 Mazzeo, supra note 19, at 1451.
93 Id.
94 Lowenthal & Haggarty, SLUSA’s Elimination, supra note 62, at 20.
does not apply to class actions that “rais[e] claims exclusively under the Securities Act (without any state law claims).”

Advocating a narrow interpretation, these plaintiffs assert that removal is limited to “covered class action[s] based upon the statutory or common law of any State.” Accordingly, removal does not apply to federal claims—those that arise under the Securities Act—and thus, defendants continue to be prejudiced by the circumvention of the PSLRA’s heightened federal standards.

District courts principally disagree about whether the amendment made by SLUSA limits removal to preempted state law securities class actions, or whether the amendment also allows for removal of claims arising solely out of the Securities Act.

Thus far, federal district courts have reached inconsistent results. Initial disagreement rested upon conflicting interpretations of SLUSA’s amendment to the anti-removal provision of the Securities Act. Some courts took a narrow approach and looked only to the “plain meaning” of the Securities Act. Other courts, noting an inconsistency between the removal and pre-emption provisions, took a broad approach and looked to legislative intent to aid in deciphering the scope of removal. However, an emerging trend has been to ignore SLUSA’s amendment to the anti-removal provision and instead look to SLUSA’s amendment to the concurrent jurisdiction provision of the Securities Act. These courts purport to align the “plain meaning” of the text with the legislative intent to create a uniform federal standard.

95 Id.
97 Snyder, Jr., supra note 45, at 670.
98 Schacter & Gearn, supra note 83.
A. Courts Denying Removal: The Narrow Interpretation of SLUSA’s Amendment to the Anti-Removal Provision

Courts that have taken a narrow approach in interpreting SLUSA’s amendment to the Securities Act focus on the effect of the “three cross-referencing provisions of the act” and, in doing so, purport to interpret the “plain meaning” of the statute.\footnote{Lowenthal & Haggarty, Jurisdictional Struggle, supra note 54, at S4. See, e.g., In re Waste Mgmt., Inc., 194 F. Supp. 2d at 591 (remanding a case alleging only violations of the Securities Act of 1933. The Court laid out a five-part test for removal of a claim under SLUSA: “(1) [that] the action is a ‘covered class action’ under SLUSA; (2) that the causes of action on their face are based on state statutory or common law; (3) that it involves a ‘covered security’ under SLUSA; (4) that it alleges defendants misrepresented or omitted material facts; and (5) that the alleged misrepresentation or omission was made ‘in connection with’ the purchase or sale of the covered security”); Irra, 2006 WL 2375472 (holding the plaintiff’s claims based solely on the Securities Act were not removable under SLUSA); Pipefitters Local 522 and 633 Pension Trust Fund, U.S. Dist. LEXIS 14202, at *6-7; Haw. Structural Ironworkers Pension Trust Fund, 2003 U.S. Dist. Lexis 15832, at *5-6 (finding that the plain language of 77p(c) limits removal to class actions based on State claims); Nauheim, No. 02-C-9211, 2003 WL 1888843, at *11 (limiting removal to class actions complaints based on State statutory or common law).} “It is well settled that courts interpreting a statute should ‘give effect, if possible, to every clause and word of the statute’.”\footnote{Costa, supra note 42, at 1213 (quoting Gade v. Nat’l Solid Wastes Mgmt. Ass’n, 505 U.S. 88, 100 (1992)).} Thus, courts adopting a narrow approach ultimately conclude that pure Securities Act claims cannot be removed from the state court in which they are brought.

Courts begin this statutory interpretation by looking to Section 22(a) of the Securities Act. Section 22(a) provides for an exception to the anti-removal provision in Section 16(c).\footnote{15 U.S.C. § 77v(a) (2012) (“Except as provided in section 77p(c) of this title, no case arising under this subchapter and brought in any State court of competent jurisdiction shall be removed to any court of the United States.”).} Section 16(c) excepts from the removal bar “[a]ny covered class action brought in State court involving a covered security, as set forth in subsection (b). . . .”\footnote{Securities Litigation Uniform Standards Act of 1998 § 101, 112 Stat. 3227, 3228 (codified as amended at 15 U.S.C. § 77p(c) (2012)) (emphasis added) (“Any covered class action brought in any State court involving a covered security, as set forth in subsection (b), shall be removable to the Federal district court for the district in which the action is pending, and shall be subject to subsection (b).”).} Finally, subsection (b), the preclusion provision, provides for preclusion from federal court of certain class actions “based upon the statutory or common law of any State or subdivision thereof . . . .”\footnote{Securities Litigation Uniform Standards Act of 1998 § 101, 112 Stat. 3227, 3228 (codified as amended at 15 U.S.C. § 77p(b) (2012)) (emphasis added) (“No covered class action based upon the statutory or common law of any State or subdivision thereof may be maintained in any State or Federal court by any private party alleging—(1) an untrue statement or omission of a material fact in connection with the purchase or sale of a covered security; (2) a breach of any duty arising under this subchapter; (3) the duty to make a prospectus or registration statement available to any person; (4) the duty to prepare, keep current, and make available to any person, any registration statement required to be filed; (5) a knowing furnishing to the Securities and Exchange Commission of a false report or statement to the SEC; (6) the liability of an issuer to any underwriter that makes a purchase of securities from the issuer; (7) the liability of an underwriter to any person who purchases securities from the underwriter at the underwriter’s direct offer price; (8) the liability of any person who disseminates or distributes any report, document, or communication of any kind that contains any of the material misrepresentations or omissions described in subsection (a); (9) the liability of any person who disseminates or distributes any report, document, or communication of any kind that is similar to, but not identical to, a report, document, or communication described in subsection (a); (10) the liability of any person who sells to an underwriter a security at the initial public offering price or other price for which the underwriter becomes a dealer only; (11) the liability of any person who, as an underwriter, lends or lends a security that the underwriter has agreed to repurchase, unless the person can establish by clear and convincing evidence that the underwriter did not knowingly receive or contribute any material misrepresentations or omissions described in subsection (a); (12) the liability of any person who, in connection with the purchase or sale of a security, sells any of the securities included in a registration statement or any of the securities registered under section 5(b) of this title; or (13) the liability of a person for a violation of any provision of this title.”).} Because Section
16(c) specifically cites to the preclusion provision of subsection (b), some courts have found that these three provisions must be coextensive and that “only actions precluded under [subsection (b)] are removable under [Section 16(c)], and the actions removable under [Section 16(c)] are the only actions removable under [Section 22(a)].” Put another way, “if an action is not preempted by [subsection (b)], then it may not be removed under Section 22(a).” But if the claim is an “entirely preempted state law action” or a Securities Act claim coupled with a preempted state law claim, then it may properly be removed under this interpretation.

This approach seems to have gained some support from the United States Supreme Court in its decision in *Kircher v. Putnam Funds Trust*. The Court, in rather “emphatic and expansive . . . dicta,” stated that there was “no reason to reject the straightforward reading: removal and jurisdiction to deal with removed cases is limited to those precluded by the terms of subsection (b).” The Court continued, “If the action is not precluded [under subsection (b)], the federal court likewise has no jurisdiction to touch the case on the merits, and the proper course is to remand to the state court . . . .” However, covered security; or (2) that the defendant used or employed any manipulative or deceptive device or contrivance in connection with the purchase or sale of a covered security.”

---

106 See Serota, supra note 20, at 169-70.
107 Lowenthal & Haggarty, *Jurisdictional Struggle*, supra note 54, at S4 (internal quotation marks omitted); see also Lowenthal & Haggarty, *SLUSA’s Elimination*, supra note 62, at 23.
109 See id.
110 *Kircher v. Putnam Funds Trust*, 547 U.S. 633 (2006). The issue in this case was whether the Seventh Circuit had jurisdiction to review a district court’s decision to remand a class action securities claim, based on state law, back to state court because the district court found that it lacked subject matter jurisdiction. *See Kircher v. Putnam Funds Trust*, No. 03-CV-0691, 2004 U.S. Dist. LEXIS 10327, at *31-32 (S.D. Ill. Jan. 27, 2004). The Supreme Court then tackled the issue of whether holders of mutual fund shares were within the section of SLUSA that “allows for the removal of claims that arise ‘in connection with the purchase or sale of a covered security.’” *See J. Tyler Butts, Removal of Covered Class Actions Under SLUSA: The Failure of Plain Meaning and Legislative Intent as Interpretive Devices, and the Supreme Court’s Decisive Solution*, 1 WM. & MARY BUS. L. REV. 169, 189-90 (2010). The ultimate holding was “procedural, not substantive, when it held that under federal law, the case should never have been heard on appeal at the federal level.” *Id.* at 190. But after the “ultimate issue . . . had been decided, Justice Souter elaborated on the scope of SLUSA and removal generally.” *Id.; see also Kircher v. Putnam Funds Trust*, 547 U.S. 633, 646-48 (2006).
111 Lowenthal & Haggarty, *Jurisdictional Struggle*, supra note 54, at S5 (internal quotation marks omitted).
112 *Kircher*, 547 U.S. at 643; *see also* Lowenthal & Haggarty, *Jurisdictional Struggle*, supra note 54, at S5.
113 *Kircher*, 547 U.S. at 644. It must also be noted that the Supreme Court was dealing specifically with the “covered security” provision under subsection (b)(1), and
the Supreme Court was faced with a state law claim, not a pure Securities Act claim that district courts currently struggle with, and perhaps “did not fully explain its rationale.”

Critics point out that this interpretation of the “inartfully drawn” statute “renders the exception to [Section] 22(a)’s anti-removal provision unnecessary, meaningless and not an exception at all.” Section 22(a) creates a general rule against the removal of cases ‘arising under’ the Securities Act, but permits an exception ‘as provided in Section 16(c).” “[T]o have meaning, [this section] must apply to some subset of cases that actually arise under the Securities Act,” but a narrow interpretation of SLUSA would limit removal to preempted claims under subsection (b), i.e., claims arising under state law. Noting this apparent inconsistency, these critics reason that state law claims obviously cannot “arise under” the Securities Act but rather, “arise under state law.” The amendments to Section 16(c) and subsection (b), standing alone, would accomplish the same result as the narrow interpretation. If Congress did in fact intend to limit removability to state law claims, SLUSA’s amendment to Section 22(a) would be superfluous and unnecessary.

Not to subsection (b) as a whole, which includes the state law limitation provision. See 15 U.S.C. § 77p(b)(1) (2012) (“in connection with the purchase or sale of a covered security”); see also 15 U.C. § 77p(b) (“based upon the statutory or common law of any State or subdivision thereof”). However, it would be internally inconsistent to allow one provision under subsection (b) to prohibit removal, while not applying the same interpretation to another provision (the state law limitation) within the same subsection. See Butts, supra note 110, at 192.

The dicta could fairly be interpreted to suggest that “if a claim fails to meet any one of the requirements of subsection (b), removal is not an option.” Id. This is in line with the narrow approach, and if the claim is not based on state law, then it is not removable under SLUSA. This interpretation was adopted by the District Court for the Northern District of Georgia in Unschuld, which assumed that “[p]resumably, the [Supreme Court] was aware of the ongoing dispute about removal of such claims.” Unschuld v. Tri-S Sec. Corp., No. 06-02931, 2007 U.S. Dist. LEXIS 68513, at *34 (N.D. Ga. Sept. 14, 2007).

114 See Snyder, Jr., supra note 45, at 693.
115 Schacter & Gears, supra note 83.
116 See Lowenthal & Haggarty, Jurisdictional Struggle, supra note 54, at S4-S5.
118 See Lowenthal & Haggarty, SLUSA’s Elimination, supra note 62, at 23.
119 See id.
120 See Lowenthal & Haggarty, Jurisdictional Struggle, supra note 54, at S5.
121 See Lowenthal & Haggarty, SLUSA’s Elimination, supra note 62, at 23.
122 See id. This argument was directly addressed in Nauheim. The defendants argued against removal, but the court found that the language was “made meaningful by . . . 77p(c)’s preemption of an expressly delineated category of state law class actions.” Nauheim v. Interpublic Grp. of Cos., No. 02-C-9211, 2003 WL 1888843, at *5 (N.D. Ill. Apr. 16, 2003). As one author puts it, however, this assertion is “simply wrong,” and the court “entirely fail[s] to justify this conclusion.” Costa, supra note 42, at 1210.
Accordingly, a broader interpretation that includes removal for pure Securities Act claims gives meaning to the word “except” in Section 22(a). The courts’ narrow interpretation has led to the “somewhat bizarre and anomalous” result of putting state law claims in federal court but leaving federal claims in state court.

However, as the Supreme Court has long acknowledged, “when the words of the statute are unambiguous[,] judicial inquiry is complete.” It is the “sole function of the courts,” when presented with a law containing unambiguous language, “to enforce it according to its terms.” When a statute can be interpreted on its face, resorting to legislative history is inappropriate, and the court must presume the statute means what it says. Courts taking a narrow interpretation often merely recite the words of the statute, followed by “conclusory summation[s],” as though the meaning is “so obvious that it needs no more explanation.” But such an unambiguous statute simply does not exist, or else courts would not be confronted with such widespread divergence of interpretation.

B. Courts Upholding Removal

1. The Broad Interpretation of SLUSA’s Amendment to the Anti-Removal Provision

The courts that reject the “plain meaning” approach apply a much broader reading to determine whether SLUSA
permits removal of pure Securities Act claims.\textsuperscript{129} They typically look beyond the text and focus on the purpose of SLUSA to supplement their interpretation of the otherwise confusing statute.\textsuperscript{130} Accordingly, these courts believe that their interpretation is in line with Congress’s decision to “remedy the PSLRA’s failure to ‘prevent abuses in private securities fraud lawsuits’” and Congress’s perceived need to enact “national standards for securities class action lawsuits.”\textsuperscript{131} Thus, courts adopting a broad approach ultimately conclude that pure Securities Act claims are removable from state court.

“When statutory language is open to more than one reasonable interpretation, courts attempt to find the meaning ‘which can most fairly be said to be imbedded in the statute, in the sense of being most harmonious with its scheme and with the general purposes that Congress manifested.’”\textsuperscript{132} Congress enacted the PSLRA to curtail “strike suits,” but it failed to achieve its goal because plaintiffs simply avoided filing in federal court.\textsuperscript{133} Congress subsequently enacted SLUSA\textsuperscript{134} “to realize the intent of the [PSLRA] . . . [and ensure] that class action suits for securities that are traded on the . . . major securities trading


\textsuperscript{130} Lowenthal & Haggerty, \textit{SLUSA’s Elimination}, supra note 62, at 24; see also TXU Corp., No. 3:02-CV-2243-K, 2003 U.S. Dist. LEXIS 7900, at *4-6 (where the court looked to legislative findings to support its interpretation of SLUSA).

\textsuperscript{131} See Serota, supra note 20, at 170 (citing \textit{Brody}, 240 F. Supp. 2d at 1124) (internal quotation marks omitted). Congress expressly set forth its goal in enacting SLUSA under the “Findings” section of the Act:

The Congress finds that . . . in order to prevent certain State private securities class action lawsuits alleging fraud from being used to frustrate the objectives of the Private Securities Litigation Reform Act of 1995, it is appropriate to enact national standards for securities class action lawsuits involving nationally traded securities, while preserving the appropriate enforcement powers of State securities regulators and not changing the current treatment of individual lawsuits.

\textsuperscript{132} Butts, supra note 110, at 186 (quoting Comm’r v. Engle, 464 U.S. 206, 217 (1984)).

\textsuperscript{133} Lowenthal & Haggerty, \textit{SLUSA’s Elimination}, supra note 62, at 21.

\textsuperscript{134} Thirteen senators co-introduced S. 1260, later enacted as SLUSA, on October 7, 1997. \textit{See} 143 CONG. REC. S10475 (1997) (statements of Sen. Gramm).
exchanges . . . [are] subject to the rules that we passed last time [in the PSLRA] and . . . go to federal court.”

Senator Phil Gramm of Texas, one of the two chief co-sponsors of SLUSA, stated that “in the case of class-action suits, . . . if a stock is traded on the national market . . . then the class action-suit has to be filed in federal court.” The other chief co-sponsor, Senator Chris Dodd of Connecticut has commented that the “one development . . . that has the potential to undermine our good work and send us back to the days of litigation frenzy . . . is the significant increase in securities fraud class actions filed in State court.” Senator Dodd was also very concerned about having to subject foreign companies to not only “very tough Federal standards on securities fraud, but also the possibility of 50 constantly changing State standards.” In the view of courts interpreting SLUSA broadly, certain statements by some members of the legislature “prove conclusively that Congress meant to remove all securities claims from State court, and simply fell victim to sloppy or misleading drafting.” These legislator’s comments reflect the view that the problem under the PSLRA was not that plaintiffs were bringing state law claims, but rather that those claims were being brought in state court. Congress enacted SLUSA in response to the perceived failings of the PSLRA, and, accordingly, “the goals of each are inextricably intertwined . . .” Thus, some argue, the only way to prevent circumvention of the PSLRA by plaintiffs asserting pure Securities Act claims is to interpret the SLUSA’s removal amendment broadly. As Thomas Bliley, Jr., U.S. Representative from Virginia, explained before the House of Representatives, “The premise of this legislation is simple: lawsuits alleging violations that involve securities that are offered nationally belong to Federal court.” It is clear that at least some

137 Id. § 10476 (comments of Sen. Dodd).
138 Butts, supra note 110, at 188.
139 See Lowenthal & Haggerty, SLUSA’s Elimination, supra note 62, at 24.
140 Snyder, Jr., supra note 45, at 696.
141 See id., at 697.
142 144 CONG. REC. S11020 (1998) (comments of Rep. Bliley). Some courts have found this quote to assist in deciphering the text and purpose of the statute. See Butts, supra note 110, at 188 (citing Brody v. Homestore, Inc., 240 F. Supp. 2d 1122, 1124 (C.D. Cal. 2003)).
members of Congress sought to create federal uniformity and eliminate class action claims in state court.\textsuperscript{144}

The Supreme Court offered its own guidance favoring a broad interpretation in \textit{Merrill Lynch, Pierce, Fenner, \& Smith Inc. v. Dabit}.\textsuperscript{145} The Court specifically addressed whether SLUSA precluded state law holder claims, but language in the opinion “suggests that the Court would support the broad reading of [Section 16(c)] and extend removal authority to [pure] Securities Act claims.”\textsuperscript{146} The Court “purported to follow congressional intent”\textsuperscript{147} and noted the “congressional preference for national standards for securities class action lawsuits.”\textsuperscript{148}

The Court stated that the “magnitude of the federal interest in protecting the integrity and efficient operation of the market for nationally traded securities cannot be overstated.”\textsuperscript{149} The Court also recognized that the purpose of SLUSA was “[t]o stem this shift from Federal to State courts and prevent certain State private securities class action lawsuits alleging fraud from being used to frustrate the objectives of [the PSLRA].”\textsuperscript{150} The narrow interpretation would certainly undercut the effectiveness of the PSLRA by permitting pure Securities Act class actions to remain in state court. While it seems clear that the Supreme Court’s broad language in \textit{Dabit} and its narrow language in \textit{Kircher} are contradictory,\textsuperscript{151} “the Court’s broad policy language . . . may be more indicative of [its] view of securities laws [in general]” than its more narrow statutory analysis of a specific provision.\textsuperscript{152} Thus it is possible that the Court could “abandon its dicta in \textit{Kircher} and construe the

\textsuperscript{144} See Lowenthal \& Haggerty, \textit{SLUSA’s Elimination}, supra note 62, at 24.


\textsuperscript{146} See Snyder, Jr., supra note 45, at 688-89 (citing \textit{Dabit}, 547 U.S. at 88-89). On the ultimate issue, the Court refused to limit preclusion of claims “in connection with the purchase or sale” to only claims involving a “purchase or sale.” \textit{Dabit}, 547 U.S. at 89; see also Snyder, Jr., supra note 45, at 690.

\textsuperscript{147} Snyder, Jr., supra note 45, at 691. “[T]he Court also appeared comfortable allowing policy to dictate statutory interpretation in this area of the law . . . [because it] had previously done so in \textit{Blue Chip Stamps}.” \textit{Id.} (citing \textit{Dabit}, 547 U.S. at 80, 84). The Court in \textit{Dabit} noted that in \textit{Blue Chip Stamps}, a rule 10b-5 case, it had “relied chiefly, and candidly, on ‘policy considerations.’” \textit{Dabit}, 547 U.S. at 84 (citation omitted); see also Snyder, Jr., supra note 45, at 691.


\textsuperscript{149} \textit{Dabit}, 547 U.S. at 78; see also Snyder, Jr., supra note 45, at 691.

\textsuperscript{150} \textit{Dabit}, 547 U.S. at 82 (internal quotation marks omitted).

\textsuperscript{151} In \textit{Kircher}, the Supreme Court stated in dicta that “removal jurisdiction under subsection (c) is understood to be restricted to precluded actions defined by subsection (b) . . . .” \textit{Kircher v. Putnam Funds Trust}, 547 U.S. 633, 643-44 (2006).

\textsuperscript{152} See Snyder, Jr., supra note 45, at 693.
[anti]-removal provision . . . to cover [pure] Securities Act claims," especially when "[s]uch a reading is also more consistent with . . . congressional intent."153

Aside from statements by some legislators, opponents of this broad interpretation argue that actually effectuating Congress’s purported intent “could not have been simpler.”154 If Congress really intended for removal of pure Securities Act claims, it could have easily added a sentence clarifying as such. While sporadic statements can be read as authorizing removal of pure Securities Act claims, “the issue is never addressed nor supported directly.”155 Perhaps, the legislative body as a whole did not agree entirely with some of the more vocal members.156 In fact, proponents of a narrow interpretation can simply point to the first sentence of SLUSA, which states that SLUSA’s purpose is “to limit the conduct of securities class actions under State law.”157 Because they see the legislative history as “murky,”158 courts that interpret SLUSA narrowly find legislative history unreliable to the extent that any analysis into it would be superfluous.159 As Judge Alex Kozinski, Chief Judge of the U.S. Court of Appeals for the Ninth Circuit, has said, “Consulting

153 Id.
154 See Costa, supra note 42, at 1217.
155 See id. at 1220.
156 See id. at 1222. It is possible that the current legislation is merely a compromise, and “[h]ad SLUSA explicitly allowed removal of all class actions arising under the Securities Act, it may . . . have lacked the political support to pass the 105th Congress.” Id. at 1122-23. The 105th Congress was controlled by a Republican majority in both houses and at the time “the Republican majority was generally concerned with federalism, and with ‘returning authority to the states.’” Id. at 1123 (citing A.C. Pritchard, Constitutional Federalism, Individual Liberty, and the Securities Litigation Uniform Standards Act of 1998, 78 WASH. U. L.Q. 435, 435 (2000)). In the 105th Congress, the Senate was composed of 55 Republicans and 45 Democrats and the House was composed of 228 Republicans, 206 Democrats, and 1 Independent. See Costa, supra note 42, at 1223 n.203 (citing S. PUB. 105-20, at 2-3 (1997)).
158 See Butts, supra note 110, at 187.
159 See Butts, supra note 110, at 187.
legislative history is like 'looking over a crowd of people and picking out your friends.'

It must also be noted that SLUSA amended both the Securities Act and the Exchange Act in an "almost identical fashion" with respect to their clauses, precluding state law claims. Some interpret this symmetry to mean that Congress really must have intended for SLUSA to cover both state law claims and pure Securities Act claims. Because the Exchange Act granted exclusive federal jurisdiction to claims arising under it, the only method plaintiffs could use to circumvent the PSLRA was to allege state law claims in state court. Accordingly,

---

160 See Alex Kozinski, Should Reading Legislative History Be an Impeachable Offense?, 31 SUFFOLK U. L. REV. 807, 813 (1998). Judge Kozinski was quoting the words of Judge Harold Leventhal, who served on the District of Columbia Circuit Court of Appeals from 1965 to 1979; see also Butts, supra note 110, at 186 n.100.

161 Snyder, Jr., supra note 45, at 688. Section 16 of the Securities Act of 1933, 15 U.S.C. § 77p (b)-(c) (2012), provides:

(b) Class action limitations: No covered class action based upon the statutory or common law of any State or subdivision thereof may be maintained in any State or Federal court by any private party alleging:

(1) an untrue statement or omission of a material fact in connection with the purchase or sale of a covered security; or

(2) that the defendant used or employed any manipulative or deceptive device or contrivance in connection with the purchase or sale of a covered security.

(c) Removal of covered class actions: Any covered class action brought in any State court involving a covered security, as set forth in subsection (b), shall be removable to the Federal district court for the district in which the action is pending, and shall be subject to subsection (b).

Section 29(f) of the Securities Exchange Act of 1934, 15 U.S.C. § 78bb(f), provides:

(1) Class action limitations: No covered class action based upon the statutory or common law of any State or subdivision thereof may be maintained in any State or Federal court by any private party alleging:

(A) a misrepresentation or omission of a material fact in connection with the purchase or sale of a covered security; or

(B) that the defendant used or employed any manipulative or deceptive device or contrivance in connection with the purchase or sale of a covered security.

(2) Removal of covered class actions: Any covered class action brought in any State court involving a covered security, as set forth in paragraph (1), shall be removable to the Federal district court for the district in which the action is pending, and shall be subject to paragraph (1).

162 See Snyder, Jr., supra note 45, at 688-90 ("The broad reading . . . gives purpose to the amendment to the Securities Act by interpreting it to extend removal authority to [pure] Securities Act claims . . . .").

163 15 U.S.C. § 78aa ("The district courts of the United States and the United States courts of any Territory or other place subject to the jurisdiction of the United States shall have exclusive jurisdiction of violations of this chapter or the rules and regulations thereunder, and of all suits in equity and actions at law brought to enforce . . . .")
SLUSA’s amendment to the Exchange Act was perfectly tailored to allow removal of these state law claims into federal court, thus making it the exclusive venue for claims that, in substance, alleged violations of the Exchange Act. Because Congress intended all federal and state law securities fraud claims relating to the Exchange Act to be in federal court, it could also be argued that they intended the same result for the Securities Act—that both state law securities fraud claims and pure Securities Act claims should be removable to federal court.164

Yet this reasoning does not necessitate the conclusion that Congress intended federal court to be the exclusive forum for all class action securities fraud claims arising under the Securities Act or the Exchange Act. SLUSA’s amendment to the Exchange Act applies only to removal of state law claims because the Exchange Act lacks a concurrent jurisdiction provision, and hence, “pure” Exchange Act claims already enjoy exclusive federal jurisdiction.165 Because there was no need to address removal for pure Exchange Act claims, the anti-removal provision simply does not apply to them. A consistent application of almost the same statutory language lends support to the argument that SLUSA’s amendment to the anti-removal provision of the Securities Act also applies to claims solely alleging state law securities fraud.

2. SLUSA’s Amendment to the Concurrent Jurisdiction Provision

There is still the possibility that Congress intended to limit SLUSA’s removal amendment only to state law claims, as a narrow reading of the amendment would suggest, and yet still intended to make federal courts the exclusive venue for class action securities fraud claims. While much of the focus in recent years has been on the effect of SLUSA’s amendment to the Securities Act’s anti-removal provision, an emerging trend among district courts has been to focus on another one of SLUSA’s amendments: the amendment to the concurrent jurisdiction provision of the Securities Act.166 Although the end result is the same as the broad interpretation of SLUSA’s amendment to the anti-removal provision (that pure Securities Act claims are removable), courts adopting this emerging concurrent jurisdiction

---

164 See Snyder, Jr., supra note 45, at 688-90.
approach have found that this method is “rooted in both the text and the congressional findings underlying the SLUSA amendments.”167 Relying on the plain meaning of SLUSA’s amendment to the concurrent jurisdiction provision, these courts have concluded that state courts are no longer “court[s] of competent jurisdiction”168 for claims arising under the Securities Act. Therefore, the removal bar does not apply to them at all so that they can be removed just like any other federal claim.

While some courts had commented in dicta on SLUSA’s change to the concurrent jurisdiction provision of section 22(a),169 it was not until the District of New Jersey handed down two rulings in 2007 that a court explicitly found that SLUSA’s amendment to the concurrent jurisdiction provision, not the anti-removal provision, allowed for removal of pure Securities Act claims.170 Section 22(a) of the Securities Act provides for federal jurisdiction that is “concurrent with State and Territorial courts, except as provided in [Section 16] with respect to covered class actions, of all suits . . . brought to enforce any liability or duty created by this subchapter.”171 Thus, the amendment provides for an exception to the general rule of concurrent jurisdiction if the suit: (1) is brought to enforce the rights and liabilities created by the Securities Act, and (2) is a covered class action as provided in Section 16.172

The scope of the exception to concurrent jurisdiction can therefore be found in Section 16 of the Act.173 Subsections (b), (c), and (d) to this provision refer to state law claims and thus, by definition, are not “brought to enforce the rights and liabilities created by the Securities Act.”174 Subsection (f),

167 Serota, supra note 20, at 171.
170 In Rovner v. Vonage Holdings Corp., No. 07 Civ. 178, 2007 WL 446658 (D.N.J. Feb. 7, 2007), and Pinto v. Vonage Holding Corp., No. 07 Civ. 0062, 2007 WL 1381746 (D.N.J. Jan. 4, 2007), Judge Freda Wolfson was the first to focus on the jurisdictional amendment as the basis for denying remand, instead of the anti-removal provision. See Lowenthal & Haggerty, Jurisdictional Struggle, supra note 54, at 55. She found that “there exists exclusive federal jurisdiction over claims which (i) are brought to enforce the rights and liabilities created by the Securities Act; and (ii) are covered class actions.” Id. Judge Wolfson ultimately concluded that pure Securities Act claims are removable, not because the SLUSA amendment provides for removal, but because “state courts lack[] jurisdiction over [pure] Securities Act claims in the first place.” See id.
172 See Lowenthal & Haggarty, SLUSA’s Elimination, supra note 62, at 22; Lowenthal & Haggarty, Jurisdictional Struggles, supra note 54, at 55.
174 Id. § 77p. See Lowenthal & Haggarty, Jurisdictional Struggles, supra note 54, at S14; see also Knox v. Agria Corp., 613 F. Supp. 2d 419 (S.D.N.Y. 2009). The plaintiffs
however, is the definition section and sets forth the meaning of a “covered class action” as:

[A]ny single lawsuit in which . . . one or more named parties seek to recover damages on a representative basis on behalf of themselves and other unnamed parties similarly situated, and questions of law or fact common to those persons or members of the prospective class predominate over any questions affecting only individual persons or members.175

On its face, a “covered class action” seems to apply to “any single lawsuit” that meets the definition, and therefore includes pure Securities Act claims. As Judge William Pauley concluded, “Section 16 . . . does not add a substantive limitation to the exception to concurrent jurisdiction in Section 22(a).”176 Thus the effect of SLUSA’s amendment is to replace “concurrent jurisdiction with exclusive federal jurisdiction over ‘covered class actions . . . brought to enforce any liability or duty created by [the Securities Act].’”177

As Judge Pauley points out, the anti-removal provision of Section 22(a) provides that “no case arising under [the Securities Act] and brought in any State court of competent jurisdiction shall be removed to any court of the United States.”178 Because SLUSA’s amendment to the concurrent jurisdiction provision of Section 22(a) replaced concurrent jurisdiction with exclusive federal jurisdiction, state courts are

---

176 Knox, 613 F. Supp. 2d at 421. It is still clear, however, that SLUSA’s amendments do not reach individual actions brought in state court. See Cal. Pub. Employees’ Ret. Sys. v. WorldCom, Inc., 368 F.3d 86, 98 (2d Cir. 2004) (“SLUSA did not in any way alter Section 22(a)’s bar on removal of individual Securities Act claims[,]”); see also Lowenthal & Haggarty, SLUSA’s Elimination, supra note 62, at 22.
177 Lowenthal & Haggarty, SLUSA’s Elimination, supra note 62, at 21 (citing 15 U.S.C. § 77v(a)).
no longer courts of competent jurisdiction. Therefore, pure Securities Act claims are no longer covered by the anti-removal provision at all.\textsuperscript{179} As a result, defendants are no longer limited by the “[e]xcept as otherwise expressly provided by Act of Congress”\textsuperscript{180} provision, and they may successfully remove pure Securities Act claims as they would any other federal claim, through the general removal mechanism of Section 1441(a).\textsuperscript{181}

Some have been critical of this recent approach, finding that it lacks sufficient support from other courts.\textsuperscript{182} It appears, however, that this interpretation is becoming more of an emerging trend, as more courts are adopting it.\textsuperscript{183} One of the most recent decisions to address the removal of pure Securities Act claims was \textit{Lapin v. Facebook, Inc.},\textsuperscript{184} where the United States District Court for the Northern District of California was faced with several putative class actions asserting pure Securities Act claims against Facebook. Despite authority in other California districts finding that SLUSA’s amendment does not provide for removal of pure Securities Act claims,\textsuperscript{185} Judge Maxine Chesney reached the opposite conclusion.\textsuperscript{186} She found that “federal courts alone have jurisdiction to hear covered class actions raising [pure Securities Act] claims.”\textsuperscript{187}

\textsuperscript{179} See Lowenthal & Haggarty, \textit{Jurisdictional Struggles}, supra note 54, at S14.
\textsuperscript{180} See 28 U.S.C. § 1441(a); see also Costa, supra note 42, at 1198 (“Congress has reserved the right to circumscribe removal in all cases in which they chose to do so in . . . 28 U.S.C. § 1441.”).
\textsuperscript{181} See 28 U.S.C. § 1441(a).
\textsuperscript{182} See Butts, supra note 110, at 195-96. Butts is also critical of the \textit{Knox} court’s dismissal of \textit{Kircher} as “inapplicable and ‘not to the contrary.’” \textit{Id.} at 196. Butts argues that \textit{Knox} misunderstood the meaning of dicta, and even though \textit{Kircher} “did not have to explicitly decide the federal law removal question in its decision, . . . [d]icta are statements made in a decision that . . . are nonetheless important to the deciding court.” \textit{Id.} Thus the \textit{Knox} court should have accorded the Supreme Court’s decision “more respect than . . . [a] cursory dismissal.” \textit{Id.}
\textsuperscript{184} \textit{Id.}
\textsuperscript{187} \textit{Lapin}, 2012 WL 3647409, at *2 (quoting Knox v. Agria Corp., 613 F. Supp. 2d 419 (S.D.N.Y. 2009)). Following this decision, motions were made pursuant to 28 U.S.C. § 1407 before the United State Judicial Panel on Multidistrict Litigation to transfer the numerous actions relating to the Facebook IPO to the United States District Court for the Southern District of New York. See Notice of Hearing Session, \textit{In re Facebook, Inc., IPO Securities and Derivative Litigation}, MDL No. 2389 (S.D.N.Y. Aug. 8, 2012). The motion sought to consolidate 15 actions from the Northern District of California, as well as actions from the Middle District of Florida and Western District of Missouri, with the 26 pending actions in the Southern District of New York.
Relying on SLUSA’s amendment to the concurrent jurisdiction provision of the Securities Act, the Court found that “the [anti]-removal provision in [Section 22(a)] ... no longer applies to ‘a covered class action’ alleging [pure Securities Act claims], and, consequently, [they are] removable pursuant to 28 U.S.C. § 1441(a).” The court noted that SLUSA’s legislative history, with its focus on correcting and strengthening the PSLRA and its goal of “national standards for securities class action lawsuits involving nationally traded securities,” supports the “plain meaning” interpretation of the concurrent jurisdiction provision.

Only months later, the District Court for the Northern District of California was given an opportunity to reinforce its decision in Lapin. On August 1, 2012, a class action against Zynga, Inc., a provider of social game services, commenced in California state court. Defendants subsequently removed the case to the District Court for the Northern District of California. On October 25, 2012, the plaintiff, Robert Reyes, filed a motion to remand the case back to state court. Despite Judge Chesney’s recent ruling in Lapin, Judge White for the Northern District of California granted the plaintiff’s motion to remand the case. Noting the “divided district courts within this district and around the country,” the court concluded that the narrow approach was more persuasive. While the defendants “raise[d] valid for pre-trial proceedings. Id. The motion included the Lapin, Spatz, Lazar, Stokes, and DeMois actions mentioned supra. See supra notes 7, 9, 10 and accompanying text.


188 Lapin, 2012 WL 3647409, at *3.
189 See id. at *2.
191 See id. at *3 (citing H.R. REP. 105-802, at 13 (1998) (“The purpose of ... [SLUSA] is to prevent plaintiffs from seeking to evade the protections that Federal law provides against abusive litigation by filing suit in State, rather than in Federal, court.”)).
192 Id.
194 See Plaintiff’s Motion to Remand at 1, Reyes v. Zynga, Inc., No. 12 Civ. 5065 (N.D. Cal. 2013).
196 Id. at 3-5.
arguments about inconsistencies between Section [22(a)] and Sections [16(b)] and [16(c)],” the court also found that “there are plausible ways to construe Section [22(a)] to avoid any such inconsistencies.” Given the “lack of clear authority from the Supreme Court or Ninth Circuit on this issue,” as well as the split among district courts, the court followed the principle “that any doubt about the propriety of removal should be resolved in favor of remand.” To further add to the inconsistency among district courts, Judge White admitted that “given the intent of SLUSA[,] it just makes no sense to prohibit the removal of federal securities class actions to federal court.”

Because the concurrent jurisdiction approach aligns the plain meaning of the SLUSA amendment with congressional intent, district courts should continue to adopt it when analyzing the removal of pure Securities Act claims. Following this trend is all the more important because federal appellate courts have yet to address the removability of pure Securities Act claims, despite the significant split authority on the issue among district courts. And even with the emerging prominence of the concurrent jurisdiction approach, it appears that no conclusive answer can be expected. District court decisions are persuasive, not binding, authority on each other, “even in the same district, and the district judges are free to resolve legal questions like these unless there is controlling circuit or Supreme Court authority.” These problems are exacerbated because district court decisions ordering remand are “not reviewable on appeal,” as “[f]ederal appellate courts can

---

197 Id. at 5. The Court also concluded that while the legislative history does suggest that “SLUSA was amended to make federal courts the exclusive venue for most class actions alleging securities fraud, . . . the legislative history is, itself, murky insofar as it suggests an answer to the question before the Court.” Id. at 5-6.

198 Id. at 6 (citing Gaus v. Miles, Inc., 980 F.2d 564, 566 (9th Cir. 1992)).

199 Id. at 6 (citing Unschuld v. Tri-S Sec. Corp., No. 6 Civ. 2931, 2007 WL 2729011, at *8-9 (N.D. Ga. 2009)).


202 MILBANK, TWEED, HADLEY & MCCOY LLP 3, supra note 200 (citing 28 U.S.C. § 1447(d) (“An order remanding a case to the State court from which it was removed is not reviewable on appeal or otherwise . . . .”)).
hear appeals from only final decisions, and denials of remand are not considered final.”203 This leaves parties who wish to appeal the decision to remand a removed case with two options: (1) a mandamus or interlocutory appeal, or (2) the issue must be raised on direct appeal at the case’s end.204 The former is a highly discretionary device, while the latter rarely arises because “most securities actions are dismissed or settled…”205 Absent congressional action on the issue, plaintiffs will be able to continually forum shop for jurisdictions where federal judges are likely to remand to the state courts.

III. INVESTOR PROTECTION

Regardless of the approach taken, allowing removal of pure Securities Act claims to federal court aligns with congressional intent and furthers the goal of investor protection, which lies at the heart of the Securities Act. Only by interpreting SLUSA to permit removal of covered class actions asserting pure Securities Act claims can the PSLRA fully accomplish its goal of “decreasing [the] vexatious strike suits” that harm investors.206 “The private securities litigation system is too important to the integrity of American capital markets to allow this system to be undermined by those who seek to line their own pockets by bringing abusive and meritless suits.”207 The “in terrorem” effect of plaintiffs bringing groundless claims encourages massive settlements to avoid the potential for more devastating judgments.208 If the trend from the first half of 2013 continues, there will be approximately 222 new securities class actions filed in 2013, each with an average settlement value of $78 million, which is more than double the six-year average of $35 million.209 Because strike suits have such a high settlement value, without regard to actual culpability, the ultimate effect is to harm not only the innocent corporation, but its owners as well, the investors.

Permitting removal of pure Securities Act claims will prevent the meritless harassment of corporations by plaintiffs

203 Id.
204 See id.
205 See id.
206 Snyder, Jr., supra note 45, at 698.
208 See Snyder, Jr., supra note 45, at 699.
that sidestep the safeguards of the PSLRA, the cornerstone of Congress’s fight against abusive securities litigation. By enacting the PSLRA, Congress hoped to protect the innocent investors from the greedy ones, those who exploit meritless claims into large settlements. Through a number of procedural and jurisdictional reforms, the PSLRA made it increasingly difficult to bring securities class actions under federal securities laws.\textsuperscript{210} For example, under the new rules imposed by the PSLRA, covered class actions brought in federal court could take three years to arrive at the same point in litigation that a comparable case brought in state court would reach in only a few days.\textsuperscript{211} The ultimate benefit of permitting removal of pure Securities Act claims will be to decrease the number of vexatious strike suits, which in turn “protect[s] investors, who are ‘always the ultimate losers when extortionate “settlements” are extracted from issuers.’”\textsuperscript{212}

SLUSA and the PSLRA also work to encourage individual action by institutional investors.\textsuperscript{213} Neither SLUSA nor the procedural protections of the PSLRA apply to individual suits in state court.\textsuperscript{214} Institutional investors, such as large stakeholders, can file their own individual suits in state court and “avoid the costly and time-consuming burdens of PSLRA.”\textsuperscript{215} Furthermore, individual suits do not pose the same risk of harassment because the individual will bear the cost of litigation himself, including the cost of losing.\textsuperscript{216} An interested institutional investor is also much more likely to be looking out for the interests of smaller investors than a “‘professional plaintiff’’s lawyer,”\textsuperscript{217} and permitting pure Securities Act claims to be remanded back to state court may encourage these institutional investors to once again become

\textsuperscript{210} See supra, part I.C.

\textsuperscript{211} See Snyder, Jr., supra note 45, at 698 (citing Jonathan F. Mack, \textit{PSLRA and SLUSA: Laws with Unintended Consequences}, \textsc{Wallstreetlawyer.com: Sec. Elec. Age} 1 (Nov. 2003)) (internal quotation marks omitted).

\textsuperscript{212} Snyder, Jr., supra note 45, at 700 (quoting H.R. Rep. No. 104-369, at 31-32).

\textsuperscript{213} Id.

\textsuperscript{214} State causes of action in general are not prohibited, rather covered class actions asserting such claims are preempted. See Merrill Lynch, Pierce, Fenner, & Smith Inc. v. Dabit, 547 U.S. 71 (2006). “This . . . reflects Congress’s additional concern for the class action’s potential for abuse.” Snyder, Jr., supra note 45, at 701.

\textsuperscript{215} Snyder, Jr., supra note 45, at 701. “[T]he combined effect . . . has been that nearly every major class action securities fraud case is now accompanied by intensely litigated individual actions by institutions that never brought such suits ten years ago.” Id. (citing Jonathan F. Mack, \textit{PSLRA and SLUSA: Laws with Unintended Consequences}, \textsc{Wallstreetlawyer.com: Sec. Elec. Age} 1 (Nov. 2003)).

\textsuperscript{216} See id.

\textsuperscript{217} See id. (quoting Keith L. Johnson, \textit{Deterrence of Corporate Fraud Through Securities Litigation: The Role of Institutional Investors}, \textsc{Law & Contemp. Prosbs.} 155, 155 (Autumn 1997)).
“satisfied to quietly wait on the sidelines and take a modest check as a class member.” Thus a narrow reading of SLUSA’s amendments “would deprive the market of the benefit of more active participation of institutional investors.”

CONCLUSION

For almost 80 years, the Securities Act and the Exchange Act have worked to protect investors from the misdeeds of corporations. They remained largely untouched until the 1990s, when Congress saw fit to protect investors from a new harm: the vexatious strike suits initiated and maintained by fellow investors. Congress enacted the PSLRA under the assumption that “both investors and the national economy suffer when innocent parties are forced to pay ‘exorbitant settlements’ in meritless lawsuits.” When plaintiffs began to circumvent the PSLRA, Congress responded by enacting SLUSA and established national standards for securities class actions involving nationally traded securities. SLUSA bolstered the effectiveness of the PSLRA by making federal court the exclusive venue for certain class actions.

These efforts were hampered, however, because of the poor and confusing drafting of the amendment. SLUSA’s amendment to the anti-removal provision of the Securities Act resulted in conflicting broad and narrow interpretations. Both interpretations have received support from the Supreme Court, but both ultimately have large flaws that prevent either from gaining traction among federal district courts. The narrow interpretation attempts to find the plain meaning of the statute, but it does so in a conclusory fashion. It fails to acknowledge the confusion that the poorly written statute creates. The broad interpretation acknowledges the confusion and instead looks to the legislative intent for clarity. While certain statements of members of the legislature certainly support the notion that SLUSA was meant to make federal court the exclusive venue for class actions, these statements alone have been insufficient to persuade many district courts. The result of this district court split has enabled the continued circumvention of the PSLRA and the continued use of vexatious

218 Id. n.217 (citing Mack, supra note 215 (internal quotation marks omitted).
219 Id. at 702. The PSLRA also encourages the appointment of such institutional investors as lead plaintiff, who have an incentive to maximize recoveries in class action lawsuits. Id. at 702-03.
220 Mazzeo, supra note 19, at 1446.
221 Id. 1450-51.
strike suits by plaintiffs to harass innocent defendant-corporations into settlements. The conflicting approaches also encourage forum shopping by plaintiffs, and the persuasive nature of district court authority makes litigation even more unpredictable and expensive for all parties involved.

As this note demonstrates, a third interpretation of SLUSA’s amendment to the Securities Act has garnered some support among the district courts. By addressing SLUSA’s amendment to the concurrent jurisdiction provision, this interpretation looks to both the plain meaning of the text and the legislative history in concluding that class actions asserting pure Securities Act claims are excluded from federal court. Instead, these claims are trusted to the federal courts, as Congress intended. District Court judges should continue to adopt this analysis because allowing removal of purely federal claims avoids the paradox of keeping federal claims in state court and furthers the goals of investor protection by allowing the protections of the PSLRA to take full effect.

Stephen O’Connor†

---

222 Serota, supra note 20, at 175.
† J.D. Candidate, Brooklyn Law School, 2014; B.S., Lehigh University, 2011. I would like to thank the Brooklyn Law Review staff for all of their help and support throughout the note writing process. I would also like to thank my family for their love and support throughout my life.
The Confluence of *Sullivan v. Harnisch* & Dodd-Frank

**ADAPTING NEW YORK’S COMMON LAW TO FILL A COMPLIANCE HOLE**

**INTRODUCTION**

In May 2012, Joseph Sullivan lost a wrongful termination suit against Peconic Partners LLC and Peconic Asset Managers LLC (together, Peconic). Sullivan, a partner of both firms, alleged that he was fired after bringing improper trading activity to the attention of Peconic’s CEO. In its decision, the New York Court of Appeals adhered to New York’s strict at-will employment doctrine, which states that “absent violation of a constitutional requirement, statute or contract, ‘an employer’s right at any time to terminate an employment at will remains unimpaired.’” The court affirmed the Appellate Division’s decision to grant the defendants’ motion for summary judgment. The court in *Sullivan* declined to create a new exception to New York’s at-will employment doctrine and rejected an application of the doctrine’s only exception to salvage Sullivan’s claim.

Although the Court of Appeals may have ruled in Peconic’s favor, its decision may ultimately undercut the ability of New York’s private investment advisers to maintain effective compliance programs. When coupled with existing state and federal whistleblower protections, the outcome in *Sullivan* only adds to the already strong incentives encouraging a similarly

---

2 *Id.* at 759.
3 *Id.* at 760 (quoting Murphy v. Am. Home Prods. Corp., 448 N.E.2d 86, 91 (N.Y. 1983)).
4 *Id.* at 761.
5 The exception pertains to attorneys within law firms, and it protects attorneys who internally report the improper behavior of colleague attorneys. See Wieder v. Skala, 609 N.E.2d 105 (N.Y. 1992).
6 *Sullivan*, 969 N.E.2d at 760.
situated employee to externally disclose a potential securities law violation instead of making the complaint in-house.\textsuperscript{7} Neither New York’s common law nor its whistleblower statute\textsuperscript{8} provide employees of private investment advisers adequate protection against retaliatory action for disclosing potential securities law violations.\textsuperscript{9} These firms include hedge funds and private equity funds, many of which under Dodd-Frank were required to register with the SEC pursuant to the Investment Advisers Act of 1940.\textsuperscript{10} As registered investment advisers, these firms must appoint chief compliance officers and create policies and procedures that promote adherence to federal securities laws.\textsuperscript{11}

To protect oneself against retaliatory action, an employee like Sullivan—one who becomes aware of a potential securities law violation by her employer but works as an at-will employee for a private company and therefore cannot rely on an employer’s promise against retaliatory conduct—faces a stark choice when considering whether to blow the proverbial whistle. If the employee is concerned primarily with retaining her position and avoiding other retaliatory action by her employer, she must either remain silent, allowing the possible securities law violation to go unaddressed, or she must externally report the potential violation to the SEC.\textsuperscript{13} Although Dodd-Frank augments the protections and incentives for would-be whistleblowers,\textsuperscript{14} this law does not extend protection to employees of private companies who limit disclosure to an internal audience.\textsuperscript{15} In order for employees like Sullivan to invoke Dodd-Frank’s protection against retaliatory action, they must externally disclose the impropriety to the SEC.\textsuperscript{16}

\textsuperscript{7} See infra Part III.
\textsuperscript{8} N.Y. LAB. LAW § 740 (McKinney 2013).
\textsuperscript{9} See infra Part I.
\textsuperscript{11} 17 C.F.R. § 275.206(4)-7(a) (2012); see infra Part II.
\textsuperscript{13} See infra Part II.
\textsuperscript{14} See 17 C.F.R. §§ 240.21F-2–F-3.
\textsuperscript{16} See Egan, 2011 WL 1672066. There, the court held as a matter of first impression that, because the plaintiff’s disclosure did fall within one of the few statutory exceptions, “the anti-retaliation whistleblower protection provisions of the
The *Sullivan* decision and current securities laws strip from the employee’s consideration the sensible option of limiting disclosure to a supervisor or the firm’s internal compliance program.

This note examines the effect *Sullivan v. Harnisch* has on securities-related disclosure among private investment advisers in New York in a post Dodd-Frank world. Depending on one’s vantage point, the combination of federal securities laws and the holding of *Sullivan v. Harnisch* either promotes a policy of external disclosure—one facilitating adherence to securities laws—or it further skews employee incentives to stay silent, one promoting an ineffective method of compliance enforcement. Although these incentives may promote the enforcement goals of the SEC, they may also run counter to the motivations and considerations of would-be whistleblowers. The goal of policymakers should be to craft a policy that creates the most efficient mechanism to ensure compliance with securities laws. In order to achieve this result, either New York or federal policy must change so that employees are protected if they choose to limit disclosure of a possible securities law violation to an internal audience.

Because the current paradigm pushing employees to make external disclosure arises from a combination of state and federal policies, many sources for a potential solution exist. Companies can create and promote strong compliance programs—systems that proscribe retaliation or that preserve anonymity. Further, Congress could amend Dodd-Frank to forbid retaliation for internally reporting violations among all financial companies. Finally, although perhaps unlikely, New York courts could adapt existing common law to protect internal disclosure. In *Wieder v. Skala*, the court recognized the role self-policing plays

---


18 See infra Part IV.
19 See infra Part IV.
20 See infra Part IV.
21 “American courts, including our own, have proved chary of creating common-law exceptions to the rule and reluctant to expand any exceptions once fashioned.” *Sullivan v. Harnisch*, 969 N.E.2d 758, 760 (N.Y. 2012) (quoting Horn v. N.Y. Times, 790 N.E.2d 753, 755 (N.Y. 2003)).

---
in the legal profession and carved out the only exception to its at-will employment doctrine for an attorney who was fired after attempting to discipline a colleague.\textsuperscript{23} By augmenting established contractual principles with this policy of self-regulation, New York courts can establish a framework that protects employees who make internal disclosure without significantly undermining its at-will employment doctrine. Employees should have the option of limiting disclosure of potential securities law violations to an internal audience. And although the adoption of a new common law policy would, in certain circumstances, infringe upon an employer’s right to discharge employees, establishing these legal protections would foster internal disclosure and improve a firm’s ability to maintain effective compliance programs.

Part I of this note will focus on New York common law and the statute relevant to would-be whistleblowers of private companies. This section will review \textit{Sullivan v. Harnisch} as well as other cases to reveal the current contours of New York’s at-will employment doctrine. This section will also address New York’s whistleblower statute and the difficulty one would face trying to invoke its protections in the context of reporting securities law violations. Part II will explore the impact of Dodd-Frank, emphasizing the heightened regulatory standards imposed upon most private investment advisers. This section will also explore the implementation of Dodd-Frank’s whistleblower protections and the related SEC rules. Part III will juxtapose New York’s employment and whistleblower policies with Dodd-Frank. This juxtaposition will demonstrate how all roads now point to external disclosure and how such a path may not align with the behavior of employees or the compliance interests of employers. Part IV will assess possible remedies, including those that would extend anti-retaliatory protections to employees that engage in internal disclosure.

I. \textbf{NEW YORK’S COMMON LAW AND STATUTORY PROTECTIONS AGAINST RETALIATORY TERMINATION}

In order to understand how the combination of current federal and state policies pushes an employee of a private investment adviser to externally disclose potential securities law violations, one must first understand the protection against retaliatory employer conduct, or lack thereof, offered to employees by New York law. To do so, the contours of New

\textsuperscript{23} \textit{Id.} at 108–09.
York’s at-will employment doctrine and its whistleblower statute must be explored.

A. New York’s Common Law Provides Little Help

1. Sullivan v. Harnisch

The New York Court of Appeals addressed the issue of internal disclosure of alleged hedge fund improprieties in Sullivan v. Harnisch.\(^{24}\) Joseph Sullivan was a 15% partner of Peconic and bore many titles at the firm, such as Chief Compliance Officer, Executive Vice President, and Chief Operating Officer.\(^{25}\) Sullivan alleged he was fired days after confronting William Harnisch, Peconic’s CEO, over Harnisch’s alleged “manipulative and deceptive trading practices”\(^{26}\) known as “front-running,” where one “[takes] advantage of investment opportunities that should first be accorded to its clients.”\(^{27}\)

In his appeal, Sullivan asserted that he was required to disclose trading activity that ran afoul of federal securities laws and the firm’s internal code of ethics.\(^{28}\) Sullivan argued that his objection to the misconduct was improper grounds for termination.\(^{29}\) The court rejected Sullivan’s position and reaffirmed its staunch adherence to New York’s at-will employment doctrine.\(^{30}\) Without a finding of breach of contract, or a violation of Sullivan’s statutory or constitutional rights, Harnisch and Peconic reserved the right to fire Sullivan for any reason.\(^{31}\)

The court also rebuffed Sullivan’s attempt to invoke the only judicially recognized exception to New York’s at-will employment doctrine.\(^{32}\) This exception, espoused in Wieder v.

---

\(^{24}\) Sullivan, 969 N.E.2d 758.

\(^{25}\) Id. at 759.

\(^{26}\) Id.

\(^{27}\) Sullivan v. Harnisch, 915 N.Y.S.2d 514, 516-17 (N.Y. App. Div. 2010), aff’d, 969 N.E.2d 758 (N.Y. 2012). Harnisch sold two-thirds of a personal $100 million position in Potash Corp.—a security also held in his clients’ funds—days before the release of a disappointing earnings report from a related company. Id. at 516. Harnisch did not execute similar trades on behalf of his clients until after the release of the earnings report. Id. As a result, Harnisch pocketed $132 per share whereas the funds earned $103 per share. According to Sullivan, this practice violated Peconic’s SEC filings and its internal compliance manual. Id. at 517.

\(^{28}\) Sullivan, 969 N.E.2d at 759.

\(^{29}\) Id.

\(^{30}\) Id.

\(^{31}\) Id. at 760.

\(^{32}\) Id. at 760–61; see Wieder v. Skala, 609 N.E.2d 105, 110 (N.Y. 1992) (discussed infra Part I).
Skala,\textsuperscript{33} recognized a breach of contract based upon an implied obligation where an attorney, bound by his professional responsibilities, was fired after insisting that a colleague face disciplinary action for malpractice.\textsuperscript{34} The court in Sullivan noted that Sullivan's compliance duties at Peconic were not “so closely linked [to his position] as to be incapable of separation,” or that compliance was not “at the very core, and indeed, the only purpose of Sullivan’s employment.”\textsuperscript{35} In dissent, Chief Judge Lippman opined that “the majority unwisely limits the exception to the at-will employment . . . . In so doing, it creates a great potential for abuse in the financial services industry.”\textsuperscript{36}

Although never raised by the parties or the court, Sullivan's conduct was not protected by New York's whistleblower statute.\textsuperscript{37} And although Sullivan's termination occurred before the passage of Dodd-Frank, the court observed that Dodd-Frank's whistleblower protections would have nevertheless been inapplicable.\textsuperscript{38}

2. The Confines of New York's Common Law

The court's decision in Sullivan v. Harnisch represents one of the latest in a line of decisions that express New York's at-will employment doctrine as highly deferential to the judgment and decisions of employers. The cases that follow articulate the contours of New York's at-will employment doctrine. As they demonstrate, plaintiffs who allege that they were wrongfully discharged without asserting a claim that is cognizable under a contractual lens likely will fail in their lawsuit.

In Weiner v. McGraw-Hill,\textsuperscript{39} however, the New York Court of Appeals held that the plaintiff brought a facially valid breach of contract action after being fired.\textsuperscript{40} Walter Lewis Weiner, a long-time employee of McGraw-Hill who steadily ascended through the ranks there before being fired, successfully persuaded the Court of Appeals that a provision in the McGraw-Hill’s personnel handbook might have constituted an express promise

\textsuperscript{33} Wieder, 609 N.E.2d 105.
\textsuperscript{34} Id. at 110.
\textsuperscript{35} Sullivan, 969 N.E.2d at 761 (quoting Wieder, 609 N.E.2d at 108 (internal quotation marks omitted)).
\textsuperscript{36} Id. at 765 (Lippman, C.J., dissenting).
\textsuperscript{37} See infra Part I.B.
\textsuperscript{38} Sullivan, 969 N.E.2d at 761 (“[Dodd-Frank] seems not to apply to conduct like that alleged in Sullivan's complaint; Sullivan does not claim to have blown a whistle—i.e., to have told the SEC or anyone else outside Peconic about Harnisch's alleged misconduct—but only to have confronted Harnisch himself.”).
\textsuperscript{40} Id. at 443.
within his employment contract. The employee handbook limited termination to instances where there was “just and sufficient cause.” The company’s application for employment stipulated that Weiner’s “employment [was] subject to the provisions of [the handbook].” These factors, along with Weiner’s alleged reliance on this for-cause provision as a basis for working for McGraw-Hill and the employer’s prior insistence that Weiner adhere to the corporate policy when faced with staffing decisions, led the court to conclude that Weiner articulated a legitimate question of whether McGraw-Hill had breached a contract it held with Weiner. 

In Murphy v. American Home Products Corporation, the court considered two separate causes of action that, when combined, form the spine of the single claim considered by the Sullivan court. Plaintiff Joseph Murphy alleged that he was terminated after internally disclosing accounting improprieties that resulted in an overstatement of the firm’s earnings. These overstatements allegedly enabled the firm’s management to receive “unwarranted” bonuses. Murphy raised several causes of action, including one sounding in wrongful, retaliatory termination and another in breach of an implied employment contract. In addressing Murphy’s retaliatory discharge claim, the court concluded that whether a wrongful termination cause of action should be created was a matter reserved for “a principled statutory scheme, adopted after opportunity for public ventilation, rather than in consequence of judicial resolution of the partisan arguments of individual adversarial litigants.” Murphy also argued that “he was required by the terms of his employment to disclose accounting improprieties” and that by his termination, his employer breached the implied “obligation on the part of the employer to deal with . . . employees fairly” that

---

41 Id. at 442–43, 45
42 Id. at 442. The pertinent text of the personnel handbook read:

[t]he company will resort to dismissal for just and sufficient cause only, and only after all practical steps toward rehabilitation or salvage of the employee have been taken and failed. However, if the welfare of the company indicates that dismissal is necessary, then that decision is arrived at and is carried out forthrightly.

Id. (alteration in original).
43 Id.
44 Id. at 445–46.
46 Id. at 87.
47 Id.
48 Id. at 89, 91.
49 Id. at 90.
permeates all employment contracts. Finding no precedent to support this proposition and wary of its broad policy implications, the court rejected Murphy’s argument and held that whether an employer’s right to terminate an at-will employment should be infringed is a matter best left to the legislature.

The only judicially recognized exception to New York’s at-will employment doctrine was espoused in Wieder v. Skala. Howard Wieder, an associate at a law firm, reported to his supervising partners the “false and fraudulent material misrepresentations” a colleague attorney made in an effort to conceal his negligence in handling a real estate transaction for Wieder. Wieder, who at that time claimed he spearheaded the firm’s most important litigation project, insisted that his firm pursue disciplinary action against his colleague before the New York Appellate Division Disciplinary Committee. He abandoned this request after his supervisors threatened to fire him. That action, apparently, did not pacify Wieder’s employer; he was fired soon after he submitted important filings related to his litigation project.

In review of his action for wrongful termination and breach of contract, the court concluded that an exception to New York’s at-will employment doctrine was warranted. The court noted “the unique function of self-regulation belonging to the legal profession” and the “essential compact that in conducting the firm’s legal practice both plaintiff and the firm would do so in compliance with the prevailing rules of conduct and ethical standards of the profession.” In its analysis, the court placed great emphasis on the fact that, as an attorney, Wieder’s conduct was governed by a code of professional responsibility. The code, in pertinent part, prohibits attorneys from keeping secret unprivileged conduct or statements made by another attorney that would impeach that attorney’s integrity and trustworthiness. The court held that the plaintiff stated a

50 Id. at 91.
51 Id.
53 Id. at 106.
54 Id.
55 Id.
56 Id.
57 Id. at 108.
58 Id. at 110.
59 Id. at 108–09.
60 Id. at 106 n.1; N.Y. COMP. CODES R. & REGS. tit. 22, § 1200.0 rule 8.3(a) (2012).
valid breach of contract claim “based on an implied-in-law obligation in his relationship with defendant.”

B. New York’s Whistleblower Law Does Not Apply to Financial Harms

Under New York’s whistleblower statute, an employee who shares or threatens to share “an activity, policy or practice of the employer that is in violation of law, rule or regulation which violation creates and presents a substantial and specific danger to the public health or safety” is protected from retaliatory action. The statute protects both internal and external disclosure by a whistleblower.

The application of New York’s whistleblower statute, however, is limited. First, “the requirements of a violation and a danger to health or safety are conjunctive, the statute does not protect an employee who reports activity presenting such a danger . . . if the activity does not violate a specific statute or regulation.” Further, “[t]he law requires that there be . . . an actual, as opposed to a possible, violation . . . Reasonable belief as a basis for protection under [New York’s whistleblower law] will not suffice.” Finally, the jurisprudential interpretation of New York’s whistleblower statute does not affiliate “corporate wrongdoing or white-collar crimes” with the type of danger to the public the law is intended to prevent. For example, in Clarke v. TRW, Inc., the court noted that a defective electrical relay that could give rise to problems with an automobile’s brakes or fuel pump qualified as a harm covered by New York’s whistleblower law. In Susman v. Commerzbank Capital Markets Corp., on the other hand, the court affirmed a lower court’s dismissal of a state whistleblower claim because the firm’s alleged

---

61 Wieder, 609 N.E.2d at 110.
62 N.Y. LAB. LAW § 740(2)(a) (McKinney 2013).
63 Id.
66 Silvia X. Liu, Note, When Doing the Right Thing Means Losing Your Job: Reforming the New York Whistleblower Statute, 7 N.Y. CITY L. REV. 61, 71 (2004). Liu argues, among other things, that New York’s whistleblower statute should be amended such that “an employee should be protected if she or he reasonably believes that a violation of the law, rule or regulation has occurred or is occurring.” Id. at 64.
68 Id. at 935–36.
illegality—engaging in transactions with the Central Bank of Iran, a federally proscribed activity—did not constitute a “substantial and specific danger to the public health or safety.”\textsuperscript{70} As one court noted, “financial improprieties [are not] placed on the same plane as threats to public health or safety.”\textsuperscript{71} One justification for this distinction was to “avoid the cumulative effect of a potential flood of lawsuits seeking to convert ordinary employment disputes into ‘whistle-blower’ protection cases.”\textsuperscript{72}

In sum, neither New York’s common law nor its whistleblower statute extend protection against retaliatory conduct to an employee who makes internal disclosure about a potential securities law violation. The tools exist, however, for New York courts presented with these circumstances to craft a common law remedy against retaliatory conduct.

II. DODD-FRANK’S IMPACT—PRIVATE INVESTMENT ADVISER REGISTRATION REQUIREMENTS AND WHISTLEBLOWER PROTECTIONS

A. Hedge Fund Compliance Requirements

Prior to the passage of Dodd-Frank, hedge funds and other private investment advisers largely avoided SEC registration and the Commission’s direct oversight by taking advantage of Section 203(b)(3) of the Investment Advisers Act of 1940.\textsuperscript{73} This provision extended a registration exemption for certain investment advisers with fewer than 15 clients in the preceding year.\textsuperscript{74} The D.C. Circuit in \textit{Goldstein v. SEC} \textsuperscript{75} molded the shape of this exemption and determined that the term “clients” referred to the “advising of

\textsuperscript{70} Id. at 7.
\textsuperscript{71} McGrane v. Reader’s Digest Ass’n, Inc., 822 F. Supp. 1044, 1046 (S.D.N.Y. 1993) supplemented, 92 CIV. 8132 (VLB), 1993 WL 525127 (S.D.N.Y. Dec. 13, 1993) (declining to extend New York state whistleblower protections to an employee that was fired after internally reporting financial improprieties). The plaintiff submitted a five-pound complaint but the court determined that it neither established a continuing fraud or a “hazard to health or safety which has not been remedied by [the employer].” \textit{Id.}
\textsuperscript{72} Id. at 1046.
\textsuperscript{74} Seth Chertok, \textit{A Detailed Analysis of Title IV of the Dodd-Frank Wall Street Reform and Consumer Protection Act}, 6 VA. L. & BUS. REV. 1, 6 (2011). Additionally, to qualify for the exemption, funds could not “hold themselves out to the public as investment advisers, and [could] neither act[] as an investment adviser to any registered investment company, nor to a company which has elected to be a ‘business development company.’” \textit{Id.}
\textsuperscript{75} Goldstein v. SEC, 451 F.3d 873 (D.C. Cir. 2006).
pooled investment vehicles rather than the individual investors therein."\textsuperscript{76} This statutory interpretation allowed a hedge fund, regardless of the number of its investors, to avoid registration so long as it managed less than 15 funds.\textsuperscript{77}

Title IV of Dodd-Frank in large part eliminated this exemption for large private investment advisers.\textsuperscript{78} Noting that “their trades can move markets,”\textsuperscript{79} Congress determined that “information regarding [a private investment adviser’s] size, strategies, and positions could be crucial to regulatory attempts to deal with a future crisis.”\textsuperscript{80} Consequently, as of March 2012, large private investment advisers must now comply with the Investment Advisers Act of 1940.\textsuperscript{81} This change resulted in a large increase in the number of registered investment advisers, expanding the number of firms regulated by the SEC.\textsuperscript{82} Registered private investment advisers must draft a code of ethics that, among other things, include “[p]rovisions requiring [the firm’s] supervised persons to comply with applicable Federal securities laws”\textsuperscript{83} and “[p]rovisions requiring supervised persons to report any violations of [the firm’s] code of ethics promptly to [the firm’s] chief compliance officer or, provided [the firm’s] chief compliance officer also receives reports of all violations, to other persons . . . designate[d] in [the firm’s] code of ethics.”\textsuperscript{84} A registered private investment adviser must also maintain “books and records” that, among other things, record employee violations of the firm’s code of ethics and the remedial actions taken in response.\textsuperscript{85} These firms must also appoint a chief compliance officer\textsuperscript{86} and create

\textsuperscript{76} Chertok, \textit{supra} note 74, at 8.
\textsuperscript{80} Id. at 72.
\textsuperscript{81} SEC Adopts Dodd-Frank Act Amendments, \textit{supra} note 73. Dodd-Frank created three new exemptions for private advisers. “Certain foreign advisers without a place of business in the U.S.” private advisers that only oversee venture capital funds, and advisers to only private funds “with less than $150 million in assets under management in the U.S.,” are exempt from SEC registration. \textit{Id.} at 3.
\textsuperscript{84} \textit{Id.} at § 275.204A-1(a)(4) (emphasis added). Hedge Funds must further include in their code of ethics a standard of conduct that reflects the firm’s and employees’ fiduciary standards, the disclosure by employees and the review by supervisors of all “personal securities transactions,” and provisions requiring the written confirmation by employees confirming the receipt of the code. \textit{Id.} §§ 275.204A-1(a).
\textsuperscript{85} \textit{Id.} at § 275.204–2(a)(12)(ii).
\textsuperscript{86} \textit{Id.} at § 275.206(4)–7(c).
policies and procedures that are “reasonably designed to prevent violation[s] . . . of the [Investment Advisers] Act.”

B. Dodd-Frank Whistleblower Provisions

Federal securities law does not extend protection to employees of private financial companies who make internal disclosure of a possible securities law violation. Dodd-Frank altered the scheme of incentives for would-be whistleblowers, but it failed to add protections for certain internal disclosures. In addition to augmenting the registration requirements for private investment advisers, Section 922 of Dodd-Frank added Section 21(F), titled “Securities Whistleblower Incentives and Protections,” to the Securities and Exchange Act 1934. This new provision builds upon and incorporates existing whistleblower protections promulgated by Sarbanes-Oxley (SOX). Under SOX, which passed in 2002 in the wake of the collapse of publicly traded firms like Enron and WorldCom, employees of a publicly traded company are protected against retaliatory actions taken by their employer for disclosing potential fraud. SOX protects employees whether they externally disclose the perceived fraud to federal authorities or internally report the issue to their immediate supervisors. Finally, SOX provides a

---

87 Id. at § 275.206(4)-(7)(a).
89 Joel D. Hesch, Whistleblower Rights and Protections: Critiquing Federal Whistleblower Laws and Recommending Filling in Missing Pieces to Form a Beautiful Patchwork Quilt, 6 Liberty U. L. Rev. 51, 105 (2011) (“In many ways, the Dodd-Frank Act supersedes SOX because it expands protections.”).
93 See 18 U.S.C. § 1514(a) (“No company [that is publicly traded] . . . may discharge, demote, suspend, threaten, harass, or in any other manner discriminate against an employee . . . because of any lawful act done by the employee—(1) to provide information . . . regarding any conduct the employee reasonably believes constitutes a
private cause of action to whistleblowers that incur retaliatory action; those who bring suit are eligible to recover compensatory damages, including reinstatement, back pay with interest, and compensation for any special damages.\textsuperscript{94}

Unlike SOX's whistleblower protections, Dodd-Frank's provisions incentivize whistleblowing-employees in both publicly traded and private companies.\textsuperscript{95} Under Dodd-Frank, an individual is a whistleblower if she provides information to the SEC about a possible violation of securities laws that "has occurred, is ongoing, or is about to occur."\textsuperscript{96} The SEC interprets the term "possible violation" as embodying a requirement that the violation be potentially actionable.\textsuperscript{97} Therefore, the information provided to the SEC cannot be frivolous but must have a "facially plausible relationship to some securities law violation."\textsuperscript{98}

Dodd-Frank's whistleblower provisions also extend anti-retaliation protection to whistleblowers.\textsuperscript{99} An employer may not take any retaliatory action—including threats, demotion, harassment, termination, or suspension—if a whistleblower provides to the SEC information about a possible securities law violation; assists the SEC, through testifying or helping in its or a judicial investigation; or makes "disclosures that are required or protected under [SOX, the Securities Exchange Act of 1934, including 10A(m) of such Act.] and any other law, rule or regulation subject to the jurisdiction of the Commission."\textsuperscript{100}

A close reading of the anti-retaliatory provisions indicates that unless one works for a publicly traded company—and is thus protected under SOX's anti-retaliatory measures—one must report the possible securities law violation to the SEC.\textsuperscript{101} The SEC rules stipulate that those who seek to avail themselves of Dodd-Frank's retaliatory protection must have a reasonable belief, "one that a similarly situated employee might reasonably possess,"\textsuperscript{102} that a possible securities law violation has,\textsuperscript{103}

\begin{footnotes}
\item[94] See 18 U.S.C. §§ 1514A(e)(2).
\item[96] 17 C.F.R. § 240.21F–2 (2012).
\item[98] \textit{Id.} at 13.
\item[99] See 15 U.S.C. § 78u-6(b)(1)(A); 17 C.F.R. § 240.21F–2(b)(iii).
\item[101] Implementation of the Whistleblower Provisions, \textit{supra} note 17, at 18 ("[T]he retaliation protections for internal reporting afforded by [15 U.S.C. § 78u-6(h)(1)(A)] do not broadly apply to employees of entities other than public companies."); \textit{see also Whistleblower Incentives, \textit{supra} note 88, at 1834.}
\item[102] Implementation of the Whistleblower Provisions, \textit{supra} note 17, at 16.
\end{footnotes}
is, or will occur and they must provide the information to the SEC through certain prescribed methods.\footnote{103}{17 C.F.R § 240.21F-2(b)(ii); see “Procedures for Submitting Original Information” 17 C.F.R. § 240.21F-9 (a whistleblower must either submit their information online through the SEC’s website or send the SEC, via mail or fax, a Form TCR (Tip, Complaint, or Referral)).} However, so long as whistleblowers do not offer a frivolous tip, they are protected under Dodd-Frank’s anti-retaliatory provisions regardless of whether the submitted information leads to a successful SEC enforcement action.\footnote{104}{17 C.F.R. § 240.21F–2(b)(iii).} Like SOX, Dodd-Frank extends a private cause of action to whistleblowers that suffer retaliatory action by their employers.\footnote{105}{15 U.S.C. § 78u-6(b)(B)(i).}

In addition to extending anti-retaliatory protection to whistleblowers, Dodd-Frank’s whistleblower provisions include significant monetary incentives to encourage disclosure of possible violations.\footnote{106}{See id. at § 78u-6(b).} Lawmakers “[r]ecogniz[ed] that whistleblowers often face the difficult choice between telling the truth and the risk of committing ‘career suicide’.”\footnote{107}{S. Rep. No. 111-176, at 111 (2010); see also Paul Sullivan, \textit{The Price Whistleblowers Pay For Secrets}, N.Y.TIMES, Sept. 21, 2012, \textit{available at} http://www.nytimes.com/2012/09/22/your-money/for-whistle-blowers-consider-the-risks-wealth-matters.html.} If one provides original information\footnote{108}{“An action brought under a statute that allows a private person to sue for a penalty, part of which the government or some specified public institution will also receive.” \textit{BLACK’S LAW DICTIONARY} 1368 (9th ed. 2009).} about a possible securities law violation to the SEC that, in turn, leads to a successful enforcement action yielding for the SEC a recovery of over one million dollars, the whistleblower who provided the information is entitled to receive a\footnote{109}{\textit{15 U.S.C. § 78u-6(a)(3).}} \textit{qui tam} award between 10\% and 30\% of the total amount recovered by the SEC.\footnote{110}{15 U.S.C. § 78u-6(b).} These large awards are important, because they are “what it takes to have an employee risk everything.”\footnote{111}{Sullivan, supra note 107 (quoting Stephen M. Kohn, whistleblower attorney).} Whistleblowers do not have to provide the damning information to their employer to be eligible for the award. Instead, one’s participation in an internal compliance program is taken into account as a factor by the SEC when deciding whether to increase a whistleblower’s total award.\footnote{112}{17 C.F.R. § 240.21F-6(a) (2012). Other factors that may affect the size of the bounty include the significance of the information provided by the whistleblower, other factors that may affect the size of the bounty include the significance of the information provided by the whistleblower, other factors that may affect the size of the bounty include the significance of the information provided by the whistleblower, other factors that may affect the size of the bounty include the significance of the information provided by the whistleblower, other factors that may affect the size of the bounty include the significance of the information provided by the whistleblower, other factors that may affect the size of the bounty include the significance of the information provided by the whistleblower, other factors that may affect the size of the bounty include the significance of the information provided by the whistleblower, other factors that may affect the size of the bounty include the significance of the information provided by the whistleblower, other factors that may affect the size of the bounty include the significance of the information provided by the whistleblower, other factors that may affect the size of the bounty include the significance of the information provided by the whistleblower.}
To further incentivize whistleblower participation in their employer’s internal compliance apparatus, the SEC’s rules include a “look back provision,” which retroactively applies the date the whistleblowers informed their compliance department as the date of the SEC submission.\textsuperscript{113}

Whether an employee who internally discloses a potential securities law violation is protected under federal law depends on the nature of the company for whom the employee works. In this context, an employee of a privately owned hedge fund like Sullivan is left without any federally prescribed recourse.

III. CONFLUENCE OF FEDERAL AND STATE POLICY HELP NEITHER EMPLOYEES NOR EMPLOYERS

A. Policy Juxtaposition—New York and Dodd-Frank

The intersection of New York law and Dodd-Frank’s whistleblowing provisions leave those who work for New York-based private investment advisers with few options to safely disclose a possible securities law violation.\textsuperscript{114} To receive federal protection, one must first statutorily qualify. In the case of SOX, one must be an employee of a publicly traded company.\textsuperscript{115} Under Dodd-Frank, employees of a private investment adviser must externally disclose an impropriety to the SEC to avail themselves of the law’s anti-retaliatory protection.\textsuperscript{116}

Further, neither New York’s whistleblower statute nor its at-will employment doctrine protect employees who make disclosures regarding possible securities law violations from retaliatory conduct.\textsuperscript{117} As noted above, New York’s courts typically

\textsuperscript{113} Implementation of the Whistleblower Provisions, supra note 17, at 89-90 (“[A] whistleblower who first reports to an entity’s internal whistleblower, legal, or compliance procedures for reporting allegations of possible violations of law and within 120 days reports to the Commission could be an eligible whistleblower whose submission is measured as if it had been made at the earlier internal reporting date.”).

\textsuperscript{114} See \textit{supra} Part I.A-B.

\textsuperscript{115} See 18 U.S.C. § 1514A (2012). SOX’s whistleblower covers employees of publicly traded companies, its contractors, subcontractors and “any subsidiary or affiliate whose financial information is included in [the company’s] consolidated financial statements.” \textit{Id}.

\textsuperscript{116} See Egan v. TradingScreen, Inc., 10 CIV. 8202 LBS, 2011 WL 1672066, at *7 (S.D.N.Y. May 4, 2011). (“[U]nless the statutory exceptions apply[,] the anti-retaliation whistleblower protection provisions of the Dodd-Frank Act require Plaintiff to show that he . . . provided information to the SEC . . . .”)\textsuperscript{116}; see also Whistleblower Incentives, \textit{supra} note 88, at 1834.

\textsuperscript{117} See \textit{supra} Part I.
analyze at-will employment disputes under a contractual framework. Unless the employer has borne an express obligation to refrain from arbitrarily terminating an employee or, and in only very narrow circumstances, there exists some type of implied obligation of self-regulation, "an employer’s right at any time to terminate an employment at will remains unimpaired." New York’s whistleblower statute, too, provides little comfort to those who detect, and seek to internally disclose, a securities law violation. Courts sustain actions brought under New York’s whistleblower statute only in a context of protecting the public from tangible dangers, such as health and safety emergencies. Even though financial scandals will continue to threaten the economic health of New Yorkers, New York’s whistleblower statute does not apply to protect those that disclose financial or securities-related infractions.

Without any state protections, New York employees like Sullivan—those who work for a private company in the financial services industry—must look to federal law for protection. “Dodd-Frank has created a two-tiered structure of protections where potential whistleblowers receive different sets of protections depending on whether they chose to report internally or externally.” Employees of publicly traded companies—those subject to SOX—who disclose possible securities law violations receive anti-retaliatory protection regardless of whether the recipient of the disclosure is the whistleblower’s boss or the SEC. Employees of private companies that are not subject to SOX, such as Peconic, however “receive no protection if they report internally.”

121 See, e.g., Clarke v. TRW, Inc., 921 F. Supp. 927, 931, 935 (N.D.N.Y. 1996) (noting, among other things, that a defective electrical relay that could affect car’s brakes or fuel pump constituted “a specific and substantial risk to the health and safety of the public”); Liu, supra note 66, at 71.
123 Whistleblower Incentives, supra note 88, at 1834.
124 Id.
125 Id.
B. Confluence of State and Federal Policies May Not Align with Whistleblower Motives

Internal disclosure should be promoted and protected. Not only is internal disclosure arguably easier for an employee, it is often the first course of action for those that detect that something may be wrong. The confluence of federal and state policies—the “all roads point to the SEC” framework—may not align with the best interests and motivations of most whistleblowers. Although a significant number of whistleblowers state that they incurred retaliatory actions by their employer, including termination in many instances, the majority of whistleblowers who ultimately pursue *qui tam* bounties, such as the one offered under Dodd-Frank, first reported the matter internally to their employer. These employees generally do not rely on employer hotlines, where confidentiality is preserved, but instead report a problem to their supervisor. In fact, one study suggests that only one in six employees who first reported a matter internally decided to later disclose the potential violation to regulators. The same study suggests that only three percent of whistleblowers pursue external disclosure as a first course of action.

This reluctance to publicly “blow the whistle” may be understandable as public disclosure can carry with it a heavy financial and emotional cost. Even after the passage of SOX, approximately 40% of employees in one survey stated that they remained silent upon detecting a problem, opting not to share the matter with anyone in their firm. This behavior may


127 Stephen Martin Kohn, Amended Remarks, *The Impact of Qui Tam Whistleblower Rewards On Internal Compliance*, in Michael D. Greenberg, *For Whom the Whistle Blows: Advancing Corporate Compliance and Integrity Efforts in the Era of Dodd-Frank* (RAND, 2011). Kohn, Executive Director of the National Whistleblower Center, noted that “[i]n cases under the False Claims Act . . . [e]mpirical data show that approximately 90 percent of employees who filed a *qui tam* case initially reported their concerns internally.” *Id.*

128 *Id.* at 42. Kohn, Executive Director of the National Whistleblower Center, cites a report from the Ethics Resource Center. *Id. See Blowing the Whistle on Workplace Misconduct*, ETHICS RESOURCE CENTER (Dec. 2010), http://www.ethics.org/files/u5/WhistleblowerWP.pdf.


130 *Id.* at 19.

131 Sullivan, *supra* note 107 (collecting opinions of prominent whistleblower attorneys).

make sense as a different study questioning financial service professionals indicated that only 35% of respondents believed that their firm would not take retaliatory action.\textsuperscript{135} Additionally, the reward one may eventually receive through a bounty program and a successful lawsuit may not fully compensate the whistleblower.\textsuperscript{134} In addition to lost wages from being fired, whistleblowers will become pariahs in their field and will likely struggle to find similar work.\textsuperscript{135}

Although Joseph Sullivan limited disclosure to Peconic’s CEO,\textsuperscript{136} his rationale for doing so may not be representative of the would-be whistleblower. At the time of the alleged trading improprieties, Sullivan and Harnisch were embroiled in an ownership dispute regarding a partnership agreement that “would have eliminated Sullivan’s ownership interest” in Peconic.\textsuperscript{137} Harnisch fired Sullivan the same day Sullivan’s attorney contacted Peconic to discuss the agreement.\textsuperscript{138} Although this context may have colored the court’s impression of Sullivan’s claim, the fact remains that an employee of a private investment adviser that limits disclosure to an internal audience is not protected against employer retaliation.

C. Confluence of State and Federal Polices May Not Align with Employers’ Interests

In addition to failing to protect the natural behavior of the common employee, neither federal nor state whistleblower policies adequately align with employers’ interests. While Peconic may have won the litigation battle in \textit{Sullivan v. Harnisch}, the ruling may cause private investment advisers in New York to lose ground in the compliance war. Many commentators and practitioners of the corporate bar bemoaned the SEC’s final whistleblower rules as creating a system that overly incentivizes whistleblowers to make external disclosure and, at the same time, one that undermines a firm’s ability to self-policing.\textsuperscript{139} Although these points of concern arose

\begin{quote}

\textsuperscript{134} \textit{Sullivan, supra} note 107.

\textsuperscript{135} \textit{Id}.


\textsuperscript{137} \textit{Id}. at 759.

\textsuperscript{138} \textit{Id}.

\textsuperscript{139} \textit{See, e.g.}, Letter from Cleary Gottlieb Stein to Sec. & Exch. Comm’n, Re Proposed Rules For Implementing the Whistleblower Provisions of Section 21F of the Securities Exchange Act of 1934—File No. S7-33-10 (Dec. 17, 2010), \textit{available at}
in a national debate concerning the implementation of federal legislation, they remain equally potent—and perhaps the predicted effects by these commentators become compounded—when assessed under a local, New York-based lens.

With the release of the final rules implementing Dodd-Frank’s whistleblower provisions, the SEC believes that it “charted a Solomonic middle ground in an attempt to prevent harming companies’ internal compliance programs while also not putting up too many obstacles for whistleblowers to overcome.”\(^{140}\)

While noting the value of internal compliance programs, the SEC also observed that, at times, these programs “cannot serve as adequate substitutes for our obligation to identify and remedy violations of the federal securities laws.”\(^{141}\) The whistleblower program, according to the SEC, the whistleblower program “encourages the whistleblower to report allegations internally, yet ultimately and appropriately leaves that decision to the whistleblower.”\(^{142}\)

Companies, as can be imagined, disagree with the SEC’s final rules. Many fear that the bounty program, which grants an award to the whistleblower of up to 30% of any SEC recovery exceeding one million dollars,\(^ {143}\) will inhibit “well-meaning, compliant” firms from investigating and rectifying compliance issues internally.\(^ {144}\) “The ability to bypass pre-existing internal procedures for reporting wrongdoing also may limit the role of internal investigations in future enforcement actions by inhibiting companies from conducting such investigations before the government becomes involved.”\(^ {145}\) Others believe that internal disclosure should not be a factor that augments an award but, rather, should be “a prerequisite to recovery by an employee-whistleblower.”\(^ {146}\) Without setting internal disclosure as a precondition to recovery, companies fear that the “bounty


\(^{143}\) *Id*.


\(^{145}\) *Id*.

\(^{146}\) Letter from Cleary Gottlieb Stein, *supra* note 139.
program will . . . drive companies to view all compliance matters equally—as matters of potentially severe consequence, given the increased risk of regulatory intervention.”\textsuperscript{147} Not only will this lead to an inefficient allocation of compliance resources, as companies would have to probe a potentially catastrophic matter with the same vigor as a frivolous issue, it may also compel companies to self-report to the SEC all matters to avoid its employees from whistleblowing, thereby flooding the SEC with problems and unnecessarily tarnishing the reputation of firms.\textsuperscript{148}

Although the current Dodd-Frank whistleblowing paradigm encourages external whistleblowing, the lack of state protection for internal disclosure demands external disclosure from employees of New York private investment advisers. The problem is not that Dodd-Frank over-incentivizes would-be whistleblowers, as some commentators have suggested,\textsuperscript{149} but rather that, at least in New York, public disclosure is the only way in which an employee of this firm type can guarantee legal protection against retaliatory action. This confluence of state and federal policies places undue pressure on employees who seek to report a potential violation and deprives employers of a meaningful opportunity to self-police securities compliance.

IV. REMEDIES

Under the current paradigm created by the confluence of Dodd-Frank and New York’s policies, those employed by New York private investment advisers with knowledge of a potential securities law violation are stuck between a rock and a hard place. To preserve their jobs, these employees face the choice of doing nothing, which may cause a problem to fester, or taking the risky path of disclosing a violation to the SEC. Changes to either New York or federal policies are necessary to efficiently balance the interests of employees, employers, and society. Fortunately, there are several possible remedies available to companies, policy makers, and the courts. The menu of available cures is divided into two parts. First, there are private remedies; steps that firms can take to induce and promote internal disclosure. Second, there are public remedies; measures policymakers and the courts can pursue to extend legal

\textsuperscript{147} Id.

\textsuperscript{148} Id.

protections to those who, but for the fear of retaliatory action, seek to internally disclose possible securities law violations.

A. Private Remedy: Establishing Strong Compliance Policies

In light of Dodd-Frank’s significant incentives to externally disclose possible securities law violations, many commentators and practitioners recommend that firms establish strong compliance programs that provide employees with “a visibly safe pathway” for internal disclosures.\textsuperscript{150} “The occurrence of external whistleblowing . . . usually indicates not only the failure in a firm’s commitment to morality but also a breakdown in its ethical structure and communication channels.”\textsuperscript{151} Practitioners recommend that firms either establish written policies that proscribe retaliatory action or provide employees with a means to confidentially relay information regarding a possible problem, and, in doing so, create a disclosure infrastructure that makes it impossible for an employer to take retaliatory action.\textsuperscript{152}

The adoption of internal policies that facilitate confidential internal disclosure and or proscribe retaliatory action is a strong solution. These policies both align with whistleblower behavior and an employer’s interest in maintaining strong compliance programs. Although this remedy is attractive, it is not sufficient. Widespread implementation of corporate policies that protect internal disclosure cannot be guaranteed. The adoption of internal policies, by its very definition, is firm dependent. Society’s interest in compliant investment advisers may be better served by the promulgation of a uniform floor of legal protection upon which employees can stand, and not through ad-hoc policies that likely vary among different firms.

\textsuperscript{150} Michael D. Greenberg, Corporate Integrity in the Wake of Dodd-Frank: How Do We Fortify Internal Compliance, Reporting, and Culture? in FOR WHOM THE WHISTLE BLOWS: ADVANCING CORPORATE COMPLIANCE AND INTEGRITY EFFORTS IN THE ERA OF DODD-FRANK 23 (RAND, 2011) [hereinafter Corporate Integrity].

\textsuperscript{151} Cavico, supra note 126, at 623.

\textsuperscript{152} See Corporate Integrity, supra note 150, at 23–24.
B. Public Remedies

1. Update Dodd-Frank

Perhaps the clearest cut remedy to protect internal disclosure is to amend Dodd-Frank. Indeed, as some commentators have noted, the inconsistency between SOX and Dodd-Frank regarding anti-retaliatory protections for those who internally disclose may be unintentional.153 This theory, however, is likely misplaced, as the SEC definitively stated: “[T]he retaliation protections for internal reporting afforded by [15 U.S.C. § 78u-6(h)(1)(A)(iii)] do not broadly apply to employees of entities other than public companies.”154 The SEC has further signaled its intent that disclosure should be publicly made, noting, “[I]nternal compliance programs are not substitutes for rigorous law enforcement.”155

Courts, too, have recognized the bifurcation in anti-retaliatory protection. Recent judicial interpretations of this area of Dodd-Frank indicate that courts will not interpret the new law so that anti-retaliatory protection extends to employees of private companies who pursue internal disclosure.156 In the first case to review Dodd-Frank’s SEC whistleblower provisions,157 the court compared the SEC’s whistleblower program and scope with a different whistleblower program promulgated by Dodd-Frank for the Bureau of Consumer Financial Protection (CFPB).158 The court noted that the CFPB’s anti-retaliation provisions159 were

153 See, e.g., Hesch, supra note 89, at 105–06.
154 Implementation of the Whistleblower Provisions, supra note 17, at 18.
155 Id. at 97.
159 The CFPB’s anti-retaliation provisions extend protection to those who disclose possible violations under the agency’s jurisdiction “to the employer, the Bureau, or any other State, local, or Federal, government authority or law enforcement agency.” 12 U.S.C. § 5567(a)(1); see also Egan, 2011 WL 1672066, at *4.
broader than the SEC’s whistleblower provisions, “indicat[ing] that Congress intended to encourage whistleblowers reporting [securities] violations to report to the SEC.”

One commentator has argued that amending Dodd-Frank to incorporate SOX’s anti-retaliation protection for internal disclosure within private firms “could certainly be one of the most important [amendments aimed toward] fulfilling the investor protection mission of Dodd-Frank and of the SEC in general.” In fact, one proposed amendment to the SEC’s whistleblower program has already been introduced before Congress. New York Representative Michael Grimm introduced the Whistleblower Improvement Act of 2011. Under this proposed legislation, whistleblowers must first disclose a possible violation to their employer in order to maintain eligibility for an award under Dodd-Frank’s bounty program. Grimm’s bill includes an exception to this default rule and allows for recovery of an award without internal disclosure when “the whistleblower alleges and the Commission determines that the employer lacks either a policy prohibiting retaliation for reporting potential misconduct or an internal reporting system allowing for anonymous reporting.”

This legislation has its sights fixed on Dodd-Frank’s bounty program. The proposed bill, however, does not adequately address the current disconnect among Dodd-Frank’s and SOX’s anti-retaliatory protections. In effect, this proposal is an unfair exchange as employers benefit at the expense of their employees. The bill’s internal disclosure requirement addresses firm’s concerns about the strength of their compliance programs but the bill does not provide something of equal value, like protection against retaliation, to employees.

---

161 Whistleblower Incentives, supra note 88, at 1836 (internal quotation marks omitted).
163 Id.
164 Id. at § 2(a)(2)(A).
165 Id. at § 2(a)(2)(D).
166 See id. at § 2(b)(1–2) (eliminating a minimum award amount and amending Dodd-Frank’s language such that the maximum award is capped but the minimum amount is left to the SEC’s discretion); see also Dana Liebelson, New Bill to Weaken Protections, Incentives for Whistleblowers Sneaks Through Committee, TRUTH-OUT.ORG (Feb. 16, 2012), http://truth-out.org/news/item/6721:new-bill-to-weaken-protections-incentives-for-whistleblowers-sneaks-through-committee.
167 See Whistleblower Improvement Act of 2011, H.R 2483, 112th Cong. § 2. The proposed legislation allows for external disclosure as a first course of action where the employer does not have a policy proscribing retaliation but the proposed legislation does not proscribe employer retaliation for internal disclosure. Id.
2. Adapt New York Common Law

New York does not recognize a common law tort claim for wrongful discharge, or a public policy exception to the at-will employment doctrine. Where an employee is at-will, “absent a constitutionally impermissible purpose, a statutory proscription, or an express limitation in the individual contract of employment, an employer’s right at any time to terminate an employment at will remains unimpaired.” New York courts, however, should recognize the effect the juxtaposition of its at-will employment doctrine and Dodd-Frank’s whistleblower provisions have on employees and their respective employer’s compliance program. Further, crafty plaintiffs should recognize that the current jurisprudence of New York’s at-will employment doctrine, when juxtaposed with the new requirements that Dodd-Frank imposes upon private investment advisers, might afford a means of obtaining protection from retaliatory termination.

As noted above, New York has carved one exception to its strict at-will employment doctrine when it recognized a claim sounding in breach of contract “based on an implied-in-law obligation” in an action brought by an attorney against his employer law firm. The exception, espoused in Wieder v. Skala, is founded, in part, upon the principle that attorneys are a self-regulating industry and are subject to a code of professional responsibility. “Erecting . . . disincentives to compliance with the applicable rules of professional conduct . . . would subvert the central professional purpose of [an attorney’s] relationship with the firm—the lawful and ethical practice of law.”

---

172 Id. at 108 ("The particular rule of professional conduct implicated here (DR 1-103[A]), it must be noted, is critical to the unique function of self-regulation belonging to the legal profession."). DR 1-103[A], now codified at N.Y. COMP. CODES R. & REGS. tit. 22, § 1200.0 rule 8.3(a) (2012), states: “A lawyer who knows that another lawyer has committed a violation of the Rules of Professional Conduct that raises a substantial question as to that lawyer’s honesty, trustworthiness or fitness as a lawyer shall report such knowledge to a tribunal or other authority empowered to investigate or act upon such violation.”
173 Wieder, 609 N.E.2d at 108.
Sullivan v. Harnisch, Chief Judge Lippman argued that the Wieder exception should apply to a hedge fund’s chief compliance officer. Just as in Wieder, Sullivan was “an employee of a business that was subject to certain legal and ethical obligations to its clients and his reason for being, as a compliance officer, was to ensure that in providing services to those clients, those rules were followed at all times.” If applied, the application of Judge Lippman’s extension of the Wieder exception would cover only those that were employed as compliance personnel for a private investment adviser. And while the implementation of such a policy would be a step in the right direction, it does not go far enough as it leaves all other non-compliance employees without meaningful protection against retaliatory conduct.

There may yet be hope for protecting non-compliance personnel from retaliatory conduct. New York also recognizes an action sounding in breach of express contract in cases involving at-will employees where an employer’s handbook or written policies prohibit arbitrary termination. In Weiner v. McGraw Hill, the New York Court of Appeals held that a discharged at-will employee had put forward a sufficient breach of contract claim because an express obligation in the company’s handbook prevented the employer from terminating him without sufficient cause and “only after all practical steps toward rehabilitation or salvage of the employee have been taken and failed.”

If the principles of self-regulation and employee obligation embedded in Wieder are combined with the contractual principles of Weiner, New York courts may be able to devise a way to extend anti-retaliatory protection for employees of private investment advisers while remaining generally adherent to New York’s current at-will employment doctrine. Among other things, Dodd-Frank requires newly registered private investment advisers to create a code of ethics that must have “[p]rovisions requiring [the firm’s] supervised persons to comply with applicable Federal securities laws” and “[p]rovisions requiring supervised persons to report any violations of your code of ethics promptly to your chief compliance officer or, provided your chief compliance officer also

---

175 Id. at 764.
176 Id. (“[W]here an employee is merely peripherally responsible for informing his or her employer (or others) of violations of certain obligations, that person is unlikely to be covered by the Wieder exception. This is not such a case.”).
178 Id. at 442.
receives reports of all violations, to other persons you designate in
your code of ethics.”179 In effect, these provisions place a duty on
private investment advisers to craft policies that require their
employees to internally report possible securities law violations.
An employee should be able to point to these provisions and argue
that they were obligated, pursuant to an express condition in
their code of ethics, to make internal disclosures.

Because, however, the provisions outlining the
requirements of a private investment adviser’s code of ethics do
not require a firm to put in place a policy proscribing retaliation
for internal disclosure, the court may not be satisfied that an
express obligation exists against which a breach of contract can be
established. “As has been observed, courts should not ‘infer a
contractual limitation on the employer’s right to terminate an at-
will employment absent an express agreement to that effect which
is relied upon by the employee.’”180 For example, in Lobosco v. New
York Tel. Co./NYNEX the court dismissed a breach of contract
claim raised by a terminated employee where the employer had a
code of conduct that both required employees to internally disclose
misconduct and provided assurances that retaliatory action would
not be taken as a result of disclosure.181 The same handbook,
however, contained an express disclaimer stating that the code of
conduct could not be interpreted as a contract.182

It is here, critically, where the courts should impute the
policy principles of Wieder. In Wieder, the court placed significant
weight on the idea that attorneys are subject to a code of
professional conduct.183 Both New York’s Rules of Professional
Conduct discussed in Wieder184 and the registration requirements
of Dodd-Frank impute a duty upon those that practice in their
respective fields to disclose possible violations. Unlike the policy
in Lobosco, the policies referenced in Wieder and required by
Dodd-Frank are universally applied across a profession. And
unlike a failure to adhere to a private corporate policy in
Lobosco, the failure by attorneys or securities professionals to

179 17 C.F.R. § 275.204A-1(a) (2012). The Sullivan court did not focus on these
requirements but instead focused its attention on 17 C.F.R. § 275.206(4)-7, which
requires companies to appoint chief compliance officers and to create policies to prevent
violations of securities laws. Sullivan, 969 N.E.2d at 761.
180 Sullivan v. Harnisch, 915 N.Y.S.2d 514, 518 (N.Y. App. Div. 2010), aff’d,
(N.Y. App. Div. 2003)).
182 Id.
184 Id.
abide by their respective universal policies could have a wide-ranging, systemic impact.

Employees of private investment advisers are required to abide by their code of ethics and disclose to their colleagues potential securities law violations. Incurring retaliation for this adherence, consequently, could be interpreted as an action taken by one party to a contract to frustrate the counterparty’s ability to satisfy its contractual obligations. As the Wieder court stated: “It is the law that in ‘every contract there is an implied undertaking on the part of each party that he will not intentionally and purposely do anything to prevent the other party from carrying out the agreement on his part.”185 Such conduct, as the Wieder court held, could constitute a breach of an implied contract.186

Although this litigation strategy likely would receive a cold reception before New York courts,187 it is worth noting that an approach that borrows from and bends existing jurisprudence is not as radical as some may fear. The strategy of applying both the Wieder exception and existing contract jurisprudence to sustain a claim of wrongful termination is inherently circumscribed by the unique characteristics arising from the obligations imposed by Dodd-Frank upon private investment advisers. Recognizing that this new approach would arise only in the context of retaliation for internal disclosure, private investment advisers may be willing to forfeit a portion of their unencumbered right to discharge employees because the adoption of this policy may further their own compliance goals. As home to half of the world’s 20 largest hedge funds,188 a New York policy that promotes and protects internal disclosure may be in the best interest of employees, employers, and the State.

---

185 Id. at 109 (quoting Patterson v. Meyerhofer, 97 N.E. 472, 473 (N.Y. 1912)).
186 Id.
187 “American courts, including our own, have proved chary of creating common-law exceptions to [the at-will] rule and reluctant to expand any exceptions once fashioned.” Sullivan v. Harnisch, 969 N.E.2d 758, 760 (N.Y. 2012) (quoting Horn v. N.Y. Times, 790 N.E.2d 753, 755 (N.Y. 2003)). The majority in Sullivan also noted that “the existence of federal regulation furnishes no reason to make state common law governing the employer-employee relationship more intrusive.” Id. at 761.
CONCLUSION

The financial crisis of 2008–2009 triggered the worst recession since the Great Depression.\(^\text{189}\) Between October 2007 and March 2009, the world watched as U.S. stock prices halved.\(^\text{190}\) With a goal of promoting stability within the U.S. financial system, Congress responded and passed the Dodd-Frank Wall Street Reform and Consumer Protection Act in 2010.\(^\text{191}\) Among the many amendments Dodd-Frank incorporated into federal securities laws, Title IV of the statute required most private investment advisers like hedge funds and private equity funds to register under the Investment Advisers Act of 1940.\(^\text{192}\) Dodd-Frank did not, however, extend SOX’s protections against retaliatory action for internal disclosure to employees of private investment advisers.\(^\text{193}\)

Under its current jurisprudence, however, New York courts may be able to formulate a method to protect against retaliatory action for employees of private investment advisers that make internal disclosure. This framework would combine the policy of industry-wide self-regulation espoused in Wieder v. Skala\(^\text{194}\) with the contractual analysis found in Weiner v. McGraw-Hill.\(^\text{195}\) The application of either case’s precedent alone is insufficient to safeguard internal disclosure; the compliance obligations imposed by Dodd-Frank upon employees of private advisers neither represent “the very core”\(^\text{196}\) of their employment nor do they create a “promise not to discharge . . . without . . . sufficient cause.”\(^\text{197}\) By applying both Wieder and Weiner simultaneously, however, the shortcomings of either may be digestible to a court that is wary of changing its at-will employment doctrine.\(^\text{198}\) This adaptation would


\(^{193}\) Implementation of the Whistleblower Provisions, supra note 17, at 18.


\(^{196}\) Wieder, 609 N.E.2d at 108.

\(^{197}\) Weiner, 443 N.E.2d at 445.

benefit both employer and employee. By erecting a legal floor upon which employees of private investment advisers can stand, New York courts can offset the incentives pulling these employees to make external disclosure. Increased internal disclosure, in turn, will improve an employer’s ability to monitor its compliance with federal securities laws.

John H. Runne†

† J.D. Candidate, Brooklyn Law School, 2014. I wish to thank my colleagues at the Brooklyn Law Review for their careful review and insightful edits. I also wish to thank my wife Amanda for her unconditional love and support.
Following Suit with the Second Circuit

DEFINING GAMBLING IN THE ILLEGAL GAMBLING BUSINESS ACT

INTRODUCTION

The federal government’s involvement in the regulation of gambling has been a demonstration of the principles of federalism. For the most part, the federal government has recognized that gambling is an area of law left to the prerogatives of the states.Absent the influence of federal law, state gambling law has developed to be extremely varied, not just from state to state, but from game to game within a state. Much of that variety is due to differing definitions of gambling, and how large a role chance, as opposed to skill, takes in gambling. The federal government has regulated gambling but it has not inserted itself into substantively defining gambling. Rather, the federal government has intervened when states have inadequately enforced their own laws. The overarching federal policy on gambling is to aid the states in enforcing what they define as gambling, rather than the federal government determining its own definition of gambling. This note argues that the Illegal Gambling Business Act (IGBA) should be read in light of this federal policy.

The IGBA prohibits “conduct[ing], financ[ing], manag[ing], supervis[ing], direct[ing], or own[ing] all or part of an illegal gambling business.” The statute then continues to define what constitutes an illegal gambling business in § 1955(b)(1), which enumerates three elements:

2 See infra Part I.
4 Blakey & Kurland, supra note 1, at 925; S. REP. No. 91-617, at 74 (1969); see also infra Parts II–III.
(1) “illegal gambling business” means a gambling business which—

(i) is a violation of the law of a State or political subdivision in which it is conducted;

(ii) involves five or more persons who conduct, finance, manage, supervise, direct, or own all or part of such business; and

(iii) has been or remains in substantially continuous operation for a period in excess of thirty days or has a gross revenue of $2,000 in any single day.6

A recent decision by the U.S. District Court for the Eastern District of New York, United States v. Dicristina, rejected the argument that the definition of an illegal gambling business is limited to the elements listed in § 1955(b)(1).7 The court in Dicristina held that § 1955(b)(2) added a fourth element to the crime. Section 1955(b)(2), which reads, reads “‘gambling’ includes but is not limited to pool-selling, bookmaking, maintaining slot machines, roulette wheels or dice tables, and conducting lotteries, policy, bolita or numbers games, or selling chances therein,” added a fourth element to the crime.8 The court interpreted § 1955(b)(2) as requiring the government to prove that in the type of game played (in this case Texas Hold ’Em), chance predominated over skill.9 In other words, in addition to the type of game violating state law, the game also had to fulfill the federal definition of gambling gleaned from § 1955(b)(2), that chance predominates skill, in order to prove a violation of § 1955. The court’s extensive statutory interpretation analysis found the plain language and legislative history to be inconclusive regarding the meaning of § 1955(b)(2) and held that, due to the rule of lenity, the defendant’s narrower interpretation must prevail.10 The court dismissed the indictment, holding that skill predominated over chance in Texas Hold ’Em.11

The Second Circuit overruled the district court’s reading that § 1955(b)(2) constitutes a fourth element and held that to be guilty of operating a gambling business in violation of federal

---

6 Id. § 1955(b).
8 Id. at 224-25, 231; 18 U.S.C. § 1955(b)(2).
9 Dicristina, 886 F. Supp. 2d at 231. This is in addition to proving that the game violates state law, which may adopt a test other than the predominance test.
10 Id. at 235. “[T]he rule of lenity requires that ambiguous criminal laws be interpreted in favor of the defendants subjected to them.” Id. at 200 (internal citations omitted).
11 Id.
law, the prosecution only has to prove the three elements articulated in § 1955(b)(1). The Second Circuit relied on the plain meaning of the statute, reading § 1955(b)(2)’s phrase “including but not limited to” as introducing a non-exhaustive list of the types of gambling businesses that violated § 1955, and not as substantively defining gambling. The court also reasoned that this reading of the plain language of the statute was bolstered by the legislative history, which was concerned with reaching gambling businesses of a certain size and character, rather than with prohibiting certain types of games. The Second Circuit held that the rule of lenity did not apply because the meaning of the statute was unambiguous. It reversed the judgment of acquittal and remanded for a reinstatement of the jury verdict.

United States v. Dicristina is the first case where a court analyzed § 1955(b)(2) in depth. Following the district court decision in Dicristina, the District Court of Guam declined to interpret § 1955(b)(2) as containing a fourth element of the crime for many of the same reasons enunciated in the Second Circuit’s opinion. Prior to Dicristina, other courts hearing cases regarding a prosecution under § 1955 assumed, with little or no analysis, that the definition of an “illegal gambling business” consisted of only those elements listed in § 1955(b)(1). This note argues that the elements listed in § 1955(b)(1) are the complete definition of “illegal gambling business,” and whether a certain type of game constitutes gambling under that statute should be determined solely by state law. Courts should follow suit with the Second Circuit’s reading of § 1955(b)(2), which is

---

12 United States v. Dicristina, 726 F.3d 92, 102 (2d Cir. 2013).
13 Id. at 98-100. The court distinguished between the use of “including but not limited to” and “means,” which is used elsewhere in the statute to define terms. It also pointed out that § 1955(e)’s language defining games of chance could have been used in § 1955(b)(2) if Congress so desired, and that there is no evidence § 1955(e) was intended to modify § 1955(b)(2). Id.
14 Id. at 102-04.
15 Id. at 104-05.
16 Id. at 106.
18 See, e.g., United States v. Atiyeh, 402 F.3d 354, 372 (3d Cir. 2005). Unlike the cases that follow, this case did not address the argument made in Dicristina. The Third Circuit found that a violation of § 1955 included conducting a gambling business as defined by the three elements in § 1955(b)(1), as distinguished from the definition of gambling provided in § 1955(b)(2). The court’s discussion, however, did not extend beyond this distinction, and the argument was not analyzed in depth. Other cases include United States v. Gotti, 459 F.3d 296, 340 (2d Cir. 2006); United States v. Truesdale, 152 F.3d 443, 446 (6th Cir. 1998); United States v. Cyprian, 23 F.3d 1189, 1199 n.14 (7th Cir. 1994); United States v. Sacco, 491 F.2d 995, 998 (9th Cir. 1974) (en banc).
consistent with the history of federal involvement in gambling. The historical practice of the federal government has been to defer to state law definitions of gambling, and the IGBA should be read in accordance with that practice.

Part I illustrates the variety among state law definitions of gambling through a case study of the popular poker game variant Texas Hold 'Em. Demonstrating that variety in Part I supports the arguments in Parts II and III, that Congress recognized and intended to defer to the states’ varying definitions of gambling. The background of the varying definitions of gambling also explain why the IGBA’s legislative history, which is devoid of any mention of the definition of gambling, should not be read to adopt one of those definitions. Part II explores the role of the federal government in gambling throughout American history, and argues that the federal government only legislates when the states cannot adequately implement their own gambling policies, and only seeks to prohibit gambling which is illegal as defined by state, not federal, law. Part III analyzes the legislative history of the IGBA, which was a part of the Organized Crime Control Act of 1970, and argues that the IGBA continued the federal tradition of aiding the states’ gambling enforcement only where the states cannot enforce their own law. It should not be read as indicating the federal government’s concern about which types of games are prohibited in a distinct federal definition of gambling. This part analyzes the language of § 1955 with the aim of showing that the statute deferred to state definitions of gambling without adding a new element in § 1955(b)(2).

I. THE VARIANCE IN STATE LAW DEFINITIONS OF GAMBLING, AS DEMONSTRATED BY THE POKER GAME VARIANT TEXAS HOLD ’EM

Poker has a long and storied history in America. It first appeared in the United States in the early nineteenth century in the port of New Orleans. The game then quickly spread by

---

19 Poker is believed to have been brought by the French settled in New Orleans, and derived from a French card game called poque. It is believed that poker also derives from the card games Bouliotte, Poch (also known as Pochen or Pochspeil), As-Nas, and Brag. The name ‘poker’ is most likely derived from non-French speaking Americans mispronouncing Poque. See Anthony Cabot & Robert Hannum, Poker: Public Policy, Law, Mathematics, and the Future of an American Tradition, 22 T. M. COOLEY L. REV. 443, 447-50 (2005); David Parlett, A History of Poker, PAGAT (Dec. 23, 2010), http://www.pagat.com/poker/history.html.
way of riverboats travelling the Mississippi River.\textsuperscript{20} Poker flourished in the western frontier states, especially in Texas.\textsuperscript{21} By the end of the nineteenth century, poker’s popularity had increased to the point where one reporter questioned whether it had supplanted baseball as “the national game.”\textsuperscript{22} By that time it was no longer “confined to the rough South-west”\textsuperscript{23} and was even played among senators.\textsuperscript{24} Poker continues to grow in popularity today, attracting more players with the accessibility of online poker. In addition, poker has become a spectator event that reaches wider audiences than ever with the advent of the televised broadcasting of major poker tournaments.\textsuperscript{25}

Despite the consistent popularity of the game, commercialized poker today remains illegal in the majority of states.\textsuperscript{26} Of the states that do not ban poker in all forms, the law
varies widely. Some states, such as Montana and Florida, allow poker but keep it under stringent restrictions. In California, poker is legal under state law, leaving local governments to determine whether to license cardrooms. In the more well-known legalized gambling cities, such as Las Vegas and Atlantic City, the states have taken different approaches. Poker is legal throughout the state of Nevada, while New Jersey has confined the game's legality to Atlantic City. Six states, mostly those on the Mississippi River, have legalized riverboat casinos. Many states, for example Kentucky and New Hampshire, allow poker games to run for the benefit of charities and nonprofit organizations, subject to certain regulations. Despite these exceptions, the game of poker is prohibited by most states. There are arguments, though, that under the various common law tests of those states, poker could be legal.

27 See Montana Card Games Act of 1974, MONT. CODE ANN. §§ 23-5-306 to 23-5-332. (West 2013). Poker is an authorized card game, id. at § 311, but is subject to restrictions contained in the rest of the Act, such as the permits needed to conduct games id. at § 306, licenses required for poker dealers, id. at § 308, a restriction on the hours of play, id. at § 307, a limit of $300 on prize money, id. at § 312, and special rules covering tournaments, id. at § 317.

28 See FLA. STAT. ANN. § 849.085 (West 2013), which allows penny-ante games of poker held in a person’s dwelling in which the winnings do not exceed $10 per hand. Section 849.086 allows the licensing of cardrooms subject to restrictions as set by that statute and by The Division of Pari-mutuel Wagering of the Department of Business and Professional Regulation. FLA. STAT. ANN. § 849.06.

29 Rose, supra note 26. California has an interesting history with poker; “stud-horse poker” was made illegal by statute in 1885, but that language was repealed from the statute in 1981, in part because no one knew what “stud-horse poker” was. Bennett M. Liebman, Poker Flops Under New York Law, 17 FORDHAM INTELL. PROP. MEDIA & ENT. L.J. 1, 22-28 (2006–2007). Those games that are not specifically prohibited by state statute, as no form of poker currently is, can be prohibited or regulated by localities without being in conflict with or preempted by state law. Id.

30 MASON & NELSON, supra note 26, at 32-33. Nevada is the only state to have legalized commercial casinos throughout the entire state. Id. Other states that have followed New Jersey’s lead in limiting legalized commercial casinos to specified geographic areas include Michigan in Detroit, Colorado in three cities, and South Dakota in Deadwood. Id. at 39. Deadwood is a tourist attraction for those wishing to see where Wild Bill Hickok was allegedly shot while holding what came to be known as the “dead man’s hand,” two pair, aces, and eights. See WILSON, supra note 20, at 3-23 (2008).

31 These six states are Iowa, Illinois, Indiana, Mississippi, Louisiana, and Missouri. Indiana has water-based casinos on the Ohio River and Lake Michigan, while the other five states have riverboat casinos on the Mississippi River. MASON & NELSON, supra note 26, at 36-9; Ronald J. Rychlak, The Introduction of Casino Gambling: Public Policy and the Law, 64 MISS. L. J. 291, 304 n.77 (1994–1995). For an in-depth discussion of Mississippi riverboat casinos, see Rychlak at 305-11.

32 Rose, supra note 26.

33 The different common law tests focus on the main issue of whether a game is one of skill or chance. Many argue that poker is a game of skill and therefore under some of the common law tests, is legal. Both the different types of common law tests, and the issue of whether poker is a game of skill or of chance, will be discussed later in this note.
A.  The Legal Significance of Skill in State Law

There is great statutory and common law variance among the states regarding the regulation of poker and other forms of gambling. States have differing regulatory schemes for gambling, and they employ different common law tests to determine whether games are a legal one of skill or an illegal game of chance. Because the state law is so varied, it makes sense that in the IGBA, Congress intended to defer to state law definitions of gambling, and not supplement them with a new, federal definition of gambling.

Whether games such as poker are categorized as a game of skill or one of chance has legal significance in determining whether the game is legal under state law. The Supreme Court has defined a lottery as having three elements: consideration, distribution of a prize, and distribution of that prize according to chance.\textsuperscript{34} Although the Supreme Court was addressing the elements of a “lottery, gift enterprise, or similar scheme,”\textsuperscript{35} these three elements also constitute the common law definition of gambling in general.\textsuperscript{36} The consideration element requires that something of value, such as money or property, be given up for a chance to win a prize.\textsuperscript{37} The expenditure of time and effort can fulfill the consideration requirement, but the amount of time and effort that must be expended varies among jurisdictions.\textsuperscript{38} The prize element requires the chance that participants win something of value.\textsuperscript{39} The prize does not have to be monetary and it does not have to have a value greater than that of the consideration given.\textsuperscript{40} The last element, chance, is defined differently in different jurisdictions.

The distinction between games of skill and games of chance was recognized long ago. Historically, governments have justified permitting games of skill to encourage their

\textsuperscript{35} Id. (internal quotation marks omitted).
\textsuperscript{36} I. NELSON ROSE, GAMBLING AND THE LAW 77 (1986); J. Royce Fichtner, Carnival Games: Walking the Line Between Illegal Gambling and Amusement, 60 DRAKE L. REV. 41, 45 (2011–2012) (“While there was once a practical difference between the use of the terms ‘gambling’ and ‘lotteries,’ any line of demarcation between the two terms has disappeared. . . . If all three elements—consideration, prize, and chance—are present, the activity constitutes gambling.”).
\textsuperscript{37} Fichtner, supra note 36, at 46.
\textsuperscript{38} Id. at 46-47; ROSE, supra note 36, at 77-79. For example, the Supreme Court in Fed. Commc’n’s Comm’n v. Am. Broad. Co. held that the consideration of expending effort by listening to the radio station in order for a chance to win the prize was insufficient to fulfill the element of consideration. 347 U.S. at 294.
\textsuperscript{39} Fichtner, supra note 36, at 47.
\textsuperscript{40} Id.
citizens to develop and perfect skills that were considered of merit.\textsuperscript{41} There are games that are seen as pure skill, such as chess (even though there is an element of chance in deciding who moves first), and those on the other end of the spectrum that are considered games of pure chance, such as bingo.\textsuperscript{42} There are a large number of games that fall in the middle of the spectrum somewhere between skill and chance, and it is in this gray area where courts employ different legal tests to determine whether the game is prohibited by law.

When there is a game of mixed skill and chance, one test a state court might employ is what has come to be known as the “pure chance” test.\textsuperscript{43} The pure chance test considers whether the participants in the game can influence the outcome of that game by their skill.\textsuperscript{44} If they can, then the game is one of skill and is legal.\textsuperscript{45} Under this approach, it is immaterial to what extent skill has influenced the outcome, so long as it played a part in determining it.\textsuperscript{46} Modern American courts have largely rejected the pure chance test.\textsuperscript{47} A few states have even adopted a rule, called the “any chance” test, which is the opposite of the pure chance test. Under the any chance test, if the outcome is determined in any measure by chance, it is considered gambling.\textsuperscript{48}

The majority of American courts determine whether a game is one of skill or chance using the predominance test, alternatively known as the dominant factor test.\textsuperscript{49} This test considers whether skill or chance predominates in the game in question.\textsuperscript{50} The test focuses on the influence of skill on the outcome of the game, not whether it is “just one part of the

\textsuperscript{41} Id. at 48-49 (quoting Corp. Org. & Audit Co. v. Hodges, 47 App. D.C. 460, 466 (D.C. Cir. 1918)); Anthony N. Cabot, Glenn J. Light & Karl F. Rutledge, Alex Rodriguez, A Monkey, and the Game of Scrabble: The Hazard of Using Illogic to Define the Legality of Games of Mixed Skill and Chance, 57 Drake L. Rev. 383, 389 (2008–2009). For example, ancient Roman societies encouraged martial games among young men to practice the skills they would use as future soldiers, and Islamic law’s prohibition on gambling did not extend to wagers on horse racing because training horses was useful for war.

\textsuperscript{42} Fichtner, supra note 36, at 49.

\textsuperscript{43} Id. (citing Secretary of State v. St. Augustine Church, 766 S.W.2d 49 (Tenn. 1989)).

\textsuperscript{44} Liebman, supra note 29, at 8.

\textsuperscript{45} Id.

\textsuperscript{46} Fichtner, supra note 36, at 49.

\textsuperscript{47} Id. at 50.


\textsuperscript{49} Cabot, Light & Rutledge, supra note 41, at 390.

\textsuperscript{50} Id. at 390-92.
larger scheme.” Some of the factors courts may consider as evidence that the game is one of skill include whether: a skillful player would, in the long-term, win more often than an unskillful player; skill can be learned from experience and/or from acquiring more knowledge by reading; knowledge of mathematics is useful in the game; knowledge of psychology in games played against other competitors can help influence the actions of others; player participation influences the outcome of the game. Some courts also consider the pool of participants, analyzing whether the people who are likely to participate in the game actually possess the skill to influence the outcome of the game, rather than whether a small class of experts would be able to do so.

In applying the predominance test, courts have differed over whether poker is a game of skill or chance. Courts in Illinois, Nebraska, North Carolina, Massachusetts, and Utah have held that poker is a game of chance. In contrast,

52 ROSE, supra note 36, at 80-81.
53 See Fichtner, supra note 36, at 51-53, & n.73 (listing cases).
54 Cabot & Hannum, supra note 19, at 459. In some jurisdictions, the question of whether a game is one of skill or one of chance is treated as a question of fact for the fact finder, while in other jurisdictions it is a question of law for the judge. Compare People v. Mitchell, 444 N.E.2d 1153, 1155 (Ill. App. Ct. 1983) (upholding jury’s finding that games of poker are not “bona fide contests for the determination of skill”), with Bell Gardens Bicycle Club v. Dep’t of Justice, Cal. Rptr. 2d 730, 750-51 (Cal. Ct. App. 1995) (treating jackpot poker as a game of chance).
55 Mitchell, 444 N.E.2d at 1153, 1155 (upholding jury’s finding that games of poker are not “bona fide contests for the determination of skill”).
56 Indoor Recreation Enters., Inc. v. Douglas, 235 N.W.2d 398, 400-02 (Neb. 1975) (holding, using a predominance test, that both poker and bridge are games of chance). After Indoor Recreation, the Nebraska legislature changed the language of the applicable statute. Am. Amusements Co. v. Neb. Dept. of Revenue, 807 N.W.2d 492, 500-01 (Neb. 2011). However, the Supreme Court of Nebraska has held that the new language simply “rewords the predominance standard,” rather than change the test. Id. at 500-02. Because the test has not been changed, it seems that Indoor Recreation remains good law.
57 State v. McHone, 90 S.E.2d 539, 540 (N.C. 1955) (upholding jury conviction of defendant who allowed a game of chance, poker, to be played and wagered on his property).
58 Chapin v. Haley, 133 Mass. 127 (1882) (upholding jury verdict convicting defendant, when the jury was charged that one of the elements required for a guilty verdict was a finding that draw poker was a game of chance). But see Commonwealth v. Club Caravan, Inc., 571 N.E.2d 405, 406 (Mass. App. 1991) (holding that, as a matter of law, video poker machines involved an element of skill, although not as much skill as live poker, that made the machines legal under the statute).
59 Collet v. Beutler, 76 P. 707 (Utah 1904) (refusing to recognize a debt because it was a gambling debt borrowed to play poker, a game of chance. The court wrote that the trial court should have directed the criminal law be enforced against these parties.).
60 Cabot & Hannum, supra note 19, at 462; Liebman, supra note 29, at 19-21.
courts in California, Oregon and Washington have held that poker is a game of skill. Some states retain a distinction between games prohibited as a lottery (a game of chance) and those prohibited as gambling; therefore, some games could be considered not a lottery (and therefore a legal game of skill), but still prohibited or otherwise regulated as gambling.

A minority of states employ a test similar to the predominance test, though more stringent, known as the material element test. Under this test, if chance is a material element (meaning it is more than incidental) in determining the outcome of the game, that game is considered gambling, even if skill predominates. Eight states—Alabama, Alaska, Hawaii, Missouri, New Jersey, New York, Oklahoma, and Oregon—adopt the material element test. Applying this test, New York courts have held that poker is a game of chance.

A few states have their own unique approaches to determining whether a game is one of skill or of chance. Ohio statutorily defines poker, as well as craps and roulette (but not bingo), as a game of chance. Texas prohibits playing for money in “any game played with cards” regardless of whether the game is one of skill or chance. In Texas there are a few available defenses to a charged violation, including one that requires the defendant to prove that the “gambling [was] in a private place,” that no one “received any economic benefit other than personal winnings” and that “except for the advantage of

---

61 Bell Gardens Bicycle v. Dep’t of Justice, 42 Cal. Rptr. 2d 730, 750-51 (Cal. Ct. App. 1995) (1995) (holding that jackpot poker is a game of chance where skill does not play the same role as it does in regular draw poker). Recall that California has a unique situation, in which certain enumerated games are prohibited by state statute (of which stud-horse used to be, but is no longer one), and other gambling games (including poker) may be prohibited or regulated as a locality so chooses without being in conflict with the state law. See Liebman, supra note 29, at 23-27.

62 State v. Coats, 74 P.2d 1102, 1106 (Or. 1938) (stating, albeit in dicta, that poker is not a lottery because it is a game of skill, even though it is prohibited as a gambling game).

63 State ex rel. Schillberg v. Barnett, 488 P.2d 255, 257-58 (Wash. 1971) (en banc) (holding that poker is not a game of chance for the purposes of a prohibited lottery, but that poker is a gambling game prohibited in the state’s criminal law). Though the definitions of lottery and gambling in many jurisdictions are the same, some jurisdictions have retained the distinction.

64 See, e.g., Coats, 74 P.2d at 1106; Barnett, 488 P.2d at 257-58.

65 Cabot, Light & Rutledge, supra note 41, at 392-93.

66 Id.

67 Id. at 392, n.64.


69 OHIO REV. CODE ANN. § 2915.01(D) (West 2013).

70 TEX. PENAL CODE ANN. § 47.02(a)(3) (West 2013).
skill or luck, the risks of losing and the chances of winning were the same for all participants.”

For the purposes of this note, the significance of the different state law tests and their applications is not so much in their substance, but rather in their variety. Given this variety, it makes sense that the IGBA defined gambling solely by reference to state law and did not create an additional, federal element to supplement these tests in § 1955(b)(2). It is also significant that even states who employ the same test have reached different results, even further counseling Congress’ deference to state law. This stance is bolstered by the fact that, while there are many possible definitions of chance in gambling, Congress did not debate any of these definitions when discussing the IGBA.

B. The History and Rules of Texas Hold ‘Em

There are many different types of poker, but none are more popular today than Texas Hold ‘Em. That popularity is due in large part to the fact that No Limit Texas Hold ‘Em is the main event in the World Series of Poker. The World Series of Poker started in 1970 in Benny Binion’s Horseshoe Casino as a chance for some of the biggest names in poker to play each other in a high-stakes game. Since then, the event has grown into a spectacle with thousands of competitors, 61 different tournaments, and a $1 million dollar buy-in for the 2012 main event. In the 21st century, the advent of online poker further increased the accessibility and popularity of Texas Hold ‘Em to amateurs, and poker tours featuring Texas Hold ‘Em continue to attract players and viewers alike. The increasing popularity of poker has led organizations, such as the Poker

---

71 Id § 47.02(b).
72 See infra Parts III.B–C.
73 Liebman, supra note 29, at 5 (citing John Scarne, SCARNE’S GUIDE TO MODERN POKER 14 (1980)).
74 Wilson, supra note 20, at 175-76. This game grew out of one held the year before, called the “Texas Gamblers Reunion,” which included the same cast of characters. Id. Tom Moore had suggested to Binion that they make it an annual game to crown a world champion. Id. The first champion, Johnny Moss, was elected by a vote of the players in 1970. Id.
Players Alliance, to advocate for the legalization of poker. The Poker Players Alliance is a nonprofit organization that lobbies Congress and state legislatures for laws favorable to poker and files amicus briefs in cases regarding poker.\textsuperscript{77}

Texas Hold 'Em is a non-house banked game, meaning that the dealer does not play and the house does not have a stake in the outcome.\textsuperscript{78} Rather, players are competing only for each other’s money.\textsuperscript{79} In Texas Hold 'Em, each individual player is dealt two pocket, or hole, cards. These cards will be combined with the five community cards, which are shared and available for all the players to use, in order for the players to create the best five-card hand possible.

The community cards are revealed in a staggered manner, which offers players multiple opportunities to make decisions on how to bet. Before pocket cards are dealt, the player to the left of the dealer button must bet the small blind, and second to the left must bet the big blind.\textsuperscript{80} After the hole cards are dealt, but before any community cards are dealt, there will be the first round of betting (pre-flop); betting occurs sequentially in a clockwise-fashion. At this time, a player has the option on his or her turn to either call or raise the bet, or to fold the hand.\textsuperscript{81} When each player has either called or folded, the dealer deals three cards face up (unless, of course, only one player remains), known as the flop. This is followed by a second round of betting, and then the fourth card, known as the turn. And then there is a third round of betting, followed by the fifth card, known as the river.

Once all the community cards have been revealed, there is one final round of betting. If two or more players remain in the game, there is a showdown and the player with the highest


\textsuperscript{78} The house usually makes its money either by a rake from the pot, typically 5% to 10%, though some may charge an hourly fee or a flat rate per hand. See Dicristina, 886 F. Supp. 2d at 173 (quoting Cabot & Hannam, supra note 19 at 452-53).

\textsuperscript{79} Id.

\textsuperscript{80} The dealer button will rotate clockwise after each hand, thereby rotating who has the mandatory small and big blinds, as well as the order of betting in the betting rounds. The big blind is usually double the amount of the small, and is the minimum bet allowed in the game. Id. at 172; WILSON, supra note 20, at 326.

\textsuperscript{81} Another way you can play Texas Hold ‘Em is to have an ante from all players in lieu of blinds. In games played under these rules there is the additional option of checking.
hand wins. Who wins the showdown is determined by the hierarchy of hands. In the event of a tie, the winner is determined by the kickers, the highest cards in the player’s five-card hand other than the cards already used to make the significant part of the hand (the pair, for instance). If there is still a tie after considering the kickers (for instance, when the players have the same kickers because the highest cards are the ones on the board), the players will split the pot. Many hands of poker never make it to the showdown because all but one player has already folded. In these cases the winner is not required to show his or her hand.

C. The Skill Employed in Texas Hold ’Em

Texas Hold ’Em can be used as an example to illustrate why states can vary widely in their determinations of whether a specific game is legal, even if they adopt the same legal standard regarding the element of chance. Even using the same legal standard for chance, there are valid arguments supporting both that poker is a game of skill and one of chance. Some argue that it is a game of chance because which cards are dealt is a chance occurrence. Others argue that poker is a game of skill. During a poker game, players utilize several different skill sets and make numerous strategic choices: bluffing, opponent modeling, unpredictability, betting strategies, risk and money management, psychology, and calculating probabilities of certain cards being dealt or certain hands being made. Because of the numerous choices being made, and because poker is a game of incomplete information, it has been used in research for computer science and artificial intelligence.

One study shows that 75.7% of the sampled 103 million hands ended before showdown. The same study also found

---

82 For a more detailed discussion of hand rankings in the context of Texas Hold 'Em, complete with illustrations, see SAM BRAIDS, THE INTELLIGENT GUIDE TO TEXAS HOLD 'EM 6-10 (2d ed. 2010).

83 Of course those who fold do not show their hands either. For the basic rules of the game, see Dicristina, 886 F. Supp. 2d at 172-73; WILSON, supra note 20, at 326-27; or one of the many books on how to play Texas Hold ’Em, such as BRAIDS, supra note 82, at 5-10.

84 This list is not intended to be all-inclusive. See Dicristina, 886 F. Supp. 2d at 173-76; Cabot & Hannum, supra note 19, at 467-83; Noga Alon, Poker, Chance, and Skill, 5-13 (unpublished manuscript), available at http://www.tau.ac.il/~nogaa/PDFS/skill4.pdf.


86 PACO HOPE & SEAN MCCULLOCH, STATISTICAL ANALYSIS OF TEXAS HOLD ’EM 5 (2009).
that 50.3% of the hands that did end in a showdown were not won by the player with the best hand (because the stronger hand had folded before the showdown). The implication of these statistics is that the games’ final results are not determined by the cards (the element of chance), but rather the decisions the players made (the skill).

Poker is a game of incomplete information because players do not know which cards their opponents hold or which cards will be dealt in the flop, turn, and river. However, because it is played with a deck of 52 cards, there are a limited number of possibilities. Therefore, the probabilities of making each type of hand can be calculated. The probabilities of making these hands change each time more cards are dealt. To use an example from Anthony Cabot, Esq. and Professor Robert Hannum, a player has a 1 in 509 chance of making a flush before any cards are dealt. If the player is dealt two spades for their hole cards, and there are two spades and a diamond on the flop, the odds against making a flush on the turn card is approximately 4.2 to 1. If the player does not make the flush after the turn card is dealt, the odds against him or her making it on the river is 4.1 to 1. The probability that you will fail to make the flush on either the turn or the river is 1.86 to 1. This kind of analysis can be utilized by skillful players to help them determine the probability of forming a certain hand for both themselves and for their opponents (who they might guess based on the community cards and their betting strategy was trying to make a certain hand), guiding the other decisions the player makes during the course of the game.

The probability of poker hands is only one of the factors that can influence the outcome of a hand. Using opponent modeling, players can categorize the type of play of their opponents (for example: their level of skill, whether they are a passive or aggressive bettor, whether they play a lot or relatively few hands) and then use that categorization to exploit the opponents’ weaknesses. Players can defend themselves against opponent modeling by being unpredictable, changing their style of

87 Id.
88 Cabot, Light, & Rutledge, supra note 41, at 396-98; Peterson, supra note 85, at 40.
89 For a chart listing the probability of making each type of hand, see Cabot & Hannum, supra note 19, at 472.
90 Id.
91 Id. at 474.
92 Peterson, supra note 85, at 40.
Another skill players may utilize is betting strategy. Players can use wagers to try to influence another player’s decision to call, raise, or fold a hand. For instance, players can make it uneconomical for their opponents to call or raise, or can give them information (possibly false) about the strength of their hand. An important factor in betting strategy is a player’s position at the table, which determines the order in which the bets are placed. A player betting later will have more information about the other players’ hands based on their opponents’ decisions, which could be combined with other information, such as knowledge of the opponent’s playing type and betting habits, to better predict the opponents’ hands.

One betting strategy commonly employed by skillful players is bluffing. The basic concept of bluffing is that a player bets aggressively in order to give opponents the impression that he or she has a strong hand, when in fact the player does not, in the hope that this will entice opponents to fold. However, this form of bluffing can be too predictable when playing skilled opponents, so experienced poker players may employ a form of bluffing known as the post oak bluff. In this type of bluff a player bets a small amount, giving opponents the impression that the player is trying to make them call because the player has a strong hand. The opponents will think the player is trying to make them call, and rather than falling into that trap, the opponents will fold, when in actuality the player did not have a strong hand. This strategy would be most effective in a situation where the opponent has modeled the player as one who makes small bets when they have strong hands. This is an example of how the different skills a player utilizes interact with one another. The various skills used in poker support the argument that poker is a game of skill, not a game of chance.

---

93 See id.; Alon, supra note 84, at 9-10.
95 Note that the order of betting on any given hand can be considered an element of chance, although the order of betting does rotate with each hand.
96 However, there are players who prefer to play up front (meaning they prefer to bet earlier rather than later in the betting order). See Dicristina, 886 F. Supp. 2d at 175; Alon supra note 84, at 11-13.
97 Dicristina, 886 F. Supp. 2d at 174-75.
98 Id.
Whether poker is a game of skill or chance, and how that is determined, varies depending on the applicable state law. State law varies greatly in how it defines the element of chance in a gambling game. Because state law is so nuanced, it would make sense that Congress deferred to state law in § 1955(b)(1) when it defined an “illegal gambling business” as one which operated in violation of state law. Further, it does not seem likely that Congress adopted one of the various definitions of gambling in § 1955(b)(2) without any discussion of which definition it was adopting. It also is in keeping with the history of federal regulation of gambling to defer to the states for the substantive definition of gambling rather than adopt a federal definition. With respect to the IGBA, this reading fits with Congress’s intent because the motivation behind the legislation was to aid the states where they have failed to enforce their own anti-gambling law, not for the federal government to take charge in the efforts to define illegal gambling.99

II. THE FEDERAL GOVERNMENT’S ROLE IN REGULATING GAMBLING

Traditionally, gambling has been an area largely left to the states to regulate.100 The federal government has only gotten involved in the regulation of gambling when it had reason to believe the states alone could not adequately enforce their own gambling laws.101 Early federal gambling law in the nineteenth century focused on the governance of lotteries, and not much else.102 These anti-lottery laws were often early battlegrounds for shaping the concept of federalism during a time when expansion of the federal government’s powers was much

99 See infra Part III.
100 See Blakey & Kurland, supra note 1, at 925.
101 Id. at 926.

The earliest congressional concern arose in response to the inability of states acting alone to control the perceived abuses of the nineteenth century state-chartered lotteries. Subsequent federal gambling legislation has manifested a variety of policies that include depriving organized crime of its gambling revenue, harmonizing federal gambling taxes with diverse state gambling policies, and developing a coherent gambling policy to govern federal enclaves.

102 See id. at 927-43 (discussing federal regulation of lotteries in the 19th century and at the turn of the century); I. Nelson Rose, Gambling and the Law®: The Third Wave of Legal Gambling, 17 VILL. SPORTS & ENT. L.J. 361, 370, 374 (2010) [hereinafter Rose, Third Wave] (discussing the first federal anti-gambling laws, describing them as weak due to the view of the federal government as not having much power, and discussing the use of the commerce clause in the late 19th century).
contested.\footnote{103} In the early twentieth century, Congress “virtually abstain[ed]” from regulating gambling, though it did prohibit broadcasting lottery advertisements in the Federal Communications Act of 1934.\footnote{104} Congress also began to regulate commodities futures, which was then considered a form of gambling.\footnote{105} After this hiatus, the federal government once again began a period of legislative activity starting in 1950 with the Kefauver Committee.\footnote{106} With this new period of activity the federal government had a new reason for involving itself with gambling—the connection between gambling and organized crime.\footnote{107} During this time, the federal government was mindful that the regulation of gambling was still mainly within the purview of the states, and sought to aid the states in curbing illegal gambling where organized crime created unique problems for state and local governments. The federal government did not seek to substantively define gambling. Instead, its primary aim was to help the states enforce their own gambling laws. Because the history of the federal regulation of gambling is a story about federalism, it is through this lens that federal legislation concerning gambling, including the IGBA, should be viewed.

A. Federalism and the Lottery

The first period of federal regulation concerning gambling focused primarily on lotteries. Lotteries were widespread in the colonial era and post-Revolutionary war, as they were an easy way for both governments and private individuals to raise capital.\footnote{108} Beginning in the 1820s and

\begin{itemize}
\item \footnote{104} Blakey & Kurland, supra note 1, at 946, 958 n.138.
\item \footnote{105} Id. at 958 n.138. Congress passed the Future Trading Act in 1921, which was later declared unconstitutional. Congress then passed the Commodity Exchange Act in 1936. Id.
\item \footnote{106} Nat’l Inst. of Law Enforcement & Crim. Justice, supra note 103, at 562-63. The Kefauver Committee, led by Senator Kefauver, was formally known as the Senate Special Committee to Investigate Crime in Interstate Commerce. Id. at 562 n.60.
\item \footnote{107} See id. at 562-64.
\item \footnote{108} Rose, Third Wave, supra note 102, at 368-70 (describing this period as the ‘first wave’ of legalized gambling); Dinan, supra note 103, at 649.
\end{itemize}
1830s, public sentiment had turned against lotteries due to the exposure of corruption and scandal in lotteries.\textsuperscript{109} The reaction was a wave of state legislation and state constitutional amendments banning lotteries; by 1862, all but two states, Missouri and Kentucky, had banned them.\textsuperscript{110} But after the Civil War, some Southern states legalized lotteries to raise funds to rebuild.\textsuperscript{111} Even though many states had banned lotteries within their borders, they could not keep other states from selling their residents lottery tickets through the mail, since regulation of the mail was under federal jurisdiction.\textsuperscript{112} To solve this problem, Congress passed legislation in 1868 that prohibited mailing lottery offers through the U.S. Post Office.\textsuperscript{113} The law was amended in 1876, and among the alterations was the deletion of the word “illegal,” which had the effect of prohibiting the mailing of all lottery-related papers, including those legal under state laws.\textsuperscript{114} This provision was controversial when debated by the Senate, as Senators argued that the legality of lotteries was the prerogative of the states.\textsuperscript{115} The law was upheld as constitutional in \textit{Ex parte Jackson}, but the Supreme Court avoided the federalism question by focusing on the power of the federal government to regulate the U.S. Post Office rather than the intrusion on states’ power, ignoring the 10th Amendment argument before the Court.\textsuperscript{116}

\textsuperscript{109} Rose, \textit{Third Wave}, \textit{supra} note 102, at 369-70; Dinan, \textit{supra} note 103, at 649.
\textsuperscript{110} Rose, \textit{Third Wave}, \textit{supra} note 102, at 369-70; For a discussion of the Supreme Court’s involvement and interpretation of the Contract Clause regarding these state law prohibitions, see Blakey & Kurland, \textit{supra} note 1, at 929-31.
\textsuperscript{111} James J. Devitt, \textit{Legal History: The Louisiana Lottery}, 55 L.A. B.J. 346 (2008). Some of the states that legalized lotteries in the antebellum period included Louisiana, Kentucky, Alabama, Georgia, and Mississippi. \textit{Id.} It seems that there was another wave of anti-lottery sentiment following that period, because “[b]y the end of the [19th] century, 35 states had prohibitions against lotteries in their constitutions and no state permitted lotteries again until 1963.” \textit{Id.}
\textsuperscript{112} Blakey & Kurland, \textit{supra} note 1, at 931. When other state chartered lotteries sold tickets through the mail to residents of states that had banned lotteries, the anti-lottery states could only enforce their laws by going after the in-state consumers, since they had no jurisdiction over the other states or the mail system. Enforcing their laws by going after the purchasers of the lottery tickets proved “difficult, expensive, and unpopular.” \textit{Id.}
\textsuperscript{113} NAT’L INST. OF LAW ENFORCEMENT & CRIM. JUSTICE, \textit{supra} note 103, at 501-03 (citing Act of July 27, 1868, 15 Stat. 194, codified as amended at 18 U.S.C. § 1302 (2012)). This law was difficult to enforce in practice because the post office could not open letters due to Fourth Amendment protections, and employees were not allowed to delay mail. See \textit{Id.} at 503.
\textsuperscript{114} Blakey & Kurland, \textit{supra} note 1, at 933.
\textsuperscript{115} NAT’L INST. OF LAW ENFORCEMENT & CRIM. JUSTICE, \textit{supra} note 103, at 505-06 n.11.
\textsuperscript{116} \textit{Id.} at 506-08.
Despite the law being upheld, some state-operated lotteries sold tickets through the postal system in open violation of the federal law, because enforcement was ineffective. Of these, the biggest problem for states that prohibited lotteries was the Louisiana Lottery, dubbed the “Serpent.” The Louisiana legislature allegedly accepted bribes in return for authorizing the lottery and giving it tax exemptions. Over 90% of the lottery’s revenue came from interstate sales. Other states were powerless to stop the sale of these tickets in their own states, and pressure mounted on Congress to take action.

President Harrison addressed Congress about the issue and urged federal legislation to help the states who could not control the effect another state’s action had on their own constituents. In 1890, Congress responded and amended the earlier law by broadening the definition of prohibited material, extending the prohibition to newspapers, and authorizing the post office to detain letters they suspected solicited lottery ticket sales. This 1890 Act was the product of “fifteen years of congressional debate” about the role of the federal government and the protection of states’ rights.

Opponents of the legislation argued that it was the states’ prerogative to legislate on moral issues, and that the federal government should not be able to undermine that legislation by criminalizing activities that were legal under state law. Proponents countered that states that prohibited lotteries were unable to protect themselves from the Louisiana Lottery. Those states needed federal legislation because they did not have the power to regulate the postal system or interstate commerce, and had done what they could within their jurisdiction to no avail.

Ultimately, because the states were unable to control the ill-effects of the Louisiana Lottery on their states without the

---

117 Blakey & Kurland, supra note 1, at 935-36.
118 Devitt, supra note 111, at 346.
119 Nat’l Inst. of Law Enforcement & Crim. Justice, supra note 103, at 512 n.24 (“The ease with which these measures were passed seem [sic] to justify the repeated claim that the company controlled every Louisiana legislature from 1868 to 1982”); Rose, Third Wave, supra note 102, at 371.
120 Devitt, supra note 111, at 346.
125 Id. at 518 n.39. Opponents were also concerned that the extension of the prohibition to newspapers threatened the freedom of the press. Id. at 513 n.29.
126 Id. at 518 n.38.
federal government’s assistance, Congress passed the legislation.\textsuperscript{127}

The 1890 amendments hurt the business of the Louisiana Lottery, and in 1892 the Louisiana legislature prohibited the further sale of lottery tickets.\textsuperscript{128} It seemed that the federal role in regulating the lotteries should have been finished. But the Louisiana Lottery moved its operation to Honduras, where it was able to reach its U.S. customers through Florida, without using the postal system and violating federal law.\textsuperscript{129} To reach the Louisiana Lottery’s operations, Congress in 1895 relied on its commerce clause power to pass 18 U.S.C. § 1301. The new law prohibited use of interstate or foreign commerce to bring lottery-related instruments into the United States.\textsuperscript{130} This was a novel and controversial use of the commerce clause power, which the Supreme Court upheld in Champion v. Ames in 1903.\textsuperscript{131}

In the second half of the twentieth century, states began to legalize lotteries.\textsuperscript{132} The federal laws that helped states keep the Louisiana Lottery out of their borders were now hindering the growing number of states that had legalized lotteries.\textsuperscript{133} The U.S. Department of Justice was ready to prosecute those states for violation of federal anti-lottery laws, but Congress passed legislation that exempted state-run lotteries from federal law.\textsuperscript{134} This exception was narrowly tailored to exempt only state-conducted lotteries authorized by state law when they were acting within their state.\textsuperscript{135} The exemption excluded those states that had authorized lotteries, while continuing to enforce the law with respect to those states that remained anti-lottery.\textsuperscript{136}

\textsuperscript{127} Id.
\textsuperscript{128} Rose, Third Wave, supra note 102, at 373. This law did not take effect until December 31, 1893. See id.
\textsuperscript{129} Id. at 373-74.
\textsuperscript{131} Rose, Third Wave, supra note 102, at 374 (citing Champion v. Ames, 188 U.S. 321 (1903)).
\textsuperscript{132} Blakey & Kurland, supra note 1, at 950 ("By 1978, fourteen states had authorized lotteries."). The trend continued: in the 1980s seventeen states and the District of Columbia legalized lotteries, and six more states did so in the 1990s. Mason & Nelson, supra note 19, at 9.
\textsuperscript{133} DEVELOPMENT OF THE LAW OF GAMBLING, supra note 90, at 540.
\textsuperscript{135} For a discussion of the specific exemptions, and the inconsistencies and problems the exemptions created, see Nat’l Inst. of Law Enforcement & Crim. Justice, supra note 103, at 545-58.
\textsuperscript{136} Blakey & Kurland, supra note 1, at 953-54.
The history of federal regulation of lotteries shows Congress’s recognition that lotteries are normally an area of the law left to the states. However, when the states were unable to enforce their own gambling policies due to action outside the state, the federal government got involved. In the case of lotteries, the federal government legislated to protect the states from the Louisiana Lottery interrupting the integrity of their anti-lottery stance during a time when the states generally disfavored lotteries.\footnote{Of course, this was at the expense of states such as Louisiana which favored lotteries.} When public opinion had turned to favoring lotteries in the twentieth century, Congress acted to exempt state-run lotteries in recognition of that sentiment.

\section*{B. Organized Crime and Federal Regulation of Gambling in the Twentieth Century}

The federal government did not involve itself much in the regulation of gambling in the first half of the twentieth century.\footnote{Blakey & Kurland, supra note 1, at 958.} Beginning in the 1950s, and lasting into the 1970s, the federal government began a foray into governing gambling. Congress legislated during this time because of a growing concern about organized crime as a nationwide problem.\footnote{\textsc{NAT'L INST. OF LAW ENFORCEMENT \\& CRIM. JUSTICE}, supra note 103, at 560.} It was concerned with organized crime’s involvement in gambling, especially because the profits from gambling were the “principal support of big time racketeering and gangsterism.”\footnote{\textsc{ORGANIZED CRIME AND LAW ENFORCEMENT: THE REPORTS, RESEARCH STUDIES AND MODEL STATUTES AND COMMENTARIES PREPARED FOR THE AMERICAN BAR ASSOCIATION COMMISSION ON ORGANIZED CRIME Vol. I}, at 13 (Morris Ploscowe, ed., Grosby Press 1952) [hereinafter \textsc{ORGANIZED CRIME AND LAW ENFORCEMENT}].} The federal government stepped in to govern gambling when they felt there was a national problem that for various reasons the states alone could not handle. Twentieth century federal regulation regarding gambling was largely enacted during three main time periods: during the Kefauver Committee from 1950-51, under Attorney General Robert F. Kennedy from 1961-62, and under the Nixon Administration from 1969-70.\footnote{Blakey & Kurland, supra note 1, at 959. The efforts under the Nixon Administration, namely the IGBA, will be discussed infra Part III.}
1. The Kefauver Committee

Before 1950, nationwide efforts to study organized crime consisted solely of the knowledge gained by the National Commission on Law Observance and Enforcement, a Prohibition-era look into bootlegging. In 1949, President Truman began to issue public statements that focused national attention on organized crime, and the FBI was asked to report on crime nationwide, adding to the attention. In February 1950, the Attorney General’s Conference on Organized Crime further illuminated the problem of organized crime.

These events led to the formation of the Special Senate Committee to Investigate Organized Crime in 1950, the first extensive national effort to study the problem of organized crime. It became known as the Kefauver Committee, after Senator Estes Kefauver, who headed the committee. The Kefauver Committee held hearings in 14 cities across the U.S., which were widely publicized and even broadcast on television. The Kefauver Committee brought the problem of organized crime onto the national stage, where it remained for the next few decades. Although the Kefauver Committee only immediately resulted in the Johnson Act being passed, its various proposals and factual findings became the basis for much of the gambling legislation passed in the coming years.

The Kefauver Committee found that gambling was the major source of revenue for organized crime, and that gambling profits supplied the capital for organized crime’s other ventures. The Committee found that illegal gambling was

---

142 ORGANIZED CRIME & LAW ENFORCEMENT, supra note 140, at 10.
143 NAT'L INST. OF LAW ENFORCEMENT AND CRIM. JUSTICE, supra note 103, at 563 n.61.
144 Id. at 562.
145 Id.; ORGANIZED CRIME & LAW ENFORCEMENT, supra note 140, at 1-2, 9-10 (containing the reports authorized by, and the suggestions of, the American Bar Association’s Commission on Organized Crime, which was created to cooperate with the federal efforts of the Kefauver Committee).
146 NAT'L INST. OF LAW ENFORCEMENT & CRIM. JUSTICE, supra note 103, at 562 n.60.
147 NAT'L INST. OF LAW ENFORCEMENT AND CRIM. JUSTICE, supra note 103, at 563 n.61. Those fourteen cities are: Washington D.C., Miami, New Orleans, Kansas City, Cleveland, St. Louis, Detroit, Los Angeles, San Francisco, Las Vegas, Philadelphia, Chicago, and New York.
148 The Johnson Act prohibits the transportation of gambling devices in interstate channels. Id. at 564.
149 Id. at 562; ORGANIZED CRIME AND LAW ENFORCEMENT, supra note 140, at 9-10 (This source summarized the findings of the Kefauver Committee, and since it is a contemporaneous source connected with the Kefauver Committee which summarized
mostly controlled by organized crime, and that illegal gambling thrived in many big and small cities, including the 14 cities where the Committee held hearings. Given these findings, it seemed necessary for the federal government to attack illegal gambling if it wanted to weaken the influence of organized crime. The Committee was mindful, however, that anti-gambling laws traditionally fell under state and local jurisdiction. One of the Committee reports recognized as much when it stated:

While channels of interstate communication and interstate commerce may be used by organized criminal gangs and syndicates, their activities are in large measure violations of local criminal statutes. When criminal gangs and syndicates engage in bookmaking operations, operate gambling casinos or slot machines, engage in policy operation, peddle narcotics, operate house[s] of prostitution, use intimidation or violence to secure monopoly in any area of commercial activity, commit assaults and murder to eliminate competition; they are guilty of violating State laws and it is upon State and local prosecuting agencies, police and courts that the major responsibility for the detection, apprehension, prosecution and punishment of offenders rests.

In the resulting legislation from both the Kefauver Committee and later legislative efforts that built on it, the federal government attacked illegal gambling where it had jurisdiction to do so—where the illegal gambling intersected with interstate commerce. The federal government had to step into an area traditionally reserved to the states because only it had the jurisdiction to do so. Congress also sought to enforce federal laws in cases where the state and local law enforcement failed to do so, for various reasons such as apathy, corruption, and lack of resources.

The Kefauver Committee found that organized crime used the channels of interstate commerce to further its gambling operations. For example, modes of interstate communication such as telephone and wire services were an integral part of illegal bookmaking (the practice of taking

---

the Kefauver Committee’s findings from the Committee’s own reports, this source will be used as the equivalent of the Kefauver Committee’s findings).

150 Paul Bauman & Rufus King, A Critical Analysis of the Gambling Laws, in ORGANIZED CRIME AND LAW ENFORCEMENT, supra note 140, at 73.

151 ORGANIZED CRIME AND LAW ENFORCEMENT, supra note 140, at 11-12.

152 Id. at 22.

153 Id. at 22-23 (quoting page 6 of a Kefauver Committee report).

154 Id. at 15-17.

155 Id. at 14.
wagers on competitions, usually sporting events such as horse or dog races, football, and basketball). In bookmaking, interstate communication was often used to place and receive wagers, as well as obtain information regarding the competitions being wagered upon.

State laws were ineffective in curbing illegal bookmaking because when they prohibited the communication of bets and wagers, bookmaking operations were able to conduct their enterprises using out-of-state communication. For example, California and Florida only had limited success when they prohibited the transfer of information used for illegal bookmaking because the bookmaking operations were able to get their information from out-of-state sources. The Kefauver Committee introduced two different bills to deal with this problem, neither of which passed. One, S. 1564, would have prohibited the transmission through interstate commerce of information on sporting contests gained through means that did not have the consent of the owner. The other, S. 1624, would have criminalized the transmission of wagers over interstate communication.

The Kefauver Committee recognized that organized crime also used interstate commerce for the distribution of gambling devices, most significantly slot machines and punchboards. To this end, in 1951 Congress passed the Johnson Act, which prohibits the transportation of gambling devices over state lines. An objection to this law based on the fear that it would infringe upon states’ rights, led by one of Nevada’s Congressmen, was allayed due to a provision which allowed the states to pass a law exempting themselves from the Johnson Act. This was in keeping with the purpose of the Johnson Act, “to support the policy of those States which outlaw slot machines and similar gambling devices, by

156 Id. at 14-15.
157 Id., at 14-15.
158 Id. at 44-45.
159 Id.
160 Id. at 45-46. The bill states that when gambling entities were prohibited from the race tracks, they found other ways of stealing the information, such as setting up observation posts overlooking the track and relaying the information from there. See id.
161 Id. at 46-47.
162 ORGANIZED CRIME AND LAW ENFORCEMENT, supra note 140, at 15.
prohibiting the use of the channels of interstate or foreign commerce for the shipment of such machines or devices into such States.”166 Like the anti-lottery laws and the Wire Act, Congress was able to achieve this goal while leaving room for the states that wished to legalize the prohibited form of gambling to do so.167

The inability of states to reach gambling enterprises operating in interstate commerce was not the sole reason the states could not handle the problem of organized crime’s involvement in gambling without federal aid. The Kefauver Committee thought of organized crime as a nationwide organization, and therefore a uniquely national problem.168 The large geographical scope of organized crime created problems of overlapping jurisdiction for local and state law enforcement.169 The Committee found that law enforcement agencies with overlapping jurisdictions often passed responsibility to one another in order to evade that responsibility and the lack of centralized coordination allowed that practice to continue without accountability.170 This problem was compounded by the corruption of local and state officials. The Kefauver Committee found that illegal gambling flourished in certain areas because the officials in that area were corrupt.171 Law enforcement officials were often bribed to look the other way.172 The Committee also found that organized crime leaders had enough political clout to put their own members into official positions or get elected officials to work for them.173 The corruption of local and state governments was a major factor in the federal government’s intervention.

2. Attorney General Robert F. Kennedy

The second period of the federal government’s active involvement in legislating gambling was in the early 1960s under Attorney General Robert F. Kennedy. In 1961, Congress passed the Wire Act, with provisions similar to the bill S. 1624, which was proposed but not passed during the Kefauver

166 NAT’L INST. OF LAW ENFORCEMENT & CRIM. JUSTICE, supra note 103, at 564 (quoting H.R. Rep. No. 81-2769, at 2 (1950)).
167 For a discussion of the Wire Act, see infra Part II.B.2.
168 ORGANIZED CRIME AND LAW ENFORCEMENT, supra note 140, at 11.
169 Id. at 16.
170 Id. (quoting page 183 of the Committee’s third interim report).
171 Id. at 15.
172 Id. at 16-17.
173 Id. at 17.
Committee period. Like the Kefauver Committee before them, Robert Kennedy and the Congress that enacted the Wire Act were concerned with the regulation of gambling because of the revelation that gambling was a major source of revenue for organized crime. The Wire Act’s purpose was to weaken that source of illegal revenue by prohibiting the use of a “wire communication facility for the transmission in interstate or foreign commerce” of a wager, information pertaining to a wager, or information that entitles one to receive money as the result of a wager. Also like the Kefauver Committee, Congress recognized that it was within the states’ power to legalize gambling, and included an exemption that read:

Nothing in this section shall be construed to prevent...the transmission of information assisting in the placing of bets or wagers on a sporting event or contest from a State or foreign country where betting on that sporting event or contest is legal into a State or foreign country in which such betting is legal.

This exemption is similar to the exemption made in the anti-lottery laws in that it retains the states’ power to legalize the form of gambling prohibited while keeping in place the federal law that reinforces the laws of those states that have chosen to prohibit that form of gambling. In this way the law fulfills its legislative purpose of aiding the states in enforcing their anti-gambling laws, without interfering with the rights of other states to legalize gambling.

With Robert Kennedy’s urging, Congress passed two more pieces of legislation under the commerce clause power. The Travel Act prohibits traveling in interstate commerce with the intent to further unlawful activity. Unlawful activity is further defined to include gambling enterprises, as well as other activities closely associated with organized crime, such as bootlegging, narcotics trafficking, bribery, and extortion. The other piece of legislation, 18 U.S.C. § 1953, prohibits the interstate transportation of wagering paraphernalia for the

---

175 See Blakey & Kurland, supra note 1, at 966 n.173.
176 18 U.S.C. § 1084(a). This statute is limited to those who are engaged in a gambling business, and does not apply to those making the bets. Blakey & Kurland, supra note 1, at 966.
178 NAT’L INST. OF LAW ENFORCEMENT & CRIM. JUSTICE, supra note 103, at 570-71.
180 Id. § 1952(b).
purposes of bookmaking, wagering pools, or numbers games.\textsuperscript{181} The statute, in keeping with the Wire Act and other federal legislation, exempts the transportation of materials into states where pari-mutuel (typically horse races) or sports betting is legal.\textsuperscript{182} In 1975, Congress amended § 1953 to exempt state-run lotteries, the parallel to the exemption created for the anti-lottery laws.\textsuperscript{183}

In sum, many of the issues that made state enforcement of their own gambling laws inadequate continued to be problematic throughout the 1960s and 1970s. In the nineteenth century the federal government regulated lotteries because they had become a national problem and the states were unable to enforce their anti-lottery laws. Likewise, in the twentieth century the federal government legislated with regard to gambling because of the growing national problem of organized crime and the inability of the states to combat this problem alone. The history of federal involvement in gambling shows that the federal government was not interested in defining illegal gambling itself but rather left that to the states, exempting from federal law that which was legal under state law. Instead, the purpose of federal action was to aid the states when larger issues made state enforcement of state law ineffective.

III. THE LEGISLATIVE HISTORY OF THE ILLEGAL GAMBLING BUSINESS ACT

The IGBA continued on the same path that the federal government had started on with the Kefauver Commission. The federal government, concerned with the growing problem of organized crime, decided to attack its gambling roots.\textsuperscript{184} It was necessary to do so because the states were failing to effectively enforce their own gambling laws.\textsuperscript{185} The motivation behind Congressional action was not to usurp state gambling

\textsuperscript{181} Id. § 1953. Numbers is alternatively known as policy or bolita.

\textsuperscript{182} Id. § 1953(b). Pari-mutuel betting is a “betting pool in which those who bet on competitors finishing in the first three places share the total amount bet,” \textit{Pari-mutuel Definition}, \textsc{Merriam-Webster Dictionary}, http://www.merriam-webster.com/dictionary/pari-mutuel (last visited Jan. 4, 2014).

\textsuperscript{183} Blakey & Kurland, \textit{supra} note 1, at 973. Recall that when states began legalizing state lotteries, Congress enacted § 1307, exempting those lotteries from the anti-lottery laws in §§ 1301-1303.

\textsuperscript{184} See \textsc{Richard Nixon, Special Message to the Congress on a Program to Combat Organized Crime in America, April 23, 1969, in Richard Nixon: 1969: Containing the Public Messages, Speeches, and Statements of the President 315, 320 (1971).}

regulation, but rather to aid the states in gambling regulation by addressing their shortcomings. A textual analysis of the language of the IGBA supports the argument that Congress did not intend to create a federal substantive definition of gambling.

In 1967, the President’s Commission on Law Enforcement and Administration of Justice issued a report called “The Challenge of Crime in a Free Society,” which included a chapter discussing organized crime.186 This influential report cited gambling as the biggest source of revenue for organized crime, deriving money from sources “rang[ing] from lotteries, such as ‘numbers’ or ‘bolita,’ to off-track horse betting, bets on sporting events, large dice games and illegal casinos.”187 It also reported that estimates of the annual intake of illegal gambling ranged from seven to fifty billion dollars, with annual profit as high as six to seven billion dollars.188 The report said that organized crime made money from illegal gambling not only by operating gambling services, but also by receiving payments from those operations which were independent of organized crime, often through intimidation.189 President Nixon, in a message to Congress in 1969, echoed many of these same findings. He reported the annual gross take of organized crime from illegal gambling to be estimated at anywhere from 20 to 50 billion dollars.190 He told Congress that illegal gambling, the “wellspring of organized crime’s financial reservoir,” helped finance the more reprehensible activities of organized crime, such as usury, bribery of police and politicians, narcotics trafficking, and infiltration into legitimate business.191 Because of this, President Nixon took the position that the effort against organized crime should focus on illegal gambling.192 In his view, “[g]ambling income is the life line of organized crime. If we can cut it or constrict it, we will be striking close to its heart.”193

---

187 Id. at 188.
188 Id. at 189.
189 Id. at 188.
190 NIXON, supra note 184, at 316.
191 Id. at 320.
192 Id.
193 Id.
A. The Need for Federal Involvement

Like earlier federal forays into gambling law, the Congress that passed the IGBA recognized that anti-gambling legislation was traditionally the purview of the states. However, the problems organized crime presented called for federal involvement. Chief among these problems were jurisdictional problems, lack of resources on the state and local levels, and the corruption of local and state officials.\(^{194}\) The IGBA sought to aid the states where they were unable to enforce state gambling law due to these issues. It was the desire to overcome these issues, rather than a desire to define gambling under federal law, that motivated this legislation.

Jurisdictional limitations are one reason that state and local enforcement of gambling laws were unsuccessful. The operation of gambling enterprises across various jurisdictional lines necessitated coordination among more than one law enforcement agency.\(^{195}\) Unfortunately this cooperation often did not exist, in part due to lack of trust for fear the other law enforcement agency was corrupted by organized crime.\(^{196}\) Because illegal gambling did not operate within state or local borders, the solution was for Congress to give jurisdiction to the federal government, which is not bound by these jurisdictional limits.\(^{197}\) This reduced the need for cooperation between possibly corrupt local and state law enforcement, although the Justice department still encouraged cooperation between states and localities.\(^{198}\)

The IGBA gives the federal government the power to prosecute crimes pursuant to § 1955 under its commerce clause powers.\(^{199}\) Part A of the IGBA contains a Congressional finding that certain gambling enterprises (those that are operated by five or more persons, are continuously operated for 30 days or more or, have a one-day revenue of $2,000 or more) affect interstate commerce.\(^{200}\) This eliminates the need for the federal prosecutor to prove any jurisdictional element of the crime other than that the gambling operation meets the requirements of Part A.\(^{201}\) Placing jurisdiction in the federal government's

\(^{195}\) The Challenge of Crime in a Free Society, supra note 186, at 199.
\(^{196}\) Id.
\(^{198}\) Nixon, supra note 184, at 317.
\(^{201}\) Id.
hands eased the problem of investigating and prosecuting illegal gambling that crossed jurisdictional boundaries.

Another issue that prevented local and state enforcement agencies from adequately enforcing anti-gambling law was their lack of resources. Investigating and developing a case for prosecution often took a significant amount of time and resources while resulting in a relatively small number of arrests. Experienced state-level prosecutors and investigators often did not stay in their positions for long, which created a lack of expertise. Simply put, the state and local law enforcement “lack[ed] . . . sufficient funds to provide adequate manpower or modern equipment” to enforce anti-gambling laws. Giving the federal government jurisdiction over gambling cases “ma[de] available to assist local efforts the expertise, manpower, and resources of the Federal agencies which under existing Federal anti-gambling statutes have developed high levels of special competence for dealing with gambling and corruption cases.” President Nixon supported the Justice Department training investigators, prosecutors, and other professionals at the state and local levels. The resources and manpower of the FBI, as well as other agencies, was one way the federal government got involved to help solve local and state enforcement issues.

The most troubling state and local enforcement issue that led to federal involvement was the corruption of state and local officials. Both President Nixon and members of Congress felt that illegal gambling could only exist in places where there was at least some level of corruption of local officials. Among

202 THE CHALLENGE OF CRIME IN A FREE SOCIETY, supra note 186, at 199.
203 Id.
206 NIXON, supra note 184, at 317.
208 See, e.g., NIXON, supra note 184, at 321 (“For most large scale illegal gambling enterprises to continue operations over any extended period of time, the corruption of corrupt police or local officials is necessary.”); see also Measures Relating to Organized Crime: Hearings before the Subcomm. on Criminal Laws and Procedures of the S. Comm. on the Judiciary, 91st Cong. 394 (1969) (statement of Will Wilson, Asst. Att’y Gen., Criminal Division, Department of Justice) (“The reason local government generally is not effective in this area is, more often than not, bribery or something like it.”); Organized Crime Control: Hearings before Subcomm. No. 5 of the H. Comm. on the Judiciary, 91st Cong. 93 (1970) (statement of Sen. McClellan) (“Title VIII provides new tools for curbing both the large-scale gambling operations themselves and the corruption of local officials on which they depend.”); S. REP. No. 91-617, at 71 (1969) (“The inevitable companion of flourishing gambling activity, moreover, is the bribery and corruption of local law enforcement officials.”).
those corrupted by organized crime were law enforcement officials, prosecutors, politicians, and judges, who “operate[d] as a ‘silent conspiracy’ in support of organized crime.”209 Organized crime considered “ice money,” money used for bribery, as a part of its operating expenses.210 Often it was the local law enforcement officers who took bribes to look the other way, leaving illegal gambling to operate freely.211 Accepting bribes was tempting for local law enforcement officers, who were often underpaid.212 In some places, bribery was not even needed to corrupt local elected officials, who were elected because it was known they would be sympathetic to organized crime.213

The corruption of local enforcement officials made local enforcement of gambling law unsuccessful, and also hindered federal enforcement. For example, in one case, the IRS unknowingly cooperated with corrupt local forces, who warned gamblers of the planned raids on their establishments.214 Organized crime’s corruption of local law enforcement and officials “destroy[ed] local law enforcement as an effective weapon against organized crime.”215 The corruption of local enforcement made state and local efforts to enforce gambling law ineffective, and necessitated federal involvement to strike at illegal gambling enterprises and organized crime in general. In fact, the IGBA includes a provision that makes it illegal to conspire to obstruct enforcement of state or local law in order to further an illegal gambling business.216 The combination of corruption, lack of resources, and lack of jurisdiction convinced the federal government that in order to strike at the heart of organized crime, federal action was needed.

B. Congressional Intent

Congress, when deliberating the IGBA, recognized and respected that gambling was an area of law traditionally

---

211 Id.
215 Id. at 71.
reserved to the states. Congress sought to aid the states in the enforcement of their traditional duties, rather than take over the regulation of gambling. As the Senate Committee Report put it,

The enforcement of criminal laws against gambling and other illegal activities is generally the responsibility of the States and local governments in our Federal system. While the intent of the committee is not to preempt this responsibility, it is its intent to make it possible for the Federal Government to intervene where local and State governments have become, in effect, incapable of law enforcement by reason of the corruption of responsible officials. This limited Federal intervention should serve to reinforce the powers of the States and local governments in our Federal system, rather than to inject the Federal Government into a responsibility traditionally left to the States.  

217

The legislative history echoes this same sentiment in its discussion of the intended role of the federal government with the passage of the IGBA. The new legislation was not intended to preempt local law enforcement, but to expand the forces fighting against illegal gambling and organized crime.  

218

Congress was attempting to encourage local enforcement by providing “an impetus for effective and honest local enforcement” and resources to aid it.  

219

In fact, Congress took the position that “it is essential that the primary responsibility for enforcement of the gambling and corruption laws remain in the hands of state and local officials.”  

220

The legislative history shows that Congress was not concerned with making anti-gambling law a primary responsibility for the federal government to define and enforce. Instead, the focus was on helping the states to enforce their existing law when the challenges presented by organized crime made that difficult for the states to achieve unaided.

When deliberating the IGBA, Congress was not preoccupied with defining the meaning of gambling or what legal standard for the element of chance should be adopted, but rather with defining gambling in terms of its influence on interstate commerce. The elements of § 1955(b)(1)(ii) and (iii), requiring the illegal gambling business be conducted by five or

\[\text{References}\]

more persons, and be in continuous operation for 30 or more days or have a gross revenue of $2,000 dollars in a single day, defined an illegal gambling business in a way that brought those enterprises under federal jurisdiction pursuant to the commerce clause.\footnote{18 U.S.C. § 1955.} The law was designed to target only gambling enterprises of “major proportions.”\footnote{S. REP. No. 91-617, at 73 (1969). The Report goes on to state that It is anticipated that cases in which this standard [referring to § 1955(b)(1)(iii)] can be met will ordinarily involve business-type gambling operations of considerably greater magnitude that this definition would indicate, however, because it is usually possible to prove only a relatively small proportion of the total operations of a gambling enterprise. Thus, the legislation would in practice not apply to gambling that is sporadic or of insignificant monetary proportions. It will reach only those who prey systematically upon our citizens and whose syndicated operations are so continuous and so substantial as to be a matter of national concern.} Debate about the meaning of gambling, therefore, revolved around the size of the gambling rather than the games it covered, or the roles of skill and chance in those games. For example, one congressman was concerned that the IGBA would reach a friendly game of poker.\footnote{116 CONG. REC. 35, 205 (1970) (statement of Rep. Mivka).} The response from Representative Poff was that the law was designed to only reach business-level gambling enterprises, not casual games.\footnote{Id. (statement of Rep. Poff). It is significant that Representative Poff responded to the concern of whether this statute would criminalize a social game of poker with a statement about the size and character of the gambling operation, not whether poker would be considered gambling under a federal definition of gambling. It emphasizes that Congress intended to define an illegal gambling business, not gambling itself.} The standards set by § 1955(b)(1)(ii) and (iii) were meant to focus federal efforts on the gambling enterprises which were most important, and therefore of utmost national concern.\footnote{Measures Relating to Organized Crime: Hearings before the Subcomm. on Criminal Laws and Procedures of the S. Comm. on the Judiciary, 91st Cong. 394 (1969) (statement of Sen. McClellan, Chairman, Subcomm. on Criminal Laws and Procedures).} As one Senator explained,

The approach of this bill is to define an ‘illegal gambling business’ in terms of the number of people involved and in terms of gross receipts and length of operation . . . . This is a sound and necessary approach . . . . [It] focuses the attack on the large-scale gambling enterprises which are the bread and butter of organized crime.\footnote{116 CONG. REC. 602-03 (1970) (statement of Sen. Yarborough).}

Congress sought only to reach the illegal gambling that was influential enough to be brought under its commerce
clause jurisdiction, and significant enough to deal a blow to organized crime when prosecuted.

C. The Meaning of § 1955(b)(2)

Congress's discussion of defining gambling in terms of the number of people involved, gross intake, and longevity stands in stark contrast to what Congress did not discuss—a federal substantive definition of gambling. In the legislative history there is no discussion of what § 1955(b)(2) means. Recall that § 1955(b)(2) states “‘gambling’ includes but is not limited to pool-selling, bookmaking, maintaining slot machines, roulette wheels or dice tables, and conducting lotteries, policy, bolita or numbers games, or selling chances therein.” 227 In the legislative record, explanations of the crime repeatedly name the three elements found in § 1955(b)(1), without any mention of § 1955(b)(2). 228 In the House, the crime was summarized as “mak[ing] large-scale gambling operations in violation of State law a federal offense,” without mention of § 1955(b)(2) adding a federal definitional element to gambling. 229 Even when § 1955(b)(2) was mentioned when summarizing the elements of the crime, Congress did not discuss the section as adding a separate element to the crime. 230 Given the variety of the common law definitions of gambling, and the different ways Congress could chose to define gambling, it seems unlikely that Congress would adopt one of these definitions with no discussion. 231 Keeping in mind that throughout the history of gambling, federal involvement in the regulation of gambling deferred to the states for defining what constituted illegal gambling, it makes sense that Congress intended for gambling to be defined entirely by state law.

Of course it is ultimately the text of the statute rather than the intention of the enacting legislative body that governs. While § 1955(b)(2) may purport to define gambling, what it says is “‘gambling’ includes but is not limited to” a list of games

---

230 S. REP. No. 91-617 at 73, 156 (1969).
231 See supra Part I.
prohibited as gambling. This language indicates Congress’s intention to create an illustrative, non-exhaustive list. The court in United States v. Dicristina examined this list of games and concluded that, because in all of the games listed chance predominates over skill, § 1955(b)(2) adopts a federal predominance test. But all of these games could also fulfill the material element test, and some of them could fulfill the pure chance test. While the list is illustrative, a categorization of the listed possibilities should not serve as a stand-in for a federal definition of gambling when it is not clear such a definition was intended. A categorization of the listed possibilities should not be made to define the characteristics of all the possibilities, morphing their similarities of the listed games into a substantive definition of gambling, when they could stand for more than one possible definition. Further, the canon of ejusdem generis—“where general words follow a specific enumeration of persons or things, the general words should be limited to persons or things similar to those specifically enumerated”—only applies when the statutory meaning is ambiguous, which it is not.

A comparison of the language in § 1955(b)(2) with the list contained in § 1955(e) further supports the argument that § 1955(b)(2) does not adopt a federal substantive definition of gambling. Section 1955(e) states, “[t]his section shall not apply to any bingo game, lottery, or similar game of chance.” It is a canon of construction that “[w]here Congress includes particular language in one section of a statute but omits it in another section of the same Act, it is generally presumed that Congress acts intentionally and purposely in the disparate inclusion or exclusion.” In § 1955(e), Congress chose to finish a non-exhaustive list with a term that describes and categorizes other

---

233 United States v. Dicristina, 726 F.3d 92, 99-100 (2d Cir. 2013) (internal citations omitted).
235 Under the material element test, all of the listed games have chance as a material element in determining the outcome. There are also arguments that bookmaking (as well as beating a bookie) requires skill, which would fulfill the pure chance doctrine. See Christopher T. Pickens, Of Bookies and Brokers: Are Sports Futures Gambling or Investing, and Does It Even Matter?, 14 GEO. MASON L. REV., 227, 250 (2006); Dicristina, 886 F. Supp. 2d at 229 (citing arguments that bookmaking and pool-selling requires substantial skill).
236 Dicristina, 726 F.3d 92 at n.8 (internal quotation and citation omitted).
games that would belong in that list, games that are “similar games of chance.” Conversely, § 1955(b)(2) gives an illustrative list but does not categorize the types of games prohibited.

If Congress were adopting a federal definition of gambling in § 1955(b)(2), it would be strange to do so without articulating that definition, especially given the pivotal role that the chance—skill dichotomy plays in determining the legality of a game. The fact that Congress did comment on the role of chance in § 1955(e), but failed to do so in § 1955(b)(2), serves to make that omission in § 1955(b)(2) even more suspect. Clearly Congress was aware that it could typify the games it sought to prohibit by commenting on the role chance played in those games; however, it did not. Additionally, § 1955(e) does not refer to or claim to modify § 1955(b)(2) in any way.

The textual differences between the IGBA and its proposed predecessor, S. 2022, further illustrate the point that the text of § 1955(b)(2) does not create a substantive definition of gambling. That bill read “[a]s used in this section, the term ‘illegal gambling business’ means betting, lottery, or numbers activity.” Representative McCulloch, who introduced the bill, said that the proposed legislation would “give the Federal Government the necessary weapons to attack all of these activities [casino-type gambling, sports bookmaking, off-track betting, bolita, policy, numbers, large dice games] with the exception of illegal casino-type gambling.” The bill defines a gambling business as these three types of gambling—betting, lottery, and numbers—and it was interpreted that way by Representative McCulloch. The change of language from S. 2022, defining gambling as meaning those three types of games, to § 1955(b)(2) defining gambling as “includes but is not limited to,” is significant. It changes the language from one defining gambling, to a non-exhaustive list that lists some, but not all, of the games that are considered gambling.

The legislative history of the IGBA shows that the main reason for federal interference in the matter of governing gambling

240 1276 F.3d 92 at 100.
241 115 CONG. REC. 10, 736 (1969) (S. 2022 read into record). S. 2022 is similar in structure to § 1955. The quoted portion is contained in the equivalent of what is § 1955(b)(1), followed by the equivalent of what became § 1955(b)(1)(i), (ii), and (iii). S. 2022 contains no provision parallel to § 1955(b)(2), rather § 1955(b)(2) and § 1955(b)(1) were condensed into proposed § 1953A(b) (the quoted language).
243 Id. It should be noted that lottery and numbers includes policy and bolita, and that betting includes bookmaking and off-track betting.
gambling was due to the inability of states and localities to adequately enforce their gambling law. This fits with the history of federal regulation of gambling, when the federal government acted when the states were unable to enforce their anti-lottery policies when challenged by the Louisiana Lottery and when organized crime made corruption, lack of resources, and lack of jurisdiction major factors in the inability of states to enforce gambling law. Congress also recognized the variance in state law regarding the legality of gambling, and did not seek to impose its own federal agenda in place of that variation. Rather, throughout the history of the regulation of gambling, Congress has sought to aid the states when they were incapable of enforcing their own gambling law. With the IGBA, Congress recognized and continued this tradition. Congress did not intend to take over deciding which forms of gaming were prohibited as illegal gambling, but rather to aid the states in enforcing their own definitions of what constituted illegal gambling.

Kaitlyn Dunphy†

† J.D. Candidate, Brooklyn Law School, 2014; B.A., Ithaca College, 2011. Thank you to the staff and editors of the Brooklyn Law Review for their hard work, especially to John Moore, Sabrina Bierer, Michael Tse, and Nathanael Everhart for their invaluable insight and feedback. To my parents, for their endless support and guidance, and to my family and friends, for their love and encouragement.
Prevent Defense

WILL THE RETURN OF THE MULTIYEAR SCHOLARSHIP ONLY PREVENT THE NCAA’S SUCCESS IN ANTITRUST LITIGATION?

INTRODUCTION

In football there is a common defensive formation called “Prevent Defense,” which teams use at the end of a game or right before halftime, in hopes of stopping an opposing team from scoring. The formation positions defensive backs and linebackers, the players responsible for pass coverage, farther away from the line of scrimmage. This strategy makes it exceedingly difficult for the offense to gain substantial yardage on any single play, but allows them to easily and consistently move the football down the field through short gains. By forcing a team to run a greater number of plays, coaches believe that time will expire before the offense has reached a scoring position.

Although teams continue to use this formation, it has received significant criticism for its ineffectiveness. The use of this strategy almost always involves switching from a successful defensive formation to a less tested one, and as a result, defenses frequently allow offenses to score points and win the game. Ignoring a valuable paradigm from one of the sports it regulates, the NCAA recently switched to a preventative defensive strategy by revising a scholarship bylaw in response to antitrust litigation brought by student-athletes.

---

2 Id.
3 Id.
5 Id.
6 Teams rarely utilize Prevent Defense throughout the game, but instead solely resort to it at the end of the game when they have the lead. Id.
The NCAA’s regulations have become “a self-protection measure for the NCAA rather than carefully thought-out rules to protect the student-athlete.”\(^7\): Not only do many of the NCAA’s bylaws fail to protect student-athletes, but many also place undue restrictions on student-athletes\(^8\) and even create harm.\(^9\) Although the NCAA considers itself committed to protecting athletes from the dangers of collegiate athletics,\(^10\) it has recently faced scrutiny for failing to live up to its self-proclaimed purpose.\(^11\) To seek redress, student-athletes have challenged various NCAA bylaws in courtrooms throughout the country, but have achieved limited success.\(^12\)

An example of an unsuccessful challenge occurred in *Agnew v. NCAA*.\(^13\) In this case, the plaintiff-appellants, former college football players, challenged the NCAA’s prohibition of multiyear athletic scholarship awards and the limit on the total number of athletic scholarships a member institution can offer.\(^14\) Plaintiff-appellants alleged that limiting athletic scholarships to one year\(^15\) created anticompetitive effects on the

---

\(^7\) Mary Grace Miller, *The NCAA and the Student-Athlete: Reform is on the Horizon*, 46 U. RICH. L. REV. 1141, 1149 (2012).


\(^9\) The NCAA’s rules fail to take necessary preventive measures to protect collegiate football players from concussions by allowing multiple full-contact practices a week. The NFL, Pop Warner, and many high schools have instituted rules restricting the number of contact practices allowed per week to one. As a result, a college football player receives approximately 70% more contact to the head per season than a professional one. *Real Sports: Think About Them* (HBO television broadcast Nov. 20, 2012).

\(^10\) This includes physical dangers, as well as the dangers of commercialization. The NCAA revised the rules of play in collegiate football to make the game safer, such as moving the kickoff starting line forward and banning the shield-blocking scheme on punts. *Rule Changes Become Official for Several Fall Sports*, NCAA.COM (Oct. 20, 2012), http://www.ncaa.com/news/ncaa/article/2012-08-27/rule-changes-become-official-several-full-season-sports; see also *Money and March Madness: Mark Emmert Interview*, FRONTLINE (Feb. 14, 2011), transcript available at http://www.pbs.org/wgbh/pages/frontline/money-and-march-madness/interviews/mark-emmert.html [hereinafter Mark Emmert Interview] (discussing how the NCAA works to protect student-athletes from professionalism).

\(^11\) Miller, supra note 7, at 1150.

\(^12\) See Christian Dennie, *Changing the Game: The Litigation That May Be the Catalyst for Change in Intercollegiate Athletics*, 62 SYRACUSE L. REV. 15, 51 (2012).

\(^13\) Agnew v. NCAA, 683 F.3d 328, 332 (7th Cir. 2012). See infra Part II.B, for an in-depth discussion of the *Agnew* case.

\(^14\) “If a student’s athletic ability is considered in any degree in awarding financial aid, such aid shall neither be awarded for a period in excess of one academic year nor for a period less than one academic year.” NCAA, 2011–2012 NCAA DIVISION I MANUAL: CONSTITUTION, OPERATING BYLAWS, ADMINISTRATIVE BYLAWS art. 15.3.3.1, at 200 (2011) [hereinafter NCAA DIVISION I MANUAL].

\(^15\) After one year, universities, through the discretion of their coaching staffs, had the option to renew a student’s athletic scholarship for an additional year. See Neil Gibson, Note, *NCAA Scholarship Restrictions as Anticompetitive Measures: The One-
market and “prevented them from obtaining scholarships that covered the entire cost of their college education,” thereby violating the Sherman Antitrust Act. The United States Court of Appeals for the Seventh Circuit ruled in favor of the NCAA and upheld the lower court’s dismissal of the claim.

Despite the dismissal, the NCAA subsequently removed its ban on multiyear scholarships. In October 2011, NCAA’s Division I board of directors adopted a proposal to permit multiyear scholarship offers to Division I student-athletes. Some critics feel the change is a “huge step toward meaningful reform,” but unfortunately this is a mischaracterization. The NCAA—following its historical priority of escaping scrutiny from courts and governmental agencies—enacted a superficial policy that merely provides schools the opportunity to offer multiyear scholarships and fails to resolve the problem of lost scholarships due to an injury or a coach’s boundless discretion.

In addition to insufficiently protecting student-athletes, the new policy undermines the NCAA’s traditional legal defenses of preservation of amateurism and maintenance of competitive balance, which it has used to thwart antitrust

---


15. Agnew, 683 F.3d at 332-34.

16. Id.

17. “This measure” was not mandated “from university presidents, court cases, or other influential sources.” Miller, supra note 7, 1155; see also Steve Wieberg, Multiyear Scholarship Rule Narrowly Survives Override Vote, USA TODAY (Feb. 17, 2012, 7:00 PM), http://usatoday30.usatoday.com/sports/college/story/2012-02-17/multiyear-scholarships-survives-close-vote/53137194/1.

18. “The new rule would allow scholarships to be awarded for as little as two years, for junior college transfers, or as long as four or five years for incoming freshmen.” Multiyear Scholarship Plan Moves On, ESPN (Feb. 17, 2012, 7:37 PM), http://espn.go.com/college-sports/story/_/id/7587582/challenge-ncaa-multiyear-scholarship-plan-falls-short.


20. See YAEGER, supra note 8, at 159-61.

litigation. The new rule erodes the NCAA’s principle of amateurism by allowing universities to compete for recruits with athletic scholarships of different lengths. This imposes a monetary value on an athlete’s ability, and inadvertently acknowledges the possibility of a labor market for student-athletes, which plaintiffs have struggled to identify in past antitrust litigation. Moreover, it encourages unconscionable employee-like contract negotiations that place student-athletes’ academic and athletic goals in direct conflict. Finally, the new rule marks the abandonment of the NCAA’s long-held position that the ban on multiyear scholarships was necessary to prevent schools with greater financial resources from gaining an unfair advantage and thus maintain a competitive balance in college athletics.

This note argues that by eroding its traditional legal defenses, the NCAA exposes itself to stronger antitrust claims by student-athletes and demonstrates that the NCAA’s policy considerations focus on protecting the commercialization of college athletics, not student-athletes. The first section of this note will provide background information that details the history of the NCAA’s athletic scholarship policies, focusing primarily on collegiate football. Part II will discuss the antitrust litigation that motivated the NCAA’s policy shift. It will also highlight the ways in which the NCAA utilizes its procompetitive justifications of amateurism and maintenance of competitive balance. Part III will discuss how the effect of the NCAA’s change in scholarship policy undermines its legal defenses, leaving the NCAA susceptible to stronger antitrust claims. Part IV will address the inability of the policy to effectuate reform that protects student-athletes and discuss the shortsightedness of the NCAA’s attempt to protect itself against antitrust litigation. Finally, this note will explore the benefits of adopting a mandatory multiyear scholarship that guarantees all student-athletes at least four years of athletic scholarship, so long as they maintain academic eligibility and a willingness to participate on the team.

25 Provided that the plaintiff proves anticompetitive effects or behavior of the defendant, “the burden then shifts to the defendant to show the merits of his or her activity by pointing out its procompetitive elements. In other words, the defendant must show that, on balance, the restraint in question functions to enhance competition.” Gibson, supra note 15, at 223-26; see infra Part II.C. The Supreme Court has recognized preservation of amateurism and maintenance of competitive balance as legitimate procompetitive justifications for regulations that create anticompetitive effects. See generally NCAA v. Bd. of Regents of Univ. of Okla., 468 U.S. 85 (1984) [hereinafter Bd. of Regents].

26 The value is determined by the number of expense-covered years the scholarship awards. See infra Part III.

27 See Agnew v. NCAA, 683 F.3d 328, 346-47 (7th Cir. 2012).
I. THE NCAA AND ATHLETIC SCHOLARSHIPS

A. The Founding of the NCAA and Its Principles

Over a century ago, the NCAA developed as the governing body for major collegiate sports, specifically college football, which it continues to service today. Due to the “rugged, violent, and deadly” nature of college football in the early twentieth century, President Theodore Roosevelt called for attempts to “reduce the brutality of the game.” In December 1905, representatives from 62 schools created the Intercollegiate Athletic Association of the United States (IAAUS), a formal organization dedicated to formulating rules and regulations for collegiate athletics. The organization, which was renamed the National Collegiate Athletic Association (NCAA) in 1912, created a football committee that focused on devising rules to alleviate the game’s violence, and ultimately make football “more palatable to the general public.”

Although reducing violence was the impetus that led to the creation of the NCAA, the immediate rise of a national market for collegiate football required the NCAA to focus its attention on amateurism and eligibility rules as early as its initial meeting. The NCAA made the determination that college athletics were for the “amateur athlete,” or someone who “competed only for [the] symbolic or intrinsic benefits” that playing a sport provides. This differentiated the amateur athlete from the paid professional athlete and led the NCAA to ban offering any financial incentive—including athletic scholarships—to recruit an athlete to attend a particular university. But the NCAA lacked the policing resources to enforce these restrictions, which essentially left the regulation

---

30 Id.
31 The NCAA wanted to keep college sports for college students, and prevent skilled professional athletes from “parading under college colors” and “receiving pay . . . for [their] athletic prowess.” The increase in a national market made this difficult because there were high stakes, such as “national prestige and large amounts of money,” available to winning programs. Id. at 34.
32 Id.
33 Article VI of the NCAA’s 1906 bylaws stated that member institutions must take measures “to prevent violations of amateur principles,” including “the offering of inducements to players to enter college or universities because of their athletic abilities or supporting or maintaining players while students on account of their athletic abilities, either by athletic organizations, individual alumni, or otherwise, directly or indirectly.” Id. at 33.
of universities to an honor system and created an atmosphere conducive to perpetual violations.\(^\text{34}\)

Without an effective enforcement mechanism, member institutions defied the regulations that were supposed to protect the principles of amateurism, resulting in college athletes who closely resembled professional athletes. Teams felt pressured to violate the rules because there were no guarantees that their competitors would abide, and compliance placed them at a competitive disadvantage.\(^\text{35}\) Having a successful team was of the utmost importance because it “created a revenue base, strong ties to their communities, willing investors, and media coverage.”\(^\text{36}\) As a result, universities moved away from the NCAA’s idealized vision of athletic programs—a place where amateur athletes played sports as beloved hobbies to supplement their education, unencumbered by contemptible financial incentives.\(^\text{37}\) Instead, universities cultivated an environment where “[a]thletes were putting in long hours of intensive and specialized training to meet the entertainment needs of thousands of discriminating fans” and were provided monetary support for their efforts.\(^\text{38}\)

**B. The Creation of the Student-Athlete and Athletic Scholarship**

As early as the 1930s, the NCAA attempted to enact regulations to maintain the illusion that college athletes were unpaid amateurs. At this time, supporters of universities, or boosters, would commonly “adopt a local high school athlete and ‘put him through college.’”\(^\text{39}\) According to a study by the Carnegie Foundation for the Advancement of Teaching, “subsidization of athletes in some form or another took place at 81 of the 112 colleges and universities studied.”\(^\text{40}\) In 1948, to stop private payments to athletes, the NCAA abandoned its previous position and endorsed athletic scholarships.\(^\text{41}\)


\(^{35}\) See SACK & STAUBROWSKY, supra note 29, at 35-37.


\(^{37}\) SACK & STAUBROWSKY, *supra* note 29, at 35.

\(^{38}\) Id.

\(^{39}\) WALTER BYERS, UNSPORTSMANLIKE CONDUCT: EXPLOITING COLLEGE ATHLETICS 65 (1995).

\(^{40}\) SACK & STAUBROWSKY, *supra* note 29, at 36.

\(^{41}\) Id. at 42; see also BYERS, *supra* note 39, at 65.
scholarship policy, known as the Sanity Code, quoted as “a student-athlete could receive tuition and fees if he showed financial need and met the school’s ordinary entrance requirements; this amounted to a merit award for athletic ability.” With the adoption of the Sanity Code, the NCAA hoped to protect its notion of amateurism and gain control over growing collegiate athletics. Similar to earlier regulations, the NCAA lacked the capability to enforce the Sanity Code and establish punishment for violations, which occurred openly. Due to its ineffectiveness, the NCAA renounced the Sanity Code in 1951, and consequently left a void in athletic scholarship regulation.

The lapse in regulation did not last long as the NCAA solidified the foundation of the modern athletic scholarship in order to avoid potentially costly litigation. Although debate continued to rage over the emergence of athletic scholarships and whether it amounted to “pay for play,” a new problem overshadowed this concern. Courts indicated that they might view NCAA athletes as employees, which posed a significant problem for colleges because such a determination would force them to provide Workmen’s Compensation benefits to injured players. According to these courts, under the Workmen’s Compensation Act, if athletes were given scholarships that paid for tuition and were contingent on their participation in a sport, then these arrangements qualified as employment contracts. Under this immense and potentially costly pressure, universities across the country united to make a

---

42 The Sanity Code was “named in part after a delegate at a previous [NCAA] Convention who called for ‘a return to sanity’ with regard to members following established rules[,] which contains strict regulations regarding financial aid, recruiting, academic standards, institutional control and amateurism.” Chronology of Enforcement, NCAA.ORG, available at http://archive.is/Ea1B (last visited May 6, 2014).
43 Byers, supra note 39, at 67. In addition, “[an athlete] could receive a scholarship exceeding tuition and fees regardless of need if he ranked in the upper 25 percent of his high school graduating class or maintained a B average in college.” Id.
44 See id. at 67-69; see also Sack & Staubowsky, supra note 29, at 44-46.
45 Seven schools refused to comply with the Sanity Code, including “Boston College, the Citadel, Villanova, Virginia Military Institute, Virginia Polytechnic Institute, the University of Maryland, and the University of Virginia,” but they were not expelled from the NCAA because the major Southern Conferences threatened to secede. Sack & Staubowsky, supra note 29, at 46.
46 Byers, supra note 39, at 68.
47 Id. at 68-69.
determination that college athletics were only for “amateurs.”\textsuperscript{50} To reinforce this notion, the NCAA created the term “student-athlete” to establish a clear demarcation between college athletes and employees (professional athletes), inserting the term pervasively throughout its rules and regulations.\textsuperscript{51} As a result of the NCAA’s propagation, the courts and the public began using the term “student-athlete” ubiquitously.\textsuperscript{52}

As an additional measure to prevent Workmen’s Compensation litigation, the NCAA instituted a revised athletic scholarship policy that covered only the expenses associated with attending college.\textsuperscript{53} The hope was that if a player did not receive compensation beyond the cost of education, then there was no payment as an employee.\textsuperscript{54} Although the NCAA alleged that the revitalization of “true amateurism”\textsuperscript{55} motivated this change, the NCAA’s then-director Walter Byers later admitted, “[T]he campaign had nothing to do with the noble ideal of amateurism, but rather addressed the practical consequences of litigation involving worker’s rights.”\textsuperscript{56} The new approach failed to stop under-the-table payments to players, diverted alumni and booster money from players to universities, and pushed collegiate athletics further down the path of commercialization, corruption, and unfairness.\textsuperscript{57} The superficial protective measures merely succeeded in establishing a legal foundation for the NCAA to protect itself from claims that alleged collegiate athletes deserved employee-status.\textsuperscript{58}

\begin{footnotesize}
\begin{enumerate}
\item \textsuperscript{50} Byers, supra note 39, at 69.
\item \textsuperscript{51} See id.
\item \textsuperscript{52} Id.
\item \textsuperscript{53} This included “room, board, tuition, fees, books, and $15 a month for laundry for nine months.” Yasser, supra note 28, at 995.
\item \textsuperscript{54} See Byers, supra note 39, at 72. A position that has stood the test of time, as current president of the NCAA explained in an interview, “We don’t pay our student-athletes . . . . We provide them with remarkable opportunities to get an education at the finest universities on earth.” Mark Emmert Interview, supra note 10.
\item \textsuperscript{55} The NCAA advised member institutions to make recruits sign a statement agreeing with the principles of amateurism, and acknowledging that no employment-duty was created from the fact that scholarships were often contingent on athletic participation. Byers, supra note 39, at 75.
\item \textsuperscript{56} Yasser, supra note 28, at 995. As one commentator observed: “[F]ull-ride athletic scholarship was a marriage of convenience for the NCAA—it made the whole arrangement ‘legal.’” Yasser, supra note 28, at 995-96.
\item \textsuperscript{57} See Byers, supra note 39, at 73.
\item \textsuperscript{58} Similarly, the NCAA is crafting new laws, such as the discretionary multiyear scholarship offer to avoid continued antitrust litigation. This is as opposed to addressing the issues that student-athletes rights are being violated. See infra Part III.
\end{enumerate}
\end{footnotesize}
C. The End of Multiyear Scholarships

The NCAA’s decision to limit the length of athletic scholarships also resulted from an attempt to minimize the costs of collegiate athletics and to increase revenue. From the 1950s to the 1970s, the NCAA’s regulations failed to limit the term of years of an athletic scholarship and the total number of athletic scholarships an institution could award.\(^{59}\) Although this allowed institutions to award multiyear scholarships, it also set the stage for the elimination of this practice.\(^{60}\)

Some colleges were offering only one-year grants to recruits, who were being wooed away by colleges offering ‘no-cut’ four-year grants.... [These] one-year recruiters, who believed that the four-year scholarship colleges had too big an advantage... motivat[ed] [a] not-so-subtle campaign among big-time coaches and athletic directors to place control of athletes’ grants in the hands of coaches instead of scholarship committees.\(^{61}\)

Players increasingly frustrated coaches when they quit or were injured because the coaches could not strip them of their scholarships.\(^{62}\) This, in conjunction with the ever-rising flood of television money and escalating rewards for winning, bred the mentality that scholarships should only go to players who contributed on the field.\(^{63}\)

In January 1973, institutions and coaches asserted their authority by eliminating the multiyear scholarship and limiting scholarships to the one-year renewable offer.\(^{64}\) The motive behind eliminating multiyear scholarships derived from cutting the cost of “deadwood”\(^{65}\) and providing coaches with more control to build winning programs.\(^{66}\) The NCAA, however, framed this as a measure to facilitate a competitive balance and ensure that the recruiting process did not disadvantage universities.\(^{67}\) They argued that a uniform scholarship rule would reduce the recruiting disparity between universities offering only one-year scholarships and those offering

\(^{59}\) Louis Hakim, The Student-Athlete vs. the Athlete Student: Has the Time Arrived for an Extended-Term Scholarship Contract, 2 VA. J. SPORTS & L. 145, 158 (2000).

\(^{60}\) See Yasser, supra note 28, at 996.

\(^{61}\) See BYERS, supra note 39, at 75-76.

\(^{62}\) Yasser, supra note 28, at 1001-03.

\(^{63}\) See BYERS, supra note 39, at 76.

\(^{64}\) Yasser, supra note 28, at 1002.

\(^{65}\) “Deadwood” is defined as: Players who received athletic scholarships, but whose contributions to the team were considered unsatisfactory by coaches because they were not as athletically gifted as anticipated or got injured. BYERS, supra note 39, at 76.

\(^{66}\) Yasser, supra note 28, at 1003.

\(^{67}\) See Hakim, supra note 59, at 158.
extended-term scholarships, thus establishing the maintenance of competitive balance legal defense. Further, the NCAA claimed the new scholarship policy supported the ideal of amateurism because individuals who maintained their scholarship, but were no longer on the team received benefits that went beyond expenses.

The one-year deal existed for more than 40 years, but coaches continued to commonly use the term “full-ride” while recruiting players. Even though the NCAA’s bylaws forbid anything more than a one-year scholarship with the option of renewal, coaches assured promising high school student-athletes that these grants would be renewed so long as the student continued to participate on the team and remain eligible. This once again sounded precariously similar to an employment contract, but the NCAA established a formal requirement that student-athletes sign a letter of intent that reinforced the amateur agreement.

With the letter of intent in place, athletic scholarships became binding contracts. Student-athletes’ protection under these contracts lasted for only one academic year, after which schools were free to release players from a team and vacate their scholarships. As a result,

One-year renewable scholarships have provided the burgeoning college sports industry with a reliable and disciplined source of cheap labor. It is difficult to overstate the kinds of demands coaches can make on players as a condition for the yearly renewal of financial aid. Coaches ask that athletes play with injury, and control their lives on and off the field. Because each season is a tryout for financial aid the next, sports takes priority. An NCAA survey carried out a few years ago found that big-time college football players spend

---

68 Id.
69 SACK & STAROWSKY, supra note 29, at 82-84.
70 Levin, supra note 24 (dispelling the notion that athletic scholarships are always a four-year guaranteed education or full ride).
71 See Hakim, supra note 59, at 172-73.
72 BYERS, supra note 39, at 75.
73 See generally Sean Hanlon, Athletic Scholarship as Unconscionable Contracts of Adhesion: Has the NCAA Fouled Out?, 13 SPORTS LAW. J. 41 (2006) (explaining how the athletic-scholarship developed into a recognized contract). The letter of intent also requires that student-athletes sign their respective school’s Statement of Financial Aid, which defines the terms, conditions, and amount of the athletic award. Although the Statement of Financial Aid is made between each school and their respective scholarship athletes, the actual Statements of Financial Aid are uniform contracts that do not vary from school to school. Id. at 69-70.
74 Yasser, supra note 28, at 1003.
an average of 44.8 hours a week on their sport in addition to time in the classroom.\textsuperscript{75}

As Ray Yasser aptly stated, despite these demands, “The school’s only obligation to an athlete who gives his or her blood, sweat, and tears is to notify promptly the athlete of the nonrenewal decision.”\textsuperscript{76}

Under this system the NCAA has grown into “a voluntary unincorporated association that governs more than 1,200 colleges, universities, athletic conferences, and sports organizations; 380,000 student-athletes; and eighty-eight championship events in three divisions.”\textsuperscript{77} A more accurate portrayal of the NCAA is a commercialized big business that benefits the NCAA, member institutions, corporate sponsors, and everyone else except those whose skills are marketed.\textsuperscript{78} The biggest collegiate sports such as football and men’s basketball “generate more than $6 billion in annual revenue,” a profit exceeding some professional sports.\textsuperscript{79} College coaches can earn a salary as high as or higher than professional coaches.\textsuperscript{80} The commercialization of collegiate athletics, focus on profit maximization, and continuous scandals\textsuperscript{81} indicate that the NCAA has strayed from its stated goals to “promote student-athletes and college sports through public awareness . . . [,] protect student-athletes through standards of fairness and integrity . . . [,] prepare student-athletes for lifetime

\begin{footnotes}
\item[75] Sack, supra note 22.
\item[76] Yasser, supra note 28, at 1003.
\item[77] Dennie, supra note 12, at 16.
\item[78] Miller noted, Yet the student-athlete sees none of the money that exchanges hands as a result of his or her performance. For instance, big college football teams . . . bring in between $40 million and $80 million in profits a year, even after paying coaches multimillion-dollar salaries. The student-athlete is granted a scholarship that often fails to cover the true cost of living, and thus he or she frequently lives below the poverty line. The student-athlete is exploited.
\item[80] “Ohio State just agreed to pay Urban Meyer $24 million over six years.” Id.
\item[81] Miller noted, Scandals have recently crowded the newspapers and sports blogs with stories of one football player or another selling his own jersey for a profit or accepting money from a booster. These scandals are unnerving because the NCAA’s bylaws strictly prohibit a student-athlete from profiting from his or her athletic performance.
\end{footnotes}
leadership, and provide student-athletes and college sports with the funding to help meet these goals.”82

D. The Return of the Multiyear Option

In October 2011, the NCAA’s Division I board of directors adopted a proposal to change their policy on athletic scholarships.83 The new rule, which allows schools the option to provide multiyear scholarships, went into effect immediately, and by National Signing Day84 in February 2012 some schools already offered multiyear scholarships.85 Although a few schools signed student-athletes to multiyear scholarships, a majority of member institutions met the overnight change of the four-decade-old scholarship policy with resistance.86 A substantial number of member institutions formally opposed the new rule and demanded a repeal vote.87 The option to offer multiyear scholarship barely survived the repeal vote—“of 330 institutions voting, 62.12 percent voted to override the legislation. A 62.5 percent majority of those voting was required to override legislation.”88 The opposing member institutions failed to gain the two extra votes necessary to repeal the new rule, and thus schools retain the option to make multiyear rather than one-year offers.89

The NCAA, led by its current president Mark Emmert, argues that elimination of the prohibition on multiyear scholarships is part of a larger initiative to enhance athletes’ welfare.90 Such an explanation ignores the tradition of the NCAA’s policy changes. History suggests that the change results from the NCAA’s attempt to avoid antitrust claims by

---

82 Dennie, supra note 12, at 16-17.
83 Multiyear Scholarship Plan Moves On, supra note 21; see also Gibson, supra note 15, at 242.
87 See Hosick, supra note 86.
88 Id.
89 Id.
student-athletes that directly challenged the bylaw, and the attention that the bylaw garnered from United States Department of Justice Antitrust Division, who “informed the NCAA a little less than two years ago that it was looking into the single-year restriction and whether it restrained competition among schools for top players.” Yet, the NCAA’s strategic attempt to insulate itself from antitrust litigation comes at the cost of abandoning its most common legal defenses: preservation of amateurism and maintenance of competitive balance.

II. ANTITRUST SCRUTINY AND THE NCAA

The NCAA continues to enforce bylaws that create restrictions and requirements for student-athletes that essentially treat them as an unpaid labor force and leave them powerless to seek recourse internally. As a result, student-athletes resort to filing lawsuits that claim the NCAA’s bylaws place unreasonable restraints on them. Because of this tension, the NCAA has been “no stranger to protracted litigation and has been involved in a plethora of lawsuits relating to nearly every conceivable area of the law.” The NCAA, however, has a strong tradition of success in the courtroom, including antitrust litigation.

Student-athletes often bring claims against the NCAA for violations of Section 1 of the Sherman Act. Section 1 of the Sherman Act states that “[e]very contract, combination in the form of trust or otherwise, or conspiracy,” that creates an unreasonable restraint on trade is illegal. To succeed in an antitrust litigation under Section 1 of the Sherman Act, a plaintiff must prove a contract, combination or conspiracy, an unreasonable restraint on trade in a relevant market, and an injury.

---

91 Agnew v. NCAA, 683 F.3d 328 (7th Cir. 2012).
92 Wieberg, supra note 20.
93 See Miller, supra note 7, at 1150.
94 Dennie, supra note 12, at 22.
95 Id.
96 “Since the mid-1970s, plaintiffs have brought a great number of antitrust claims against the NCAA before federal courts. Only twice, however, have these courts recognized NCAA violations of the Sherman Act, first in NCAA v. Board of Regents of the University of Oklahoma, and later in Law v. NCAA.” Gibson, supra note 15, 208 n.22. This is in large part due to the NCAA’s time-honored legal defenses of amateurism and maintenance of competitive balance. See Dennie, supra note 12, at 22.
98 See generally Standard Oil Co. v. United States, 221 U.S. 1 (1911) (interpreting 15 U.S.C § 1).
99 Agnew v. NCAA, 683 F.3d 328, 335 (7th Cir. 2012) (quoting Denny’s Marina, Inc. v. Renfro Prods., Inc., 8 F.3d 1217, 1220 (7th Cir. 1993)).
The NCAA faced challenges under Section 1 of the Sherman Act with increased frequency following Board of Regents,100 the seminal and successful application of antitrust principles to the NCAA.101 In Board of Regents, the Supreme Court determined that the NCAA violated Section 1 of the Sherman Act by restricting “both the quantity of college football games televised and the number of televised games allowed to a given team in a single season.”102 The Court also established a precedent that “the NCAA is not exempt from the strictures of the Sherman Act merely because it is a nonprofit entity,” and further indicated that “all the regulations passed by the NCAA are subject to the Sherman Act.”103 Despite the Court’s language that “all” regulations are subject to antitrust scrutiny, courts continue to struggle to apply the Sherman Act to the NCAA’s bylaws.104

A. A Dichotomous Antitrust Approach to the NCAA

The difficulties courts face in applying the Sherman Act to the NCAA’s bylaws largely stem from the dichotomous approach adopted after Board of Regents.105 Rather than apply a single approach to all NCAA regulations, the courts established a “two-pronged antitrust approach.”106 The first approach applies to cases that involve obvious commercial restraints, such as output and price restraints on televised college football.107 With obvious commercial restraints, Board of Regents established precedent to apply a stringent balancing test that weighs the plaintiff’s anticompetitive complaint against the defendant’s procompetitive justifications to determine if the regulation creates an unreasonable restraint.108

The second approach applies to regulations that promote noncommercial goals, such as rules of play and eligibility.109 This

101 Id.
102 Gibson, supra note 15, at 228.
103 Agnew, 683 F.3d at 338-39 (describing the interpretation and legacy of the Board of Regents decision).
105 See generally id. (explaining the dichotomous approach courts use in assessing antitrust litigation against the NCAA).
106 Id. at 340.
108 See Lazaroff, supra note 104, at 340.
approach derived from Justice Stevens’s “now famous (perhaps infamous) dicta” 110 in Board of Regents:

The identification of this “product” with an academic tradition differentiates college football from and makes it more popular than professional sports to which it might otherwise be comparable . . . . In order to preserve the character and quality of the “product,” athletes must not be paid, must be required to attend class, and the like. And the integrity of the “product” cannot be preserved except by mutual agreement; if an institution adopted such restrictions unilaterally, its effectiveness as a competitor on the playing field might soon be destroyed. Thus, the NCAA plays a vital role in enabling college football to preserve its character, and as a result enables a product to be marketed which might otherwise be unavailable. In performing this role, its actions widen consumer choice—not only the choices available to sports fans but also those available to athletes—and hence can be viewed as procompetitive. 111

Courts interpreted this language to establish a more lenient standard for the NCAA in advancing procompetitive justifications. Despite the fact that noncommercial regulations may place economic restraints on student-athletes, courts accept, without demonstration by the NCAA, that these rules are justified by preservation of amateurism or maintenance of competitive balance. 112 That a number of district courts held that various bylaws pertaining to student-athlete eligibility do not violate antitrust regulation exemplifies the leniency of this approach. 113

The Supreme Court, however, has never determined “whether and when the Sherman Act applies to the NCAA and its member schools in relation to their interaction with student-athletes.” 114 Because the Supreme Court has not weighed directly on the issue, student-athletes continue to use antitrust law as an avenue to challenge the restrictions imposed upon them. Recently, “lower federal courts are also beginning to blur the distinction between restraints on players and restraints on other actors.” 115 One of the claims that appears strongly situated to

---

110 Lazaroff, supra note 104, at 339.
111 Bd. of Regents, 468 U.S. at 101-02.
112 See Lazaroff, supra note 104, at 339 (referencing Bd. of Regents, 468 U.S. 85).
113 See, e.g., Smith v. NCAA, 139 F.3d 180 (3d Cir. 1998); Banks v. NCAA, 977 F.2d 1081 (7th Cir. 1992); McCormack v. NCAA, 845 F.2d 1338 (5th Cir. 1988); Gaines v. NCAA, 746 F. Supp. 738 (M.D. Tenn. 1990).
114 Agnew v. NCAA, 683 F.3d 328, 339 (7th Cir. 2012).
115 In fact, in some cases, antitrust claims have been rejected summarily because jurists have determined that antitrust laws have no application to restraints on amateur student-athletes. In other cases, courts have engaged in antitrust analyses but concluded that the NCAA acted lawfully in imposing restraints.
demonstrate that an NCAA bylaw violates Section 1 of the Sherman Act is the challenge to the previous ban on multiyear scholarships.

B. Antitrust Challenges to Athletic Scholarship Bylaws

Agnew v. NCAA\textsuperscript{116} applied Section 1 of the Sherman Act to NCAA scholarship bylaws, and provided the context in which the NCAA utilizes its legal defenses of amateurism and maintenance of a competitive balance. In this case, NCAA student-athletes Joseph Agnew and Patrick Courtney directly attacked the NCAA’s limitation on athletic-scholarships, claiming that the NCAA’s bylaws that limit athletic scholarships to one-year\textsuperscript{117} and the total number of athletic scholarships available\textsuperscript{118} violated Section 1 of the Sherman Act.\textsuperscript{119} Both Agnew, who played football for Rice University in 2006, and Courtney, who played for North Carolina A&T in 2009, received one-year athletic scholarships to play football at their respective universities.\textsuperscript{120} Unfortunately, Agnew and Courtney suffered career-ending injuries while playing football during their college tenures, and their universities exercised the right to not renew these players’ scholarships.\textsuperscript{121} Agnew and Courtney sued the NCAA, claiming the imposed cap “on the number of scholarships given per team and the prohibition of multi-year scholarships prevented them from obtaining scholarships that covered the entire cost of their college education.”\textsuperscript{122} The plaintiffs alleged that this violated Section 1 of the Sherman Act because, absent these restrictions, colleges would offer multi-year scholarships to stay competitive, and they would have received them.\textsuperscript{123} In response, “the NCAA filed a motion to dismiss claiming that the plaintiffs failed to identify a relevant market, failed to allege facts sufficient to show that the NCAA injured

\textsuperscript{116} See Agnew, 683 F.3d 328.
\textsuperscript{117} NCAA DIVISION I MANUAL, supra note 14, at 200.
\textsuperscript{118} Id. at 207.
\textsuperscript{119} Agnew, 683 F.3d 328.
\textsuperscript{120} Id. at 332.
\textsuperscript{121} Id.
\textsuperscript{122} Id.
\textsuperscript{123} Id.

Further, some courts have suggested that, at least at the preliminary stages of litigation, NCAA athlete claims can move forward.

Lazaroff, supra note 104, at 344.
competition in a relevant market, and failed to allege facts sufficient to show an injury.”

The first element in an antitrust challenge requires a plaintiff student-athlete to demonstrate “a contract, combination, or conspiracy.” The NCAA is a voluntary unincorporated association that governs more than 1,200 colleges, universities, athletic conferences, and sports organizations, which promulgates “rules and regulations to monitor a variety of issues facing member institutions, conferences, student-athletes, and coaches, including bylaws governing amateurism, recruiting, eligibility, financial aid, and practice and playing seasons.”

As the court in Agnew stated, the member institutions have unquestionably agreed to abide by these rules and regulations, and therefore the showing of an agreement is not an issue when student-athletes challenge a bylaw. The second element requires a plaintiff student-athlete to demonstrate “an unreasonable restraint of trade in a relevant market.” To do so, the plaintiff must first establish a relevant market. Agnew and Courtney attempted to challenge the NCAA scholarship regulation as a restriction on the market for bachelor’s degrees. This is not typically the focus of challenges to Section 1 of the Sherman Act and it proved fatal to Agnew and Courtney’s claim. The district court held that the bachelor’s degree market was not a cognizable market under the Sherman Act because bachelor’s degrees cannot be bought through tuition payments. Rather, bachelor’s degrees are earned by satisfying requirements, and student-athletes are only provided an opportunity to fulfill these requirements. There is no exchange of a bachelor degree for participation on the athletic field. The district court also foreclosed the possibility that a student-athlete labor market could be a cognizable market and dismissed the plaintiffs’ complaint.

124 Id. at 333.
125 Id. at 335 (quoting Denny’s Marina, Inc. v. Renfro Prods. Inc., 8 F.3d 1217, 1220 (7th Cir. 1993)).
126 Dennie, supra note 12, at 16.
127 Id. at 17.
128 Agnew, 683 F.3d at 335.
129 Id. (quoting Denny’s Marina, Inc. v. Renfro Prods., Inc., 8 F.3d 1217, 1220 (7th Cir. 1993)).
130 Id. at 333.
131 Id. at 338.
132 Id.
133 Id.
134 Id.
On appeal, the court emphasized that the plaintiffs must “describe the rough contours of the relevant market in which anticompetitive effects may be felt.” 135 Plaintiffs failed to meet this burden for both the market for bachelor’s degrees and the market for student-athlete labor. The Agnew court suggested that “[t]he proper identification of a labor market for student-athletes . . . would meet plaintiff’s burden of describing a cognizable market under the Sherman Act.” 136 This contradicted a prior decision, which dismissed the argument that scholarship athletes could be considered a labor market because “schools do not engage in price competition for players,” 137 “the value of [a] scholarship is based upon the school’s tuition and room and board,” 138 and supply and demand does not determine the worth of student-athletes’ labor. 139 The Agnew court recognized that a market was certainly at play, stating “a transaction clearly occurs between a student-athlete and a university: the student-athlete uses his athletic abilities on behalf of the university in exchange for an athletic and academic education, room, and board.” 140 This dictum provides support for recognizing a nationwide labor market for student-athletes under the Sherman Act, 141 and contradicts the belief that bylaws affecting student-athletes, such as scholarship policies are not commercial. 142 Similarly, the Agnew court stated, “No knowledgeable observer could earnestly assert that big-time college football programs competing for highly

135 Id. at 345.
136 Id. at 346.
137 Id. at 346 (citing Banks v. NCAA, 977 F.2d 1081, 1091 (7th Cir. 1992)).
138 Banks, 977 F.2d 1091.
139 Agnew, 683 F.3d at 346 (citing Banks, 977 F.2d 1081, 1091).
140 Id. at 338.
141 It is important to note that it would not be enough for a plaintiff class to simply “write the words ‘nationwide labor market for student athletes’ on paper.” Order Granting Defendant’s Motion to Dismiss at 16, Rock v. NCAA, 928 F. Supp. 2d 1010, (S.D. Ind. 2013) (No. 12-CV-1019). Instead, a plaintiff “must properly identify the labor market at issue, plead its rough contours, or account for the commercial reality of the transaction.” Id.
142 The belief that scholarship and eligibility rules are not commercial is “an outmoded image of intercollegiate sports that no longer jibes with reality.” Agnew, 683 F.3d at 340 (quoting Banks, 977 F.2d at 1099 (Flaum, J., dissenting)). The Seventh Circuit seems to accept this dictum. The Seventh Circuit followed the Agnew court’s guidance in resolving the NCAA’s motion to dismiss in Rock v. NCAA, stating that:

[T]he NCAA’s one-year scholarship limit and the cap on the number of scholarships are financial aid rules, not eligibility rules. As financial aid rules, those bylaws are not inherently or obviously necessary for the preservation of amateurism, the student-athlete, or the general product of college football. Accordingly, unlike eligibility rules, financial aid rules are not deserving of a procompetitive presumption . . . at the motion-to-dismiss stage.

sought-after high school football players do not anticipate economic gain from successful recruiting program.”

Although the Agnew court recognized that the labor market for student-athletes may be cognizable under the Sherman Act, which will likely provide guidance to future student-athlete plaintiffs to properly identify relevant market, such as the class in Rock v. NCAA, it stated that the NCAA bylaws prohibiting multiyear scholarships and limiting the number of scholarships do not necessarily violate the Sherman Act. Future plaintiffs still need to prove the additional component that these regulatory controls are an unreasonable restraint. Despite the fact that the climate is changing, the legacy of the Supreme Court’s decision in Board of Regents suggests that there is still “a presumption in favor of certain NCAA rules when it stated: It is reasonable to assume that most of the regulatory controls of the NCAA are...procompetitive because they enhance public interest in intercollegiate athletics.” Furthermore, the Court suggests that many of the NCAA bylaws are necessary to distinguish and preserve the “character and quality of the product.”

C. Unreasonable Restraint and the Rule of Reason

Typically, the focus of Section 1 Sherman Act cases is whether a regulation poses an unreasonable restraint. To determine whether a restraint is unreasonable, courts “focus on the competitive effects of challenged behavior relative to such

---

143 Agnew, 683 F.3d at 341.
144 The named plaintiff John Rock represents a class of individuals who received an athletic-based scholarship for at least one year and had it reduced or not renewed, and subsequently was forced to pay tuition. Rock was an accomplished high school quarterback, who attended Gardner-Webb University in North Carolina on a football scholarship. Despite having been the team’s starting quarterback and captain, Rock’s scholarship was revoked when the school replaced the head coach. Like in Agnew, the plaintiff alleges that had it not been for the NCAA’s prohibition on multiyear scholarships and the limit on the overall number of scholarships a university can offer, he would have received a scholarship that covered the full cost of his education. The focus of the complaint addresses the labor market for student-athletes as the relevant market, and the bylaws as an unreasonable restraint on that market, attempting to correct the shortcomings of the plaintiff’s in Agnew. Complaint at 3-6, Rock v. NCAA, 928 F. Supp 2d 1010 (S.D. Ind. 2013) (No. 12-CV-1019). The plaintiffs in Rock eventually amended the original complaint to narrow the proposed market to the “market for the labor of Division I football student athletes.” See infra note 173.
145 Agnew, 683 F.3d at 341.
146 Id. (quoting NCAA v. Bd. of Regents of Univ. of Okla., 468 U.S. 85, 117 (1984)).
147 The product the court refers to is college football. Bd. of Regents, 468 U.S. at 117.
alternatives as its abandonment or a less restrictive substitute.”

Courts apply a balancing test, known as the “rule of reason,” to assess whether the NCAA bylaw that prohibits multiyear scholarships creates an unreasonable restraint. The rule of reason analysis has three criteria: First,

the plaintiff [must] show that the agreement has a substantially adverse effect on the competition; [second], the defendant presents some evidence of the procompetitive virtue of the challenged behavior; [and third], the plaintiff shows that the challenged conduct is not necessary to achieve the procompetitive justifications put forth by defendant or that those justifications can be achieved in a less restrictive manner.

Under the first step in the rule of reason, the plaintiff must show an actual restraint on the quantity and quality of output and price. The restraint on multiyear scholarships prevents member institutions and student-athletes from constructing scholarship agreements that each might find more favorable.

The Rock complaint indicated, picking up where the Agnew plaintiffs left off, that Bylaw 15.3.3.1 is “a blatant price-fixing agreement and restraint between member institutions of the [NCAA]. For years, NCAA member institutions unlawfully conspired to maintain the price of student-athletes’ labor at artificially low levels by agreeing never to offer student-athletes athletics-based scholarships of a duration in excess of one year.” As the Rock complaint and the Agnew court suggest, absent the limitation on scholarship offers, “member schools would choose to alter the price of the opportunity being sold by

---

148 Agnew, 683 F.3d at 335 (quoting Phillip Areeda, Antitrust Law ¶ 1500, at 362-63 (1986)).
149 As one commentator points out, “Both the sport cases and the pervasive trend in antitrust jurisprudence support th[e] conclusion” that courts will evaluate an NCAA mandate under the rule of reason, even if price fixing, usually analyzed under the per se approach, was implicated. Yasser, supra note 28, at 1010. The Court in Board of Regents elucidated: “What is critical is that this case involves an industry in which horizontal restraints on competition are essential if the product is to be available at all.” Bd. of Regents, 468 U.S. 85, 101 (1984); Some horizontal restraints are necessary to the NCAA and its member institutions if the product is to exist at all such as rules of the game, size of the fields, etc. See generally Yasser, supra note 28.
150 Pekron, supra note 109, at 32-33.
151 Yasser, supra note 28, at 1011-25.
152 Id. at 1012.
153 See supra note 144.
154 Bylaw 15.3.3.1 limits scholarships to one year. It reads, “If a student’s athletics ability is considered in any degree in awarding financial aid, such aid shall neither be awarded for a period in excess of one academic year nor for a period less than one academic year.” NCAA DIVISION I MANUAL, supra note 14, at 200.
offering multiyear scholarships in order to compete more effectively for talented players against schools that choose to offer only one-year deals.”

The second step of the rule of reason shifts the burden to the defendant to provide evidence that the restraint offers a justifiable procompetitive effect. Under this step, the burden falls on the NCAA to demonstrate that the one-year scholarship rule provides “procompetitive effects that outweigh the anticompetitive ones.” To meet its burden, the NCAA will undoubtedly argue that the bylaw limiting scholarships to one year “preserves amateurism and helps to maintain a competitive balance.” Not only were these the principles on which the NCAA instituted the rule, but the NCAA has also successfully used these defenses in a number of cases. As the court stated in Board of Regents, “maintaining a competitive balance among amateur athletic teams is legitimate and important.” Additionally, courts have upheld the “NCAA’s efforts to maintain a discernible line between amateurism and professionalism and protect amateur objectives,” as a procompetitive justification, despite the fact that “the NCAA has not distilled amateurism to its purest form.”

The last step of the rule of reason allows the plaintiff an opportunity to demonstrate that the actual restraint is not necessary or that it is overly restrictive. Although courts have determined that the NCAA regulations dictating eligibility “fall comfortably within the presumption of procompetitiveness,” courts have also indicated that the prohibition on multiyear scholarships falls into a separate category of rules. Unlike eligibility regulations, the scholarship bylaw fails to distinguish between professional and amateur sports. This distinction is

---

156 Yasser, supra note 28, at 1012.
157 Should a court, as the dicta in Agnew suggests, recognize the labor market for student-athletes. Agnew v. NCAA, 683 F.3d 328, 341 (7th Cir. 2012).
158 Yasser, supra note 28, at 1013.
159 Id.
160 Id.
161 See, e.g., Smith v. NCAA, 139 F.3d 180 (3d Cir. 1998); Banks v. NCAA, 977 F.2d 1081 (7th Cir. 1992); McCormack v. NCAA, 845 F.2d 1338 (5th Cir. 1988); Gaines v. NCAA, 746 F. Supp. 738 (M.D. Tenn. 1990).
163 Gaines, 746 F. Supp. at 743.
164 McCormack, 845 F.2d at 1345.
165 Pekron, supra note 109, at 31-34.
166 See Agnew v. NCAA, 683 F.3d 328, 343 (7th Cir. 2010); see also In re NCAA I-A Walk-On Football Players Litig., 398 F. Supp. 2d 1144 (W.D. Wash. 2005).
necessary to preserve amateurism, because courts “consider players who receive nothing more than educational costs in return for their services to be ‘unpaid athletes.’”\(^{167}\) Therefore, whether an athlete receives one or four years of scholarship grant does not affect a court’s interpretation of whether he is a professional or amateur.\(^{168}\) The court in \textit{Agnew} further indicated that the bylaw does not implicate maintenance of a competitive balance, because the NCAA survived without a prohibition on multiyear scholarships until 1973. Further, numerous other less restrictive methods exist to achieve a competitive balance such as restricting alumni donation or recruiting budgets.\(^{169}\) If a court were to apply the rule of reason to the NCAA’s prohibition on multiyear scholarships, a student-athlete plaintiff has a reasonable case to prove that the bylaw fails the analysis.

\section*{D. Injury and Future Cases}

Finally, to succeed in demonstrating that the NCAA’s regulation violated Section 1 of the Sherman Act, a plaintiff student-athlete must demonstrate an accompanying injury.\(^{170}\) Plaintiffs challenging Bylaw 15.3.3.1 alleged that they would have been able to secure a guaranteed four- or five-year scholarship, which would have protected them from losing their aid once they were injured or a coaching change occurred, had it not been for prohibition on multiyear scholarships.\(^{171}\) Despite the \textit{Agnew} plaintiffs’ failure to assert a relevant market, the court indicated that it is likely that the one-year scholarship violates Section 1 of the Sherman Act.\(^{172}\) With the \textit{Agnew} court’s willingness to recognize a labor market for student-athletes if a plaintiff properly identifies it, the \textit{Rock} class has a better opportunity to successfully challenge the bylaw than ever before.\(^{173}\)

\begin{flushright}
167 \textit{Agnew}, 683 F.3d at 344.
168 \textit{Id.}
169 \textit{See Complaint at 8, Rock v. NCAA, 928 F. Supp 2d 1010 (S.D. Ind. 2013) (No. 12-CV-1019).}
170 \textit{Agnew}, 683 F.3d at 335.
171 \textit{Id.}, at 332-33; \textit{see also Complaint, supra note 169, at 23.}
172 \textit{Agnew}, 683 F.3d at 345-47.
173 Initially, the \textit{Rock} class struggled to get through the pleading stage. The court granted the NCAA’s motion to dismiss on the grounds that the plaintiffs failed to identify a legally cognizable market. The plaintiff’s alleged market, “labor market for student athletes,” was fatally broad. The proposed market was too broad because it lumped “all student-athletes into the same labor market without accounting for germane differences such as gender and sport played.” \textit{Rock v. NCAA, 928 F. Supp. 2d 1010, 1022 (S.D. Ind. 2013).}
\end{flushright}
III. PREVENT DEFENSE: A SHIFT FROM LEGAL DEFENSES

By lifting the ban on multiyear scholarships in 2011, the NCAA has attempted to insulate itself from antitrust litigation. The new policy, which gives member institutions the option to offer multiyear scholarships, comes at the cost of eroding the NCAA’s most common legal defenses: preservation of amateurism and maintenance of competitive balance. The new rule betrays the NCAA’s ideal of amateurism by inadvertently acknowledging a labor market for student-athletes; the rule quantifies a price on their athletic ability, promotes competition over student-athletes, and demonstrates that supply and demand govern the market. In addition, the rule encourages unconscionable employee-like contract negotiations that place student-athletes’ academic and athletic goals in direct conflict. More blatantly, the new rule abandons the NCAA’s argument that the one-year scholarship provided the procompetitive effects necessary to maintain competitive balance. The harsh resistance and attempt to repeal the new policy by member institutions exemplifies this abandonment and highlights the “legitimate concerns,” raised by these schools, when the NCAA regulates haphazardly.\(^{174}\)

A. Abandonment of Amateurism

Critics admonish the NCAA’s antiquated notion of amateurism, argue that commercialization permeates NCAA, and indicate that the only individuals prevented from benefitting from the system are the student-athletes who generate billions of dollars in revenue for the NCAA’s member institutions.\(^{175}\) In an

\(^{174}\) Wieberg, supra note 20.

\(^{175}\) These critics may have evidence on their side because the NCAA is riddled with scandals of players already being paid under the table. In addition,
interview with the Public Broadcasting Service, Mark Emmert, current president of the NCAA, addressed these concerns and suggested that the NCAA’s most important priority was to prevent the commercialization of college athletics. Even though Mr. Emmert recognizes the challenge of preserving the amateur status of student-athletes in the modern era of commercialized college athletics, his promotion of the reform of the one-year scholarship rule undermines the very principle of amateurism by identifying a student-athlete labor market and by introducing employee-like contracts.

1. Identifying a Student-Athlete Labor Market

Prior to Agnew v. NCAA, courts opined, “[T]he market for scholarship athletes cannot be considered a labor market, since schools do not engage in price competition for players, nor does supply and demand determine the worth of student-athletes’ labor.” By contrast, the Agnew court suggested that there is obviously a market at play, and explained that the only reason schools do not “engage in price competition for student-athletes is that other NCAA bylaws prevent them” from doing so. With the enactment of the option to offer multiyear scholarships, NCAA bylaws no longer prohibit, and in fact encourage, price competition for student-athletes, demonstrating

[i]n 2010, despite the faltering economy, a single college athletic league, the football-crazed Southeastern Conference (SEC), became the first to crack the billion-dollar barrier in athletic receipts. The Big Ten pursued closely at $905 million. That money comes from a combination of ticket sales, concession sales, merchandise, licensing fees, and other sources—but the great bulk of it comes from television contracts.

Taylor Branch, The Shame of College Sports, ATLANTIC MONTHLY, Oct. 2011, available at http://www.theatlantic.com/magazine/archive/2011/10/the-shame-of-college-sports/308643; see also Nocero, supra note 79 (“College football and men’s basketball have become such huge commercial enterprises that together they generate more than $6 billion in annual revenue, more than the National Basketball Association.”).  

Mark Emmert stated,

I think the biggest challenge that faces intercollegiate athletics right now is, in fact, trying to protect the notion of intercollegiate athletics as a place where student-athletes compete... So when we talk about the creeping commercialization of it, what we’re concerned about—what I’m concerned about—is making sure that we maintain that preprofessional amateur status of the student-athletes while recognizing that there’s increasingly greater interest in the whole nature of athletics in America.

Mark Emmert Interview, supra note 10.

Agnew, 683 F.3d at 346 (referencing Banks v. NCAA, 977 F.2d 1081, 1091 (7th Cir. 1992)).

Id.
that the value of the scholarship is based upon the supply of and demand for players.

The NCAA’s new scholarship rule provides universities with a template to formally engage in classical price competition over student-athletes. Prior to October 2011, student-athletes could “neither be awarded [an athletic scholarship] for a period in excess of one academic year nor for a period less than one academic year.”\(^{179}\) To comply with this bylaw, member institutions could only offer identical one-year renewable scholarships to high school athletes. Alternatively, the new rule provides member institutions the discretion to offer one-year renewable scholarships or up to a five-year guaranteed scholarships.\(^{180}\) Colleges have taken,\(^{181}\) and will continue to take advantage of this rule by offering athletic-scholarships of various lengths, inevitably using multiyear scholarships as a recruiting tactic to persuade highly sought-after student-athletes to attend their school over another. An athletic scholarship awarded for up to four or five years guarantees a gifted athlete funding for a full education and provides “a significant incentive to select a university offering a four-year aid package over other schools offering only one-year scholarships with merely the possibility for renewal.”\(^{182}\) According to CNN in 2011, “[t]he sticker price of living and studying for a year at a typical private college rose 4.3% to $42,224.”\(^{183}\) Each year this cost rises as tuition at public, community, and private colleges across the country escalates,\(^{184}\) forcing many students to take out loans to pay for their education and pushing our nation closer to the brink of a student debt crisis.\(^{185}\) College recruiters will certainly take advantage of this frightening reality, by emphasizing the particular monetary value on the offer they are extending to prospective recruits, which now can vary significantly.

As a representative from Indiana State University astutely indicated in opposition to the new scholarship policy, “to get into bidding wars where one school offers a 75 percent

---

\(^{179}\) NCAA DIVISION I MANUAL, supra note 14, at 200.

\(^{180}\) See supra note 86.

\(^{181}\) Multiyear Scholarship Plan Moves on, supra note 21.


\(^{184}\) Id.

(scholarship) for two years and the other school then offers 85 percent for three years, etc.,” creates unfair advantages in the recruiting process. To understand the “bidding wars” or price competition that will occur as recruiters vie for prospective student-athletes one need only conduct simple arithmetic. Using the estimates from 2011, a school that offers a one-year renewable scholarship would only be guaranteeing a player approximately $42,224 toward his or her education. On the other hand, another university might offer a five-year scholarship, which includes an added $2,000 expense award per year, promising over $220,000 guaranteed toward that athlete’s education. Thus, the NCAA transformed athletic scholarships into a bargaining chip that demonstrates concretely that universities engage in price competition over student-athletes, and that provides further evidence of the ways in which recruiters already engage in such competition.

In addition, the new rule recognizes a labor market for student-athletes by highlighting the reality that “the value of the scholarship is based upon . . . the supply and demand for players.” The NCAA’s regulation on the amount of

---

187 See Clark, supra note 183.
188 See Grasgreen, supra note 90.
189 Grasgreen stated,

Athletes can receive additional scholarship funds of up to $2,000 or the full cost of attendance, whichever is less. Depending on the institution, the gap ranges from $200 to nearly $11,000 per year, and is the result of miscellaneous costs incurred on top of the tuition and fees, room and board, and books that full athletic scholarships currently cover. The $2,000 limit will be in place for at least three years, the board said, but in the future will be adjusted according to the consumer price index.

Id.

190 In Agnew, the court stated that

[C]olleges do, in fact, compete for student athletes, though the price they pay involves in-kind benefits as opposed to cash. For instance, colleges may compete to hire the coach that will be best able to launch players from the NCAA to the National Football League, an attractive component for a prospective college football player. Colleges also engage in veritable arms races to provide top-of-the-line training facilities which, in turn, are supposed to attract collegiate athletes. Many future student-athletes also look to the strength of a college’s academic programs in deciding where to attend. These are all part of the competitive market to attract student-athletes whose athletic labor can result in many benefits for a college, including economic gain.

Agnew v. NCAA, 683 F.3d 328, 347 (7th Cir. 2012). Moreover, universities have resorted to illegal forms of price competition, like offering cash or other incentives, which results in countless scandals.

191 Banks v. NCAA, 977 F.2d 1081, 1091 (7th Cir. 1992).
scholarships has varied over the years. Currently, the NCAA limits the amount of scholarships universities can offer in every sport. While the NCAA maintained the one-year scholarship rule, the demand for student-athletes at a given university remained restricted to the limited number of scholarships and the fixed value of one-year scholarships. Therefore, plaintiff student-athletes struggled to “allege that NCAA colleges purchase labor through the grant-in-aid athletic scholarships offered to college players when the value of the scholarship is based upon the school’s tuition and room and board.” Because schools can now award athletic scholarships of varying amounts, the value of a scholarship is no longer related to the expense of attending a university, but to the perceived athletic value of a student-athlete to that school. Furthermore, because teams that win are more profitable, if the new scholarship rule remains in place, schools will learn the best combination of differently valued athletes to create more successful teams. This will affect a school’s demand for a particular one-year or two-year or five-year guaranteed scholarship caliber athlete, and consequently an individual student-athlete’s contribution to a program will be valued accordingly and reflected in his scholarship.

The value of scholarships awarded each year will also depend on the supply of quality student-athletes graduating high school each year. Organizations like Max Preps, Rivals, ESPN, and others dedicate portions of their websites to recording statistics of high school athletes and to ranking them. These rankings assess the top overall recruits in the country and the best players by position, track a player’s scholarship offers and commitment, and grade universities on their eventual recruiting class. These analysts travel the country attending camps or combines—held by universities, independent organizations, and corporations such as Nike—where student-athletes preform drills and play games to put their talents on display. For these students, the goal is to

193 Banks, 977 F.2d at 1091.
194 The greater the athletic ability of a prospect the more scholarship money that prospect will likely be awarded.
196 Id.
make a “top-list” or to receive a five-star ranking, because this translates to multiple scholarship offers. Now that the NCAA member institutions can offer varying scholarships, the supply of “five-star” caliber student-athletes will affect the amount of multiyear scholarships offered each year.

Moreover, the new rule will likely increase the supply of student-athletes in the overall market. Each year a number of high-school student-athletes choose to go to schools that do not provide athletic scholarships. Often this is because these institutions are some of best academic institutions in the country, such as those in the Ivy League. But some recruits choose to forgo an athletic scholarship because they fear being unable to compete athletically or sustaining an injury. At a Division I school, this meant their scholarship might not be renewed. Now with the possibility of receiving an athletic scholarship that guarantees the full cost of an athlete’s education, these individuals might be persuaded to reenter the market.

2. Engaging in Contract Negotiations

Due to the athletic scholarship, “[c]ourts and scholars already overwhelmingly recognize the contractual nature of the relationship between student-athletes and their institutions.”

198 Bill Pennington, Financial Aid Changes Game as Ivy Sports Teams Flourish, N.Y. TIMES, Dec. 22, 2011, available at http://www.nytimes.com/2011/12/23/sports/financial-aid-changes-game-as-sports-teams-in-ivies-rise.html?pagewanted=all&_r=0. “We’re seeing a significant change in the caliber of the student-athlete,” said Steve Bilsky, the University of Pennsylvania’s athletic director, one of more than 50 Ivy League administrators and coaches interviewed. “It’s not even the same population because the pool has widened. We see a considerable number of student-athletes turning down athletic scholarships from places like Stanford, Northwestern or Duke to come to Penn.”

199 Recruits like Christian Webster, who chose to go to Harvard University instead of taking one of his twenty-five athletic-scholarship offers to play basketball. Id.

200 Instead, these students choose to pay to attend schools that they would consider better academic institutions. As Christian Webster explained, “It’s a sacrifice but it’s doable. . . . It’s not free, but it’s also not the full price of $50,000 or more. To me it was a 40-year life decision, not a four-year decision.” Id. (internal quotation marks omitted) (explaining his decision to attend Harvard University for $20,000 a year, instead of attending a university that offered him a one-year renewable scholarship).

201 “[S]tudent-athletes contemplating scholarship offers likely include economic factors in their decision-making process, such as the value of a given degree or the increased potential for entry into professional football.” Agnew v. NCAA, 683 F.3d 328, 341 (7th Cir. 2012).

202 Hakim, supra note 59, at 169.

The Letter of Intent, Statement of Financial Aid, and university bulletins and brochures provide the basis for the contractual relationship between the
Despite the incessant call from critics and reformers that the NCAA should acknowledge the employer-employee relationship between universities and pay their athletes—thereby ridding the NCAA of the student-athlete myth—\(^{203}\) the NCAA adamantly rejects the notion that an employment relationship exists and refuses to create one. President Mark Emmert stated he “can’t say often enough, obviously, that student-athletes are students; they are not employees . . . [and it] would be utterly unacceptable . . . to convert students into employees . . . . We don’t pay our student-athletes.”\(^{204}\) The NCAA’s position stems from the history of the athletic scholarship and its creation of a legal argument that a scholarship cannot be considered payment—and that student-athletes cannot be considered employees—so long as the scholarship does not exceed the cost of tuition and miscellaneous expenses.\(^{205}\) Although “the NCAA has crafted a body of case law that provides a position that student-athletes will remain simply student-athletes and will not obtain employee status,”\(^{206}\) in light of one of the most “tumultuous years in college sports”\(^{207}\) the NCAA’s position is vulnerable.

\(^{203}\) See generally McCormick & McCormick, \textit{supra} note 182; see also Stephen L. Ukeiley, \textit{No Salary No Union, No Collective Bargaining: Scholarship Athletes Are an Employer’s Dream Come True} 6 \textit{SETON HALL J. SPORT L.} 167, 177-78 (1996); Nocero, \textit{supra} note 79; Branch, \textit{supra} note 175. Among these critics are college athletes themselves, who formed the labor organization College Athletes Players Association (“CAPA”), and the Regional Director of the National Labor Relations Board, Region 13, Peter Sung Ohr, who ruled in favor of CAPA holding that scholarship football players at Northwestern University are employees under the National Labor Relations Act and may conduct an election to unionize. \textit{See generally Northwestern Univ. v. College Athletes Player Ass’n, N.L.R.B No. 13-RC-121359 (Mar. 26, 2014).}

\(^{204}\) Mark Emmert Interview, \textit{supra} note 10.

\(^{205}\) \textit{See supra} Part I.

\(^{206}\) Dennie, \textit{supra} note 12, at 46.

\(^{207}\) “[W]hich included conference realignment motivated by greed, several lawsuits that challenged the NCAA on antitrust grounds, and a massive scandal at Penn State that raised questions about the role of big-time college sports in university governance.” Sack, \textit{supra} note 22.
The adoption of a scholarship rule that creates dramatically different scholarship offers and negotiations between student-athletes and member institutions is incongruous with the NCAA’s unwavering stance that no employee-like relationship exists. With the reintroduction of the multiyear athletic scholarship, students can shop themselves to different universities to discover the price of their abilities indicated by the amount of guaranteed years offered in their scholarships—the potential contract salary for their labor. Even though players with lesser skill may still have their one-year scholarships renewed and eventually receive the same amount as the player guaranteed a five-year scholarship upfront, this process exacerbates the emphasis placed on athletic ability and performance, and eliminates any consideration of helping to subsidize a player’s education. This resembles the problem the NCAA faced when it first instituted athletic scholarships, and courts used the athletic scholarship as an indicator that the players qualified for worker’s compensation benefits. Consequently, athletic scholarships can now, more than ever, be paralleled to employment contracts, eroding the notion of amateurism before athletes even step foot on campus and enter the commercialized world of big-time college athletics.

The new scholarship policy contradicts the notion of amateurism not only because scholarship offers now resemble employment contracts, but also because it places academic goals at odds with athletic goals during the recruiting process. That student-athletes receive a valuable education is critical to the notion of amateurism. At the earliest stages of its organization, the NCAA posited an idealized notion of amateurism in college athletics, where an athlete focused primarily on something other than sports. The role of academics endures today as President Emmert denied the claim that student-athletes are employees on the basis that the NCAA “provide[s] [student-athletes] with remarkable opportunities to get an education at the finest universities on earth—that’s American universities and colleges.” But almost since its inception, the NCAA has struggled to maintain the illusion that student-athletes are

A representative from St. Francis College stated that as a result of the new scholarship policy, “prospective student athletes shop themselves around for the best deal in terms of length and compensation.” Complaint at 14-15, Rock v. NCAA, No. 12-CV-1019 (S.D. Ind. 2012).


See SACK & STAUVORSKY, supra note 29, at 33.

Mark Emmert Interview, supra note 10.
primarily students.\textsuperscript{212} In the commercialized NCAA that exists today, it is apparent that athletes in big-time sports play a limited role as students, exemplified by the amount of hours dedicated to their sport, the limitation on academic choices, the reduced standards to which they are subjected, the weak curricula they assume, and the low graduation rates they achieve.\textsuperscript{213} As many people believe, student-athletes are “just brought in to play some games. They don’t get a very good education, if they get one at all.”\textsuperscript{214}

Amid all this criticism, the NCAA’s new scholarship policy pits athletics against academics by asking high school students to choose guaranteed education over athletic glory. Schools will try to entice athletes with four or five-year scholarships that practically provide a guaranteed paid education. But these schools face opposition from more successful programs that play up the appeal of winning, becoming a professional, playing in bowl games, and learning from premiere coaches in premiere facilities (all football-related benefits). Although less than two percent of collegiate athletes make it to professional sports,\textsuperscript{215} this is obviously a huge attraction for many high school students, whose dream since childhood has been to become a professional athlete in their chosen sport.\textsuperscript{216} The situation will certainly arise in which a student chooses a one-year renewal offer from a school whose football team is consistently ranked in the top 25, instead of a school whose team is less successful, but guarantees four fully funded years of education. This places the decision on teenagers to choose between the guarantee of an academic degree and potential athletic fame.\textsuperscript{217}

\begin{itemize}
\item \textsuperscript{212} Sack & Staurowsky, supra note 29, at 33.
\item \textsuperscript{213} McCormick & McCormick, supra note 182, at 120; see also Branch, supra note 175; Nocera, supra note 79.
\item \textsuperscript{214} Ukeiley, supra note 203, at 209.
\item \textsuperscript{215} Except for baseball where 11.6\% of collegiate baseball players play professionally. Tom Manfred, Here are the Odds Your Kid Becomes a Professional Athlete, BUS. INSIDER (Feb. 10, 2012, 4:21 PM), http://www.businessinsider.com/odds-college-athletes-become-professionals-2012-2?op=1.
\item \textsuperscript{216} See id.
\item \textsuperscript{217} Although it may appear that the new scholarship policy provides students with a valuable bargaining chip, it fails to sufficiently protect those it will most likely affect, namely young black men who have been exploited by the recruiting process since the 1970s. As Gerald D. Higginbotham explains,
\end{itemize}

Young black children develop deep aspirations for sports because images of successful black figures in the media are usually limited to popular black entertainers and sportmen and in the black communities athletic achievement is rewarded more than any other activity. Sport, being focused more on physical and athletic ability rather than academic knowledge or social has lead to social mobility for immigrants and minority groups who faced discrimination other
B. Maintenance of Competitive Balance

The Supreme Court recognized maintenance of a competitive balance as a legitimate procompetitive justification for an NCAA bylaw under the rule of reason. The court assessed that in some instances regulations that fostered “equal competition will maximize consumer demand for the product.” When the NCAA adopted the one-year renewable scholarship in 1973, it supported the policy by claiming it would create competitive balance. The NCAA proposed that a uniform scholarship rule would even competition between universities offering only one-year scholarships and universities offering extended-term scholarships, thus preventing the recruiting process from disadvantaging any one university. The NCAA reaffirmed its stance in Agnew v. NCAA arguing “that multi-year scholarships would make it too difficult for less wealthy schools to compete in the recruiting market.” The NCAA, however, abandoned this position when it revised its bylaws to allow universities the discretion to offer single or multiyear scholarships.

The removal of the one-year scholarship rule marks a complete abandonment of the NCAA’s longstanding position that the rule was necessary to maintain a competitive balance. By no longer regulating the length of a scholarship offer, the NCAA explicitly states that a single and mandatory type of scholarship is not necessary to ensure that all universities have an equal opportunity to have successful programs. Unsurprisingly, member institutions met this policy change with resistance. These institutions argued that for decades they had operated under the belief, as the NCAA had purported, that a universal limit on scholarships prevented prospects from being wooed from their university by a school offering a more beneficial financial package. In opposition to the


219 The product of college sports. See *Bd. of Regents*, 468 U.S. at 119-20.
220 Hakim, *supra* note 59, at 158.
221 *See id.*
222 Agnew v. NCAA, 683 F.3d 328, 344 (7th Cir. 2012).
223 *See Hakim, supra* note 59, at 158.
new rule, some member institutions argue that more financially capable programs, whether through success or devotion of resources, are afforded an unfair advantage in the recruiting process because they will have the ability to offer more multiyear scholarships. As a result, they will be able to attract the most gifted athletes and create dominant teams, which will ultimately decrease the product of collegiate sports, because fans want to see exciting games and not one-sided games.

The Agnew court and critics expressed skepticism that the ban on multiyear scholarships will affect the overall product of collegiate football because the sport flourished prior to the institution of the ban. But this argument appears to be tenuous given the dramatic evolution of college football, specifically in the recruiting processes, over the last four decades. Absent any empirical evidence, it is difficult to quantify the effect the new rule will have on universities’ abilities to compete in the recruiting process. Yet, it is not farfetched to imagine that schools that devote more financial resources to their athletic programs may receive an advantage from having more multiyear scholarships at their disposal.

Additionally, the divergence from the justification of the one-year scholarship as a necessity for the maintenance of a competitive balance uncovers the ulterior motive behind the long-contested policy. Member institutions expressed outrage over the NCAA’s shift in scholarship policy because it hinders their ability to decline to renew a student-athlete’s scholarship. Over 62% of the member institutions voted to override the legislation. A representative of one institution stated that the new policy:

Create a recruiting disaster . . . institutions will be competing for recruits by making the best deal . . . in order to be competitive, institutions may offer multiyear awards so they can sign higher level recruits. However, there is never a guarantee that the incoming student-athlete will be a good fit for the program and the institution. If it is a poor fit the program is put in a difficult situation to continue to keep a student-athlete on scholarship.

Member institutions that oppose the rule believe that the burdens associated with preserving a scholarship for a player who lacks “athletic usefulness” inhibits their ability to produce a

---

224 See Levin, supra note 24.
225 See Agnew, 683 F.3d at 344.
226 Id.
227 See Hosick, supra note 86.
According to these member institutions, the new policy not only eliminates a necessary regulation to maintain a competitive balance, but also enacts a new policy that actually undermines their ability to compete.

IV. UNSOUND POLICY: FAILURE TO SUPPORT NCAA PRINCIPLES

The option to allow universities to offer student-athletes multiyear scholarships not only marks a retreat from the NCAA’s classical legal defenses, but it also fails to provide the reform needed to protect student-athletes and eliminate the harm that gives rise to antitrust claims.

Although it was important for the NCAA to reform its scholarship policy, its adopted policy suffers from a number of critical shortcomings. First, the new rule fails to address the problem created by one-year renewable scholarship offers, which caused the injury to the plaintiffs who brought forth the antitrust litigation. Universities are not required to offer multiyear scholarships, but rather have the option to do so. As a result, many student-athletes will continue to fear that their scholarship will not be renewed, and thus sacrifice their academics and risk playing with injury. Even an NCAA Presidential Taskforce concluded that, under the one-year scholarship policy: “[A]thletes may be legitimately concerned that their continued access to education depends on sports success. This can create a conflict of incentives that may lead to an emphasis on athletics at the cost of academics.” Moreover, the continued “fear of losing a scholarship and the economic hardship associated with expensive tuition incentivize injured student-athletes to resume playing before full recovery.” It is possible that the new scholarship policy might in fact magnify these problems for many student-athletes because coaches might not renew a one-year scholarship to make room for a multiyear

---

229 Id.
230 See, e.g., Agnew, 683 F.3d at 332-33; see also Complaint at 19-25, Rock, 928 F. Supp. 2d 1010 (S.D. Ind. 2013) (No. 12-CV-1019).
231 Levin, supra note 24.
232 See Branch, supra note 175 (At “informal football workouts at the University of Iowa just after the season-ending bowl games—workouts so grueling that 41 of the 56 amateur student-athletes collapsed, and 13 were hospitalized with rhabdomyolysis, a life-threatening kidney condition often caused by excessive exercise.”).
234 Id.
scholarship recruit. By ignoring the harm experienced by student-athletes, the NCAA fails to discourage student-athletes from bringing suits for compensatory damages.

Second, without representation from agents, student-athletes are ill-prepared to engage in the bargaining process for contracts that fully protect them. The NCAA recently removed the restrictions that limited college coaches’ ability to contact recruits, further aiding their already zealous recruiting behavior. Now, as early as the end of a prospect’s sophomore year “there will be no restrictions on phone calls, text messaging or contacting recruits via social media messengers.” Although this form of official contact is limited to the end of sophomore year, that does not prevent coaches from recruiting as early as middle school. This summer at their football camp, Louisiana State University’s football coaches offered a “soon-to-be eighth grader” a scholarship to be a member of the class of 2017. Student-athletes rely on these offers and the promises made by recruiters, but lack any means to guarantee that they will be fulfilled. Coaches are free to renege on their offers, “regardless of [whether a recruit] verbally committed or signed their National Letter of Intent.” This typically occurs as a result of oversigning, where “[s]chools often sign more players than they have available roster spots under the assumption that not all of the signees will qualify for the financial aid award.” When more recruits sign than there are spots available, “lesser regarded signees are told there is no room for them.” The addition of gradations in scholarship guarantees provides coaches with an additional incentive to entice and exploit student-athletes in an already


236 Miller, supra note 7, at 1155.


238 “Indeed, even if he accepted the offer, Dylan Moses couldn’t officially sign with LSU for another five years.” David Helman, LSU Courts Middle Schooler, ESPN (July, 26, 2012, 10:03 PM), http://espn.go.com/college-football/story/_id/8199497/soon-8th-grader-dylan-moses-offered-lsu-tigers-scholarship.


240 Id.

241 Id.
inequitable system. Without the assistance of counsel or an agent to help negotiate for a beneficial scholarship, the change to allow multiyear scholarship offers fails to “truly protect[] the student-athlete’s academic or athletic pursuits.”

Instead, the new rule benefits the NCAA by insulating the organization from further litigation and perpetuating its history of self-protective measures under the guise of reforms to protect student-athletes. The option to award multiyear scholarships only revises a bylaw that the NCAA feared courts might have found violated antitrust law. The new rule effectively removes the unreasonable restriction by allowing member institutions the freedom to award any scholarship that they choose. As a result, future plaintiffs will not be able to prove injury, as was claimed in the Agnew and Rock complaints, because the student-athletes cannot allege that they would have been awarded a multiyear scholarship but for the bylaw. This eliminates a cause of action for student-athletes to challenge a university’s unfair failure to renew a scholarship, and demonstrates that “scholarship is still an area where the NCAA . . . fail[s] in its mission to protect student-athletes.”

Ultimately, the NCAA’s “Prevent Defense” from this particular antitrust challenge is shortsighted because it willfully adopts policy that erodes its legal defenses and fails to rectify harm caused to the student-athletes it vows to protect. Such policy decisions uncover a long history of building and protecting a commercialized big business. The reaction of member institutions to the new policy substantiates the critique that the NCAA’s scholarship policy for the last four decades attempted to reduce costs, rather than maintain competition between universities. Since the inception of the multiyear scholarship ban, coaches and universities denied cancelling scholarships due to poor athletic performance or injury. But, as one critic questioned, “[i]f they were telling the truth, why did so many oppose this [multiyear scholarship option]?” The opposition by member institutions reveals that the primary purpose of the scholarship policy was to reduce costs associated with scholarships and enable coaches to “run-off” players they no longer wanted. These cost-cutting benefits came at the expense and exploitation of student-

---

242 Miller, supra note 7, at 1156.
243 Id.
244 Sack, supra note 22.
245 See Hakim, supra note 59, at 167.
The one-year scholarship fueled the practices that enable modern-day coaches to control every aspect of a student-athlete’s life from the time he arrives on campus to his graduation (if a player is among the minority to reach graduation). As result, student-athletes devote over 50 hours a week to their sport in season and offseason, play through injuries, miss class or give up certain majors, without a promise of continued education, leading to terribly low graduation rates for athletes who play revenue-generating sports.

The exposure of a significant regulatory area where the NCAA not only failed to protect student-athletes, but also facilitated their exploitation, uncovers the hypocrisy of the NCAA’s bylaws. This creates increased vulnerabilities for other bylaws, including those that have existed for long periods of time. For example, in In re NCAA Student-Athlete Name & Likeness Licensing Litigation, an ongoing antitrust case, a class made up of current and former NCAA student-athletes is challenging the NCAA restriction on allowing athletes to profit from the use of their names and likenesses. In an attempt to have the case dismissed, the NCAA argued the claims present “nothing more than cost-cutting by itself is not a valid procompetitive justification.” Yasser, supra note 28, at 1013 n.188.

This is because student-athletes fear that they will lose their scholarship, and that “their continued access to education depends on sports success.” Id.

See McCormick & McCormick, supra note 182; see also Luke DeCock, Football Graduation Gap Remains a Chasm, CHARLOTTE OBSERVER, Sept. 24, 2012, available at http://www.newsobserver.com/2012/09/24/2367364/decock-football-graduation-gap.html (“Three years after the University of North Carolina’s College Sport Research Institute started tracking graduation rates based not on raw numbers but on how athletes performed when compared to other students, nothing has changed. Football players are still graduating about 20 percent less than regular students. The latest edition of the study, planned for release Tuesday, found that FBS football players were 17 percent less likely to graduate than their male peers, down from 20 percent last year, with a three-year rolling average of 19 percent [graduation gap].”).

This vulnerability even extends to the viability of the scholarship itself. In two different district courts, California and New Jersey, classes of student-athlete plaintiffs have recently filed claims alleging that the athletic scholarship artificially caps collegiate athletes’ compensation to the cost of tuition, room, board, and books, and thus violate the Sherman Act. See Complaint, Alston v. NCAA, No. 3:14-cv-01011 (N.D. Cal. Mar. 4, 2014); see also Complaint, Jenkins v. NCAA, No. 3:33-av-00001 (D.N.J Mar. 17, 2014).

than a challenge to the NCAA’s rules on amateurism,” which the
NCAA posited were protected under Board of Regents. The court
rejected this argument stating that Board of Regents does not bar
the student-athletes antitrust claims and that the NCAA must
demonstrate that the ban “serves some procompetitive purpose.”
The District Court for the Northern District of California made
clear that to defeat the antitrust claims the NCAA would need to
rely on its traditional procompetitive justifications of amateurism
and maintenance of competitive balance. Because the court utilized
Agnew and Rock to deny the NCAA’s motion to dismiss, the
plaintiff class can use these cases to demonstrate that the NCAA
has recently weakened its traditional legal defenses, and therefore
those defenses are no longer sufficient to justify the restraint on the
market at issue in In re NCAA Student-Athlete Name & Likeness
Licensing Litigation. Although the NCAA and its “member schools
have downplayed the antitrust risks that stem from their current
mode of business[,]” it appears courts no longer accept that the
“NCAA plays a vital role in enabling college football to preserve
[the] character” of collegiate athletics, but rather treat the NCAA
like a comparable profit-maximizing business.

CONCLUSION: ADOPTION OF THE MANDATORY MULTIIYEAR DEAL

The NCAA’s decision to revive the multiyear scholarship by
allowing member institutions the option to provide scholarships for
more than one year is an unsound policy because it undermines the
NCAA’s traditional legal defenses to antitrust litigation and fails to
protect student-athletes. In reality, the NCAA could rectify the
problems of its newly adopted scholarship policy simply by

252 Id. at *6.
253 “In recent years, courts have held that NCAA rules restricting the size and availability of student-athletes’ scholarships and financial aid grants may be challenged under the Sherman Act, even though they relate to forms of student-athlete compensation.” Id. at *12 (citing Rock v. NCAA, 2013 WL 4479815, at *14 (S.D. Ind. 2013)). The court continued to refute the NCAA’s argument by stating,

Although the plaintiffs in Agnew focused on the NCAA’s scholarship rules, rather than its rules prohibiting student-athletes from licensing their publicity rights, the court’s rationale for distinguishing Board of Regents is still persuasive here: in short, Board of Regents did not address the impact of the NCAA’s horizontal restraints on student-athletes.

Id. at *6.
254 Edelman, supra note 250.
eliminating one-year renewable scholarships entirely and only allowing member institutions to offer multiyear scholarships.

By instituting a mandatory multiyear scholarship that provides the guaranteed cost of attendance for all student-athletes to graduate,256 the NCAA can revive its legal defenses and protect student-athletes. As Louis Hakim points out, “The formation of the extended-term scholarship contract will essentially follow the requisites of the scholarship agreement under the current system. The critical difference is that the parties will promise to be bound for four or five years rather than simply one year.”257

First, the mandatory multiyear policy would prevent coaches from cancelling scholarships of players who they no longer want or who suffered injuries,258 thereby eliminating the harm experienced by the plaintiffs in Agnew and Rock. To appease coaches and universities, players who quit without cause would become eligible to lose their scholarship because a student-athlete would have to remain eligible and willing to participate in athletics to maintain his or her athletic scholarship.259 Second, a mandatory four-year scholarship reinforces NCAA’s ideal of amateurism— “[s]tudent-[a]thletes shall be amateurs in an intercollegiate sport, and their participation should be motivated primarily by education and by the physical, mental[,] and social benefits to be derived”260—by shifting the emphasis to ensuring an education.261 Third, a rule that regulates evenly and limits universities to a singular type of scholarship option reinstates a competitive balance.

Critics of such a policy argue that this will decrease the quality of play on the field because coaches will be forced to carry players who they feel do not have the ability to make an impact. Among these critics are many college coaches.

256 This would most likely range from three to six years depending on whether a student-athlete graduated early or elected to use a redshirt and medical redshirt.

257 Hakim, supra note 59, at 170.

258 Id. at 167.

259 Id. at 165. The NCAA would need to create a provision or assign a committee to review these terminations to ensure that coaches were not running off undesirable players.


261 The mandatory scholarship rule must be an initial step to shift the paradigm in college athletics to emphasize the student-athlete’s education. Such a shift would require critical subsequent measures, including scaling back the number of allowable hours devoted to sport, eliminating the reduced academic standards for college athletes, and increasing the available academic support, in order to eradicate the low graduation rates and provide the meaningful education promised. See supra notes 213-14 and accompanying text.
themselves, who prefer one-year scholarships because it provides them overwhelming discretion over their teams and ultimately the fates of their student-athletes. These coaches also fear that guaranteed scholarships increase the costs associated with scholarships.²⁶² But,

[th]e average compensation for head football coaches at public universities, now more than $2 million, has grown 750 percent (adjusted for inflation) since the Regents decision in 1984; that’s more than 20 times the cumulative 32 percent raise for college professors. For top basketball coaches, annual contracts now exceed $4 million, augmented by assorted bonuses, endorsements, country-club memberships, the occasional private plane, and in some cases a negotiated percentage of ticket receipts.²⁶³

Coaches should not bemoan developing the players they recruited considering that is the job for which they receive such significant salaries. Moreover, if mandatory multiyear scholarships, which ensure that student-athletes graduate, do in fact increase costs for athletic programs, it justly reallocates the revenue to the individuals who generate it. Finally, any reduction of quality of play at the cost of “enhancing academic integrity and educational primacy in intercollegiate athletics”²⁶⁴ should be welcomed from organizations whose claimed principles are “educational values and academic integrity,” such as the NCAA and its member institutions.²⁶⁵

Vincent J. DiForté†

²⁶² Branch, supra note 175.
²⁶³ Id.
²⁶⁴ Hakim, supra note 59, at 168-69.
²⁶⁵ Id. at 164-65.
† J.D. Candidate, Brooklyn Law School, 2014; B.A. Amherst College, 2010. Thank you to the editors and staff of Brooklyn Law Review for their insights and guidance. I would also like to thank my family and Amanda Barrow for their love and support.
The JOBS Act and Crowdfunding

HOW NARROWING THE SECONDARY MARKET HANDICAPS FRAUD PLAINTIFFS

INTRODUCTION

Social networking has dominated the beginning of the twenty-first century. Its utility has extended beyond contacting friends to becoming a powerful marketing tool. Most recently, startup companies have begun to recognize its capability as a tool for acquiring funds by soliciting small donations from other internet users. This mechanism, by which funds for projects are procured by appealing to the internet community, is aptly called crowdfunding.

In recognition of its potential, crowdfunding has received further legitimation as a means for raising necessary capital from the federal government. Federal securities law has been amended to specifically legalize issuing equity-based securities through crowdfunding. In 2012, Congress passed the Jumpstart Our Business Startups Act (JOBS Act). The JOBS Act contains “game chang[ing]” provisions intended to allow small businesses, especially tech startups, to solicit investors. Among the provisions that have generated excitement from

7 Id.
investors and small businesses are those that create the new crowdfunding exemption. After certain disclosure requirements are met, small businesses may solicit investors through approved brokers or funding portals. The JOBS Act exempts crowdfunding issuers and investors from SEC reporting requirements so long as the transactions fall within statutory boundaries.

Nevertheless, while the JOBS Act creates new investment opportunities, it also creates the potential for investment fraud, posing unique challenges to investor fraud claims. Statutes such as the Private Securities Litigation Reform Act pose additional hurdles to those pursuing a fraud securities action. Furthermore, crowdfunding securities will occupy a novel space in common law jurisprudence on federal securities fraud litigation. Although the securities are generally available to the public, they cannot be resold within the first year of ownership. This restriction reduces liquidity of those securities, which diminishes the viability of a vibrant secondary market. Secondary markets have played an important role in the viability of currently established securities markets. A robust secondary market is a key factor in determining whether a market is efficient, that is, whether the information is reflected by the stock price. In turn, this makes it harder to establish that the stock price reflects information, in essence, making it hard to prove loss causation. If investors cannot show loss causation, then their fraud lawsuits will ultimately fail. Thus, without proper guidance by the Securities and Exchange Commission (SEC), plaintiff crowdfunding investors may be left with nothing but worthless stock.

This note addresses challenges to fraud litigation created by the JOBS Act. Part I provides background on the

---

10 Id. For a securities transaction to be exempted under § 77d, the “aggregate amount sold to all investors by the issuer . . . is not more than $1,000,000,” and “the aggregate amount sold to any investor by an issuer” cannot exceed “the greater of $2,000 or five percent of the annual income or net worth of such investor if either the annual income or the net worth of the investor is less than $100,000.” Id. § 77d(6)(B) (2012).
11 See infra Part III.
13 Infra Part III.
15 Infra Part III.
16 Infra Part III.
17 Infra Part III.
18 Infra Part III.
JOBS Act and introduces the crowdfunding provisions of the Act. Part II explores the statutory and common law background of fraud litigation. Part III examines litigation options available to defrauded investors in light of the JOBS Act’s crowdfunding provisions, as well as the undesirable impact of the Act’s one-year resale restriction on shareholder litigation, including loss causation. Part IV offers suggestions for dealing with this negative impact, more specifically, showing that, if the restriction on resale is loosened by allowing resale of shares to members already registered with complying funding portals, the loss causation problem discussed in Part III becomes less problematic.

I. BACKGROUND

The JOBS Act has garnered a lot of excitement from investors who impatiently await the SEC’s implementation of certain key provisions, including those relating to crowdfunding. But, the Act has been criticized for loosening regulations at the expense of the investor. This section will explore the historical background of the JOBS Act, and describe certain key provisions.

A. The Economic Environment That Gave Rise to the JOBS Act

Understanding the environment that gave rise to the JOBS Act helps explain the purpose behind its provisions. In 2008, the U.S. economy entered into the worst recession in recent history.
Additionally, the first decade of the twenty-first century saw several financial scandals. Many commentators attributed the economic crisis to deregulatory measures, such as the repeal of the Glass-Steagall Act. In response to the crisis, Congress passed the Dodd-Frank Act in 2010, which imposed additional regulatory requirements on securities issuers. Several years later, the SEC is still struggling to implement the Dodd-Frank Act.

In 2012, the economy remained stagnant. The economy’s anemic growth was attributed to investor weariness, which some argued caused investors to withhold capital. All the while, technological companies began lobbying Congress to ease compliance requirements to allow them to enter the public market through Initial Public Offerings. Small businesses, which had been hit hard by the recession, were viewed by some as the way out of the recession. America’s fixation with small businesses and its hope that those businesses would revive the economy were enough to create bipartisan support for the passage of JOBS Act.


26 For a detailed explanation on the delay of Dodd-Frank, see id., at 1019.


Crowdfunding in many ways symbolizes America’s hope in small businesses.\textsuperscript{32} Crowdfunding is a combination of crowdsourcing and microfinance, enabled by social networking.\textsuperscript{33} Crowdfunding websites solicit investments from ordinary people for projects.\textsuperscript{34} These projects usually involve artistic or gaming endeavors.\textsuperscript{35} Statistics also show that crowdfunding tends to engage younger people.\textsuperscript{36}

Prior to the JOBS Act, small businesses could not solicit investors through crowdfunding websites by promising equity in return for capital.\textsuperscript{37} This is because soliciting investors by promising equity or profits in exchange for buying shares from crowdfunding websites would have violated general solicitation provisions of federal securities laws.\textsuperscript{38} Congress’s ban on general solicitation in the Securities Exchange Act of 1933 was meant primarily to protect investors from fraudsters.\textsuperscript{39} Regulation D, Rule 502 bans non-exempt issuers—businesses that issue shares in exchange for equity—from advertising their offerings over several mediums\textsuperscript{40} Some have lamented the provisions that ban


\textsuperscript{33} See Bradford, supra note 4, at 27. Microfinancing “involves lending very small amounts of money” whereas crowdsourcing involves combining the efforts of numerous individuals to “achieve a goal.” Id.


\textsuperscript{35} For a more detailed explanation of the history of crowdfunding see Bradford, supra note 4, at 14-27.


\textsuperscript{37} See 17 C.F.R. § 230.502(c) (2011). “[A]n investment contract for the purposes of the Securities Act means a contract, transaction or scheme whereby a person invests his money in a common enterprise and is led to expect profits solely from the efforts of a promoter or a third party . . . .” SEC v. W.J. Howey Co., 328 U.S. 293, 298-99 (1946). For a more detailed analysis as to how crowdfunding can count as securities, see Joan MacLeod Heminway & Shelden Ryan Hoffman, \textit{Proceed at Your Own Peril: Crowdfunding and the Securities Act of 1933}, 78 TENN. L. REV. 879, 954 (2011).

\textsuperscript{38} 17 C.F.R. § 230.502(c) (“[N]either the issuer nor any person acting on its behalf shall offer or sell the securities by any form of general solicitation or general advertising.”).

\textsuperscript{39} See, e.g., Thomas Lee Hazen, \textit{Crowdfunding Or Fraudfunding? Social Networks And The Securities Laws—Why The Specially Tailored Exemption Must Be Conditioned On Meaningful Disclosure}, 90 N.C. L. REV. 1735, 1747-48 (2012) (“In large part as a response to these so-called ‘pump and dump’ schemes, the SEC amended Rule 504 to prohibit . . . general solicitation.”).

\textsuperscript{40} This includes “[a]ny advertisement, article, notice or other communication published in any newspaper, magazine, or similar media or broadcast over television or
general solicitation. Critics have cited the transaction costs of finding investors without generally soliciting them. Specifically, critics hope that with easier access to capital, small businesses can thrive.

Further, even if crowdfunding was not per se illegal, small businesses could not feasibly use it as a source of funding. This is because investments that promise equity—profits in exchange for capital—are securities under federal securities law. Thus, many small businesses would be required to register and report to the SEC and become a publicly traded company. The SEC’s compliance requirements are rigid and very expensive for the vast majority of small companies.

To address these issues, the JOBS Act contains provisions that legalize equity crowdfunding, subject to certain conditions. B. Crowdfunding Provisions

Under the JOBS Act, there are four requirements that must be met to exempt a securities transaction from reporting requirements and the Rule 502 ban on solicitation:

(A) the aggregate amount sold to all investors by the issuer, including any amount sold in reliance on the exemption provided under this paragraph during the 12-month period preceding the date of such transaction, is not more than $1,000,000;

(B) the aggregate amount sold to any investor by an issuer, including any amount sold in reliance on the exemption provided under this paragraph during the 12-month period preceding the date of such transaction, is not more than $100,000.

radio” and “any seminar or meeting” with attendees organized by such solicitation. 17 C.F.R. § 230.502(c) (2011).


The Jumpstart Our Business Startups Act § 302.
paragraph during the 12-month period preceding the date of such transaction, does not exceed—

(i) the greater of $2,000 or 5 percent of the annual income or net worth of such investor, as applicable, if either the annual income or the net worth of the investor is less than $100,000; and

(ii) 10 percent of the annual income or net worth of such investor, as applicable, not to exceed a maximum aggregate amount sold of $100,000, if either the annual income or net worth of the investor is equal to or more than $100,000;

(C) the transaction is conducted through a broker or funding portal that complies with the requirements of section 77d-1(a) of this title; and

(D) the issuer complies with the requirements of section 77d-1(b) of this title.  

Section 77d-1 places several requirements on issuers who offer securities through crowdfunding portals. An issuer must meet several disclosure requirements by disclosing its name, the names of its directors, “a description of the business of the issuer,” and its business plan. It must also go through various background checks depending on the size of its securities offering. With some exceptions, securities purchased through crowdfunding cannot be sold for a period of one year.

49 See id. § 77d-1(b).
50 Id. An issuer must:

file with the Commission and provide to investors and the relevant broker or funding portal, and make available to potential investors—(A) the name, legal status, physical address, and website address of the issuer; (B) the names of the directors and officers (and any persons occupying a similar status or performing a similar function), and each person holding more than 20 percent of the shares of the issuer; (C) a description of the business of the issuer and the anticipated business plan of the issuer; (D) a description of [the issuer’s] financial condition . . . ; (E) a description of the stated purpose and intended use of the proceeds of the [securities] offering; (F) the target offering amount, the deadline to reach the target offering amount, . . . ; (G) the price to the public of the securities or the method for determining the price . . . [and] a reasonable opportunity to rescind the commitment to purchase the securities; [and] (H) a description of the ownership and capital structure of the issuer . . . .

Id.

51 § 77d-1(b)(1)(D). Depending on the amount, the level of disclosure ranges from income taxes at a minimum to audited financial statements at a maximum. Id.
52 Id. § 77d-1(e)(1). Securities can be transferred to the “issuer of those securities,” “an accredited investor,” “as part of [a] [registered] offering,” and to family members. Id.
53 Id.
The JOBS Act calls for the creation of funding portals. These portals are where the crowdfunding securities will be transacted. Section 77d-1 places requirements on funding portals and issuers that must be met to capitalize on the exemptions. These portals must “ensure that each investor . . . reviews investor-education information, [and] . . . affirms that the investor understands that the investor is risking the loss of the entire investment, and that the investor could bear [the] loss.” The investor must answer questions that demonstrate an “understanding of the level of risk” associated with startup investments, illiquidity, and other investment matters. Regulation of these funding portals is delegated to the Financial Industry Regulatory Authority (FINRA).

C. The Ingredients for Fraud

Among the chief concerns of opponents of the JOBS Act is that its crowdfunding exemption creates a new means to defraud investors. There is merit to this concern. The JOBS Act exemptions bear some resemblance to the old Rule 504, which allowed “non-reporting issuers to offer and sell securities to an unlimited number of persons without regard to their sophistication or experience and without delivery of any specified information.” The old Rule 504 enabled widespread fraud and had to be amended by the SEC. Additionally, internet transactions may cater to impulse decision-making rather than careful deliberation over the financial security of investing in a particular project. These factors create fertile ground for fraud.

54 § 77d-1(a).
55 §§ 77d-1(a)-(b).
56 § 77d-1(a)(4).
57 Id.
59 See Hazen, supra note 39, at 1769.
61 See Karmel, supra note 21 (“Due to widespread fraud in the use of this exemption from registration, the SEC amended Rule 504 in 1999 by providing that securities issued under the rule are ‘restricted’ and prohibiting general solicitation and general advertising unless certain conditions are met.”).
1. The JOBS Act and its Similarity to the Old Rule 504

The exemptions created by the JOBS Act do in fact share several similarities to the old Rule 504. Like the new exemption in the JOBS Act, the old Rule 504 limited the aggregate offering price to $1 million in “[a] twelve month period.” The impetus behind the old Rule 504 is also similar to that behind the JOBS Act—that small businesses should have easier access to capital. These similarities have bolstered the argument that crowdfunding will enable fraud.

Nevertheless, the two rules are not identical. There are several notable differences between the old Rule 504 and the general solicitation exemption provided by the JOBS Act. First, as a general matter, the JOBS Act limits general solicitation to “accredited investors.” Second, it dictates the manner in which these companies could solicit capital by imposing several requirements on funding portals and crowdfunding sites. Third, it requires that the issuer “take reasonable steps to ensure” that investors are accredited. Finally, it imposes requirements to ensure that the pool of investors is educated. Thus, the Act takes steps to limit the class of persons who are being solicited, thereby limiting the number and kind of persons who may be defrauded. These steps help to ensure that the investor consults the crowdfunding portal, which holds pertinent information on the securities, before transacting in such securities. Yet, even with these differences, the fact remains that the JOBS Act, by virtue

---

64 Rule 504: Fact Sheet, supra note 60.
65 Id. (“[T]he limited offering exemption under Regulation D, is designed to help small businesses raise ‘seed capital.’”).
67 § 77d(a)(5).
68 § 77d-1(a).
69 § 77d “Modification of Exemption Rules.”
70 § 77d-1. The intermediary must:

ensure that each investor (A) reviews investor-education information, in accordance with standards established by the Commission, by rule; (B) positively affirms that the investor understands that the investor is risking the loss of the entire investment, and that the investor could bear such a loss; and (C) answers questions demonstrating—(i) an understanding of the level of risk generally applicable to investments in startups, emerging businesses, and small issuers; (ii) an understanding of the risk of illiquidity; and (iii) an understanding of such other matters as the Commission determines appropriate, by rule.

§ 77d-1(a)(4).
71 Compare §§ 77d & 77d-1, with Rule 504 Fact Sheet, supra note 60.
72 § 77d-1(a).
of allowing activity that was before illegal, will also create more opportunities for fraud.\textsuperscript{73}

2. “Pump and Dump” Schemes, Internet-User Impulsivity, and Unrealistic Investor Expectations

“Pump and dump” schemes are an avenue to swindle investors. Share prices are “pumped” by building excitement through exaggerated statements and financials, often through cold calls, e-mail solicitations, and other internet media.\textsuperscript{74} Once shares reach a high enough price, they are sold, or “dumped.”\textsuperscript{75} When the truth about the state of the company hits the market the shares become worthless leaving duped investors hanging.\textsuperscript{76} Importantly, the primary means of building excitement for these “pump and dump” schemes is the internet.\textsuperscript{77} Crowdfunding securities, which will be dealt primarily through the internet,\textsuperscript{78} may be vulnerable to such schemes.

The JOBS Act contains provisions to help curb these schemes, but they may not be enough to prevent “pump and dumps.” These funding portals need to follow statutory and FINRA requirements meant to deter widespread fraud.\textsuperscript{79}

\textsuperscript{73} See Rule 504: Fact Sheet, supra note 60.


\textsuperscript{75} See id.

\textsuperscript{76} In the My Baby Vintage scam, the share price went from “40 cents to $2.88.” Barrett, supra note 74. By the time of the dump, one dollar could buy five thousand shares. Id. In 2008, the SEC brought a suit against three defendants involved in the scheme. Id. Although the defendants purported to make almost $9 million in profit, id., ultimately the SEC could only obtain $2 million in assets to satisfy the judgment against the defendant. SEC v. Reynolds, 3:08-CV-438-B, 2011 U.S. Dist. LEXIS 26886, at *26 (N.D. Tex. Mar. 16, 2011).

\textsuperscript{77} A cursory look at CircleUp’s about page is not reassuring.

As part of the sign up process for CircleUp, each user must specify . . . which of the accredited investor requirements he/she meets. . . . When making an investment, each investor again represents and warrants in the signed purchase agreement that they are an accredited investor. Finally, prior to accepting any investment, [CircleUp] performs additional identity checks on each investor as required by law.

Frequently Asked Questions on Compliance, CIRCLEUP.COM (last visited Oct. 12, 2013), https://circleup.com/entrepreneur-education/compliance-faq/ (emphasis added). Furthermore, another risk to using the internet to invest is that there may be too much information. Heminway & Hoffman, supra note 37, at 934.

\textsuperscript{78} § 77d(a)(6)(C).

\textsuperscript{79} The Jumpstart Our Business Startups Act § 301 (“This title may be cited as the ‘Capital Raising Online While Deterring Fraud and Unethical Non-Disclosure Act of 2012’ . . . .”).
portals will ensure education of the investor remains to be seen. Yet, the current model seems to evince a lack of such education.\footnote{See, e.g., \textit{Beginner’s Guide to Investing: Online Publications at the SEC, SEC, \& EXCH. COMMN}, \url{http://www.sec.gov/investorpubs/begininvest.htm} (last visited Oct. 13, 2013).}

In keeping with norms of social media and internet usage,\footnote{See generally Panek, supra note 62.} crowdfunding websites cater to impulse decision-making.\footnote{Circleup.com is one such website. One internet source describes circleup.com as a “private social networking setting.” Colleen Taylor, \textit{Backed With $1.5M, CircleUp Aims To Be The AngelList For Consumer And Retail Startups}, \textit{TechCrunch} (Apr. 18, 2012) (emphasis added), \url{http://techcrunch.com/2012/04/18/circleup/}.} The internet has notably enabled and increased impulsive behavior, as demonstrated by the fact that most people tend to click through long online contracts without reading them or really considering their implications.\footnote{See generally Rebecca Smithers, \textit{Terms and Conditions: Not Reading the Small Print Can Mean Big Problems}, \textit{Guardian} (May 11, 2011), \url{http://www.guardian.co.uk/money/2011/may/11/terms-conditions-small-print-big-problems}.} If all an investor needs to do is click through some basic requirements,\footnote{See Frequently Asked Questions on Compliance, \textit{CircleUp.com}, \url{https://circleup.com/entrepreneur-education/compliance-faq/}.} then it is quite likely that some investors will skim or ignore lengthy disclosures, reacting to the internet-related impulse to accept most offers at face value.\footnote{See Smithers supra note 83.} Although this is not inherently bad in the context of donations and gratuitous investments, an investor looking for profit will care when the price of their stock falls.\footnote{As one JOBS Act’s proponent puts it, “people don’t want to believe that they’re wasting their time.” Spinrad, supra note 32.}

Fraudsters are already taking advantage of the JOBS Act. The SEC has initiated actions against penny stock companies that seek to capitalize on the hype of the JOBS Act.\footnote{See, e.g., Floyd Norris, \textit{Fraud Case Delayed By 2 Months}, \textit{N.Y. Times} (Nov. 1, 2012), \url{http://www.nytimes.com/2012/11/02/business/sec-charges-company-that-filed-under-jobs-act-with-fraud.html?pagewanted=all&_r=0/}.} For example, at the end of October, the SEC charged Caribbean Pacific Marketing with securities fraud.\footnote{See id.} Caribbean Pacific Marketing marketed itself online as an “emerging growth company,”\footnote{See id.} defrauding investors who were convinced that there was some value in the company. Caribbean Pacific Marketing sold its shares to insiders and then began a public offering that would allow insiders to recoup “over 100 times what [they] paid.”\footnote{Id.} This and similar occurrences seem to substantiate claims by those who regard rampant fraud in the crowdfunding market as inevitable.\footnote{See, e.g., Chris Gay, \textit{Equity Crowdfunding: Good for Capitalism or for Fraudsters?}, \textit{U.S. News} (Nov. 21, 2012), \url{http://money.usnews.com/money/personal-}}
After five years, half of all small businesses die. However, many crowdfunding investors will invest with the expectation (realistic or not) that the company invested in will become the next Facebook. So what are aggrieved investors to do if instead they become victims of a “pump and dump” scheme or are otherwise tricked into buying shares of a failing company? They sue.

II. CROWDFUNDING SECURITIES FRAUD CAUSES OF ACTION

Shares transacted through crowdfunding websites will fall within the scope of federal securities laws. Prior to the JOBS Act, a myriad of remedies existed for defrauded investors under these laws. Thus, investors will benefit from protections provided under both the JOBS Act as well as other federal securities laws. This Part will summarize available securities fraud causes of action, indicating courts’ treatment of similar liabilities to those provided for under the JOBS Act. This

---

finance/mutual-funds/articles/2012/11/21/will-crowdfunding-unleash-innovation-encourage-securities-fraud-or-both.

92 See, e.g., Scott A. Shane, Failure is a Constant in Entrepreneurship, N.Y. TIMES, http://boss.blogs.nytimes.com/2009/07/15/failure-is-a-constant-in-entrepreneurship/ (last updated July 17, 2009) (“According to U.S. Census data, only 48.8 percent of the new establishments started between 1977 and 2000 were alive at age five.”).

93 See, e.g., Tanya Prive, Inside the JOBS Act: Equity Crowdfunding, FORBES (Nov. 6, 2012) (“Business savvy individuals can now dream of being one of the first seed investors in the next Facebook . . . .”), http://www.forbes.com/sites/tanyaprive/2012/11/06/inside-the-jobs-act-equity-crowdfunding-2/.

94 See, e.g., id. Facebook’s IPO, enabled by the JOBS Act, has failed in the eyes of many investors, and Facebook is consequently facing a flurry of claims against it. “More than 40 lawsuits have been filed [against Facebook].” Id.; see also Securities Class Action Filings: 2010 Year in Review, CORNERSTONE RESEARCH 22 (2011), http://securities.stanford.edu/clearinghouse_research/2010_YIR/Cornerstone_Research_Filings_2010_YIR.pdf (“According to University of Florida Professor Jay Ritter’s dataset of IPOs, there were a total of 3,510 IPOs between January 1, 1996, and December 31, 2009. Out of these companies, 648 were defendants in at least one securities class action between 1996 and 2010, which corresponds to 18.5 percent of the sample of IPOs.”) (citations omitted).

The SEC may establish funds if it does manage to win in its own suit. See Investors Claims Funds, SEC (Oct. 5, 2000), http://www.sec.gov/answers/clmfund.htm. SEC actions, however, are generally slow and by the time they are initiated, if at all, the damage is substantial. See, e.g., Mark Williams, Why Did the SEC Fail to Spot the Madoff Case?, REUTERS (Jan. 6, 2009), http://blogs.reuters.com/great-debate/2009/01/06/why-did-the-sec-fail-to-spot-the-madoff-case/.


96 This note will focus on remedies that relate to material misstatements. However, it is important to briefly mention other forms of relief available to aggrieved investors. Generally, there are a variety of remedies under state law for various actions that are not covered or preempted by federal securities laws. See, e.g., Winer Family Trust v. Queen, 503 F.3d 319, 339 (3d. Cir. 2007) (affirming dismissal of state law causes of action brought on behalf of shareholder class); Gantler v. Stephens, 965 A.2d 695, 714 (Del. 2009) (refusing to allow plaintiff shareholders to proceed with their claims of breach of fiduciary duty and disclosure).
overview of the applicable laws provides necessary background for Part III which discusses how the JOBS Act affects this framework for securities fraud litigation.

It is noteworthy that, based on the mandated restrictions on individual amounts of investment per year under the JOBS Act, most crowdfunding plaintiffs will likely have little incentive to sue individually. This is because the JOBS Act mandates that securities transactions be relatively small in order to bypass reporting requirements.\(^\text{97}\) Thus, most crowdfunding plaintiffs will be reliant on the ability to file class action suits against their defrauders.

A. JOBS Act § 77d-1 versus Securities Act § 12(a)(2) Liability

One JOBS Act provision seeks to remedy a form of fraud or misstatement that arises from the crowdfunding portal itself. The JOBS Act creates liability that is analogous to § 12(a)(2) of the Securities Act of 1933.\(^\text{98}\) This liability is codified in 15 U.S.C.A. § 77d-1(c)(2):

An issuer shall be liable in an action under paragraph (1), if the issuer—
(A) by the use of any means or instruments of transportation or communication in interstate commerce or of the mails, by any means of any written or oral communication, in the offering or sale of a security in a transaction exempted by the provisions of section 77d (6) of this title, makes an untrue statement of a material fact or omits to state a material fact required to be stated or necessary in order to make the statements, in the light of the circumstances under which they were made, not misleading, provided that the purchaser did not know of such untruth or omission; and (B) does not sustain the burden of proof that such issuer did not know, and in the exercise of reasonable care could not have known, of such untruth or omission.\(^\text{99}\)

Courts characterize § 12(a)(2) as “impos[ing] essentially strict liability for material misstatements contained in registered securities offerings.”\(^\text{100}\) However, § 12(a)(2) liability extends only

\(^{97}\) See § 77d.

\(^{98}\) § 77d-1(c)(1)(B). (“An action brought under this [77d-1] shall be subject to the provisions of section 77l (b) of this title and section 77m of this title, as if the liability were created under section [12] of [the Securities Act].”)

\(^{99}\) § 77d-1(c)(2).

\(^{100}\) See, e.g., NECA-IBEW Health and Welfare Fund v. Goldman Sachs & Co., 693 F.3d 145, 148 (2d Cir. 2012). The Second Circuit court held that “[n]either scienter, reliance, nor loss causation are elements of § 12(a)(2) claims and thus “give[s] rise to liability more readily’ than § 10(b) [claims].’’ Id. at 154 (quoting In re Morgan Stanley Info. Fund Sec. Litig., 592 F.3d 347, 359-60 (2d Cir. 2010)). Although liability is easier with a § 12(a) violation, in reality, there will likely be less violations in virtue of the fact that § 12(a) liability applies to required disclosures. 15 U.S.C. § 77d-1 (c)(2)
to the issuer. In addition, defendants may raise defenses to § 12(a)(2) liability. Notably, one of these defenses, loss causation, also plays a pivotal role in § 11 liability cases.

B. Section 10 Liability

Aggrieved plaintiffs will have a cause of action under § 10 of the Securities and Exchange Act for violations of Rule 10b-5. Rule 10b-5 prohibits “the sale of any security” by means of “any manipulative or deceptive device or contrivance in contravention [of SEC Rules] or [if] appropriate in the public interest or for the protection of investors.” The Supreme Court has ruled that § 10 of the Securities and Exchange Act grants investors a private right of action connected to the fraudulent sale of securities in violation of Rule 10b-5. In order to win a § 10 claim, the plaintiff must meet these six elements:

1. a material misrepresentation (or omission), . . .
2. scienter, i.e., a wrongful state of mind, . . .
3. a connection with the purchase or sale of a security, [i.e., reliance]
4. reliance, often referred to in cases involving public securities markets (fraud-on-the-market cases) as “transaction causation,” . . .
5. economic loss, . . .
6. “loss causation,” i.e., a causal connection between the material misrepresentation and the loss, . . .

(emphasis added). Because the JOBS Act exempts many of these companies from reporting requirements, see 15 U.S.C. § 77d(a)(6), it follows that there will be less of a chance to trigger § 12(a)(2) liability.

See Gustafson v. Alloyd Co., Inc., 513 U.S. 561, 584 (1995) (“In sum, the word ‘prospectus’ is a term of art referring to a document that describes a public offering of securities by an issuer or controlling shareholder. The contract of sale, and its recitations, were not held out to the public and were not a prospectus as the term is used in the 1933 Act.” (emphasis added)).

See, e.g., In re Britannia Bulk Holdings, Inc. Sec. Litig., 665 F. Supp. 2d 404, 418 (S.D.N.Y. 2009) (“Defendants may assert the absence of loss causation as an affirmative defense to claims under Sections 11 and 12(a)(2) by proving that the allegedly misleading representations did not cause the depreciation in the stock’s value.” (internal citations omitted)). The defendant seller may also affirmatively establish that “in the exercise of reasonable care could not have known, of such untruth or omission’ which is ‘necessary in order to make the statements, in the light of the circumstances under which they were made, not misleading.” In re Worldcom, Inc. Sec. Litig. 346 F. Supp. 2d 628, 663 (S.D.N.Y. 2004) (quoting 15 U.S.C. § 77(a)(2)).


Dura, 544 U.S. at 341-42 (citations omitted).
1. *Janus* and the Authority Requirement

The Supreme Court has recently narrowed Rule 10b-5 liability. In *Janus Capital Group, Inc. v. First Derivative Traders*, the Court required that the “maker of a statement [be] the entity with authority over the content of the statement and whether and how to communicate it.”¹⁰⁸ This requires a plaintiff to prove that a defendant had ultimate authority over the fraudulent statement’s content and communication.¹⁰⁹ In today’s corporate landscape, corporations operate through various subsidiaries and limited liability companies. Thus, determining ultimate authority is increasingly difficult.¹¹⁰

2. Scienter and the Private Securities Litigation Reform Act

Securities traded through JOBS Act portals are “covered securities” subject to the Securities and Exchange Act of 1934.¹¹¹ Thus, fraud litigation with respect to crowdfunding securities will

---

¹⁰⁸ 131 S. Ct. 2296, 2303 (2011). In *Janus* the plaintiffs had sued Janus Capital Group, Inc. (JCG) and its mutual fund investment adviser, Janus Capital Management LLC (JCM), for material misstatements on a prospectus drawn up by JCM but issued by JCG’s investment fund. *Janus*, 131 S. Ct. at 2300. At issue was whether the plaintiffs could state a claim as to JCM with respect to whether JCM made the misstatement in violation of Rule 10b-5. *Id.* at 2301.

¹⁰⁹ *Id.* at 2303. Justice Thomas’ majority opinion tried to shed light on this meaning by comparing it to a presidential speech. *Id.* at 2302. Breyer, in his dissent, stated that the majority opinion would extend immunity to managers of a corporation who write a prospectus that the corporation itself has control over. *Id.* at 2310 (Breyer, J., dissenting). The majority, as the dissent points out, fails to address the level of control between JCM and JCG’s investment fund. *Id.* at 2312. JCM is the one that regularly managed JCG’s investment fund portfolio. JCM “furnish[ed] advice and recommendations concerning [investments and administrative compliance]” and JCM employees may have withheld information from JCG about the “market timing facts.” *Id.* at 2312. *Janus* may also have implications for crowdfunding. For one thing, the JOBS Act requires certain documents to be filed by the issuer with respect to financial health. 15 U.S.C. § 77d-1. The JOBS Act does not define who can audit a crowdfunder and what liability that auditor may incur. Section 10(b) liability might be the catch all, but under the *Janus* ruling, an auditor would likely not be liable since it is ultimately the crowdfunding issuer who has authority over the disclosure of the audit. See *Janus*, 131 S. Ct. at 2303; see also § 77d-1.

¹¹⁰ Some have pointed out that the *Janus* decision creates a loophole that corporations could exploit to avoid Rule 10b-5 liability. See, e.g., William A. Birdthistle, *The Supreme Court’s Theory of the Fund*, 37 J. CORP. L. 771, 786 (2012) (“If [a parent corporation] for example, created an external management firm, shifted all current [corporate] assets to a newly formed shell company, and then provided all executive management of the business via contract between those two entities, then could it not also limit its exposure to securities suits by citing *Janus*?”).

¹¹¹ 15 U.S.C. § 77r(b)(4) provides that “A security is a covered security with respect to a transaction that is exempt from registration under this subchapter pursuant to— (C) section 77d(a)(6) of this title[,]” Section 77d(a)(6) deals with JOBS Act securities transactions.
be subject to federal statutory and common law.\textsuperscript{112} One such statute is the Private Securities Litigation Reform Act.\textsuperscript{113} The Act heightens pleading requirements with respect to class actions.\textsuperscript{114} One heightened requirement relates to pleading scienter.\textsuperscript{115} Plaintiffs must allege facts that raise a “strong inference” that the issuer intended to defraud the investor.\textsuperscript{116} In \textit{Tellabs, Inc. v. Makor Issues & Rights, Ltd.}, the Supreme Court held that the Act requires a plaintiff to plead facts such that “a reasonable person would deem the inference of scienter . . . at least as compelling as any opposing inference one could draw from the facts alleged.”\textsuperscript{117}

3. \textit{Dura} and Loss Causation

In \textit{Dura Pharmaceuticals, Inc. v. Broudo}, the Supreme Court held that proving loss causation goes beyond simply showing that a misrepresentation affected the price of a stock.\textsuperscript{118} A plaintiff must go beyond alleging that “the price of the security on the date of the purchase was inflated because of the misrepresentation.”\textsuperscript{119} The Court found that, “as a matter of pure logic, at the moment the transaction takes place, the plaintiff has suffered no loss” since the inflated price is “offset by ownership of a share that at that instant

\begin{flushleft}
\textsuperscript{114} See, \textit{e.g.}, Tellabs Inc. v. Makor Issues & Rights, Ltd., 551 U.S. 308 (2007) (“As a check against abusive litigation by private parties, Congress enacted . . . [e]xacting pleading requirements . . . in the PSLRA.”).
\textsuperscript{115} See \textit{Ernst & Ernst v. Hochfelder}, 425 U.S. 185, 193 (1976). Justice Blackmun dissented to this requirement stating that “[i]f negligence is a violation factor when the SEC sues, it must be a violation factor” in a private action. \textit{Id.} at 217-18 (Blackmun, J., dissenting). The PLSRA also limits liability through its safe harbor provision. 15 U.S.C. \textsection 77z-2(c). Under this provision a defendant is not liable if “the forward-looking statement is identified as a forward looking statement and is accompanied by meaningful cautionary statements identifying important factors that could cause actual results to differ materially from those in the forward looking statement; or the plaintiff fails to prove that the forward looking statement was made with actual knowledge that the statement was false or misleading.” Plumbers and Pipefitters Local Union No. 630 Pension Annuity Trust Fund v. Allscripts Misys Healthcare Solutions, Inc., 778 F. Supp. 2d 858, 874 (N.D. Ill. 2011) (alteration in original) (emphasis added) (citing 15 U.S.C. \textsection 78u 5(c)(1)).
\textsuperscript{116} 15 U.S.C. \textsection 78u-4(b)(2).
\textsuperscript{117} \textit{Tellabs, Inc.}, 551 U.S. at 324.
\textsuperscript{119} \textit{Id.} at 338 (quotations omitted).
\end{flushleft}
possesses equivalent value.” The misrepresentation must cause, not merely “touch upon a loss.” Thus, if bad market conditions affect the price of shares, the plaintiff will ultimately be unable to make a showing of loss causation.

4. Materiality and Reliance

Reliance, that is, showing a connection between a material representation or omission and the sale or purchase, has been a complicated issue for courts to grapple with. In addition to showing that a representation was material, the misrepresentation must have induced the transaction. “Fraud-on-the-market” gives class action plaintiffs some leeway in arguing reliance. In a securities fraud class action that involves publicly traded securities, courts may grant the plaintiffs a rebuttable presumption of reliance. For a plaintiff to raise this presumption, he or she must show that the securities were traded in an efficient market.

Courts differ on what facts a plaintiff must show to trigger the presumption. The majority approach requires that the plaintiff show that “the market price of the stock fully reflects all publicly available information.” The court has defined “fully reflect,” to “mean [when] market price responds so quickly to new information that ordinary investors cannot

120 Id. at 342.
121 Id. at 343.
122 See, e.g., Phillips v. Scientific Atlanta, Inc., 489 F. App’x 339, 340-41 (11th Cir. 2012) (affirming summary judgment for defendants since plaintiffs failed to properly allege loss causation where they failed “to disentangle the effect of new information regarding customer inventory levels from [Scientific-Atlanta’s misrepresentation]”).
123 A misrepresentation or omission is material if it is “so obviously important to an investor that reasonable minds cannot differ on the question of materiality.” TSC Indus., Inc. v. Northway, Inc., 426 U.S. 438, 450 (1976) (quoting Johns Hopkins Univ. v. Hutton, 422 F.2d 1124, 1129 (4th Cir. 1970)).
124 See Basic Inc. v. Levinson, 485 U.S. 224, 243 (1988) (“[R]eliance is an element of a Rule 10b-5 cause of action. Reliance provides the requisite causal connection between a defendant’s misrepresentation and a plaintiff’s injury.” (citation omitted)).
125 See id. at 242 (“Requiring proof of individualized reliance from each member of the proposed plaintiff class effectively would have prevented respondents from proceeding with a class action, since individual issues would have then overwhelmed the common ones.”).
126 See id. at 247. The presumption can be rebutted if the defendant is able to prove that absent the misrepresentation, the plaintiff would have transacted the securities anyway. Id. at 248 (“[I]f, despite petitioners’ allegedly fraudulent attempt to manipulate market price, news of the merger discussions credibly entered the market and dissipated the effects of the misstatements, those who traded Basic shares after [denial of merger discussions] would have no direct or indirect connection with the fraud.”).
127 Id. at 248 n.27.
128 In re PolyMedica Corp. Sec. Litig., 432 F.3d 1, 19 (1st Cir. 2005).
make trading profits on the basis of such information.” The courts look to the following factors, the first five of which are known as the Cammer factors, to make this determination:

1. the average weekly trading volume expressed as a percentage of total outstanding shares;
2. the number of securities analysts following and reporting on the stock;
3. the extent to which market makers and arbitrageurs trade in the stock;
4. the company’s eligibility to file SEC registration Form S-3 (as opposed to Form S-1 or S-2);
5. the existence of empirical facts “showing a cause and effect relationship between unexpected corporate events or financial releases and an immediate response in the stock price”;
6. the company’s market capitalization;
7. the bid-ask spread for stock sales; and
8. float, the stock’s trading volume without counting insider-owned stock.

The court must review market efficiency with respect to the contested shares, not general market efficiency. This usually involves a battle of the experts. The holder of this presumption gains an important tool in shareholder litigation.

5. Class Action Requirements

The shareholder class action is the primary way for shareholders to litigate fraud. To certify a class action, Rule...
23 of the Federal Rules of Civil Procedure requires that the class be numerous, have common “questions of law or fact,” possess typical claims or defenses, and have its interests “fairly and adequately” represented.\textsuperscript{136} In addition, Rule 23(b)(3) requires that “there be questions of law or fact common to the members of the class [that] predominate over any questions affecting only individual members, and that a class action is superior to other available methods for the fair and efficient adjudication of the controversy.”\textsuperscript{137} Thus, Rule 23(a) requires common questions of law or fact for § 10(b) class claims.\textsuperscript{138}

Rule 23(b) requires that a court ensure that the class action provide common answers to the plaintiffs’ common questions.\textsuperscript{139} With respect to commonality, classes are certified based on a “fraud-on-the-market theory.”\textsuperscript{140} In \textit{Amgen Inc. v. Connecticut Retirement Plans and Trust Funds}, the Supreme Court clarified that proof of materiality is not appropriate during the class certification stage.\textsuperscript{141} Thus, courts will only require a prima facie showing of reliance to trigger the fraud-on-the-market presumption and save materiality for after certification.\textsuperscript{142}

III. HOW THE JOBS ACT CHANGES THIS LANDSCAPE

A. The One-Year Restriction on Resale Hurts Investors

By opening themselves to the public, corporations are able to raise more capital for projects in hopes that they experience any plan which is not just and reasonable.” (citations omitted) (internal quotation marks omitted)).\textsuperscript{135} The reason this note has not mentioned state law class actions is because, with the exception of derivative suits, fraud class actions will be unavailable to plaintiffs. This is because, under the statute, crowdfunding shares are “covered securities.” 15 U.S.C. § 77r(b)(4) (2012) (covered securities include exempted transactions under § 77d). Consequently, such class actions are precluded by SLUSA. See \textit{Kircher v. Putnam Funds Trust}, 547 U.S. 633, 636-37 (2006).

\textsuperscript{136} \textit{See Fed. R. Civ. P. 23(a).}

\textsuperscript{137} \textit{Fed. R. Civ. P. 23(b)(3).}


\textsuperscript{139} Walmart Stores, Inc. v. Dukes, 131 S. Ct. 2541, 2551 (2011) (“What matters to class certification . . . is not the raising of common ‘questions’—even in droves—but, rather the capacity of a classwide proceeding to generate common answers apt to drive the resolution of the litigation. Dissimilarities within the proposed class are what have the potential to impede the generation of common answers.” (citations omitted)).

\textsuperscript{140} \textit{See Amgen Inc. v. Conn. Ret. Plans and Trust Funds}, 133 S. Ct. 1184, 1195 (2013) (“[W]ithout the fraud-on-the-market theory, the element of reliance cannot be proved on a classwide basis through evidence common to the class.”).

\textsuperscript{141} 133 S. Ct. at 1204.

\textsuperscript{142} \textit{See id.} at 1203-04.
major growth.\textsuperscript{143} In return, investors expect returns on the shares, in part because they can sell them whenever they want.\textsuperscript{144} On the other hand, some corporations are closed corporations.\textsuperscript{145} Closed corporations are businesses that use the corporation as a legal scheme to avoid liability that they may incur as a partnership.\textsuperscript{146} They may elect to put restrictions on these shares and limit transferability through various statutory tools.\textsuperscript{147} The primary reason is so that control of a closed corporation stays within the family or a limited group of persons.\textsuperscript{148} Thus, closed corporations sacrifice a ready market for their shares, and thus value of their shares,\textsuperscript{149} in exchange for control.

Crowdfunding securities occupy an awkward crevice between public and private markets. These securities, by virtue of the fact that they are available on online funding portals,\textsuperscript{150} possess a public quality. Yet, the only companies that can avail themselves of capital bear resemblance to closed corporations by virtue of their small size. This awkwardness is reflected in some of the JOBS Act provisions. Specifically, the JOBS Act contains a provision that places a restriction on transferability of crowdfunding shares.\textsuperscript{151} Crowdfunding shares:

\begin{quote}
may not be transferred by the purchaser of such securities during the 1-year period beginning on the date of purchase, unless such securities are transferred—(A) to the issuer of the securities; (B) to
\end{quote}

\textsuperscript{143} Frequently Asked Questions: What are the advantages and disadvantages for a company going public?, INVESTOPEDIA, http://www.investopedia.com/ask/answers/09/ipoadvantagedisadvantage.asp#ixzz2BrV6VoOy (last visited Jan. 18, 2012). Other benefits of going public include publicity, increased market share, and the possibility of an exit strategy for founders of a successful small business. Id.

\textsuperscript{144} See Knowledge@Wharton, Are Public Corporations Passé?, TIME (Oct. 12, 2012), http://business.time.com/2012/10/12/a-premature-eulogy-for-public-companies/.


\textsuperscript{146} Compare DEL. CODE ANN. tit. 6, § 15-801 (West 2013) [Uniform Partnership Act] (expressing that a partnership is dissolved when a partner dissociates or leaves the partnership), with DEL. CODE ANN. tit. 8, § 275 [Delaware General Corporate Law] (stating that unless the requisite voting takes place or a court intervenes, corporations do not need to wind up their business because a shareholder decides to leave).

\textsuperscript{147} See, e.g., F.B.I. Farms, Inc. v. Moore, 798 N.E.2d 440, 445 (Ind. 2003) (“Indiana, like virtually all jurisdictions, allows corporations and their shareholders to impose restrictions on transfers of shares.”).

\textsuperscript{148} See, e.g., Galler v. Galler, 203 N.E.2d 577, 583 (Ill. 1964) (“For our purposes, a close corporation is one in which the stock is held in a few hands, or in a few families, and wherein it is not at all, or only rarely, dealt in by buying or selling.”).

\textsuperscript{149} E.g., Thomas J. Andre, Jr., Restrictions on the Transfer of Shares: A Search for Public Policy, 53 TUL. L. REV. 776, 785 (1979) (“That transfer restrictions may have a negative impact on the value a shareholder will receive for his shares may be conceded; the value of shares without transfer restrictions will often be greater than the same shares with restrictions.”).


\textsuperscript{151} § 77d-1.
an accredited investor; (C) as part of an offering registered with the Commission; or (D) to a member of the family of the purchaser or the equivalent, or in connection with the death or divorce of the purchaser or other similar circumstance, in the discretion of the Commission; and

(2) shall be subject to such other limitations as the Commission shall, by rule, establish.\(^{152}\)

Some contend that the one-year resale restriction serves an important purpose, arguing that the restriction acts as a deterrent against fraud.\(^{153}\) Specifically, the restriction on resale would curb the rise of a fraudulent secondary market on these crowdfunding securities.\(^{154}\) The argument is that, in these secondary markets, the buyers “find [themselves] attenuated from an accurate and complete source of information.”\(^{155}\) But the provision would dissuade many institutional investors from purchasing these shares, because restrictions on re-sale reduce liquidity of the shares, curbing a major source of investment.\(^{156}\) Thus, while the one-year resale restriction tenuously serves an anti-fraud purpose, it ultimately harms investors more than it helps them. The next Section explores the implications of the one-year resale provision on litigating fraud.

**B. The Resale Restriction Severely Impacts the Market for Crowdfunding Securities**

A cursory look at crowdfunding securities may lead one to believe that the fraud-on-the-market presumption may be available to crowdfunding plaintiffs. In *Basic*, the Court found

\(^{152}\) § 77d-1(e). It is interesting to note that the SEC is tasked with adding other rules to this restriction in transferability, § 77d-1(e)(2). One rule that could be considered is to allow companies to impose further limitations on transferability. This is not to protect the “control” of the business, but because the statute requires that an issuer who reaches two thousand shareholders become a reporting company. 15 U.S.C. § 78l(g)(1)(A)(i).

\(^{153}\) See Bradford, *supra* note 4, at 144 (“Heminway and Hoffman argue that such restrictions are necessary because a resale market may not provide new investors with direct access to the information available on the crowdfunding site itself, so resales are more conducive to fraud.”).

\(^{154}\) See Heminway & Hoffman, *supra* note 37, at 954.

\(^{155}\) Id.

\(^{156}\) There are other reasons why crowdfunding may be unattractive to institutional investors. Venture capitalists, investors who invest in and manage start ups, may be dissuaded from investing in firms that are owned by a multitude of inexperienced investors. Rohit Arora, *7 Reasons to Avoid Crowdfunding*, FOX BUSINESS (Oct. 23, 2012)http://smallbusiness.foxbusiness.com/finance-accounting/2012/10/23/7-reasons-to-avoid-crowdfunding/ ; see also Venture Capital definition. INVESTOPEDIA, http://www.investopedia.com/terms/v/venturecapital.asp#axzz2BwwoEatU (last visited Nov. 9, 2012).
that the lack of “face-to-face transactions” justified the “fraud-on-the-market” presumption.\textsuperscript{157} Importantly, the Court found that, for the question of reliance, the question in modern securities is not whether the information misrepresented by the fraudulent actor induced the transaction.\textsuperscript{158} This is because the investor is not evaluating this information; the market is evaluating this information.\textsuperscript{159}

The absence of a face-to-face relationship would also exist between crowdfunding issuers and investors. First, the crowdfunding issuer is not selling directly to investors.\textsuperscript{160} The JOBS Act itself requires a third person, the funding portal intermediary, to be involved in this transaction.\textsuperscript{161} Although the issuer is required to provide the intermediary with information, the issuer must also provide that information to investors.\textsuperscript{162} But a defrauded investor may also argue that he or she relied on the funding portal because the statute requires the intermediary “take such measures to reduce the risk of fraud.”\textsuperscript{163} This clause does not bestow ownership over the issuer’s fraudulent misrepresentations. It only requires that the intermediary follows SEC rules that reduce the risk of fraud.\textsuperscript{164} The role of an intermediary is similar to that of a broker; it is merely a third party that facilitates the transaction between the issuer and the investor.\textsuperscript{165} Thus, intermediaries do not by themselves create a secondary market.

\begin{footnotesize}
\begin{enumerate}
\item Basic Inc. v. Levinson, 485 U.S. 224, 243-44 (1988).
\item \textit{Id.}
\item See \textit{id.} at 244 (quoting \textit{In re LTV Sec. Litig.}, 88 F.R.D. 134, 143 (N.D. Tex. 1980) (“The market is acting as the unpaid agent of the investor, informing him that given all the information available to it, the value of the stock is worth the market price.” (internal quotation marks omitted)).
\item The statute interposes an intermediary funding portal for crowdfunding transactions. 15 U.S.C. § 77d(a)(6).
\item § 77d(a)(6)(C). Furthermore, the statute says that the issuer cannot advertise their offering, but only point the potential investor to the funding portal. 15 U.S.C. § 77d-1(b)(2).
\item § 77d-1(b)(1).
\item § 77d-1(a)(5).
\item \textit{Id.} Even if this statute imposed some criminal liability, after \textit{Janus} it is unlikely that the court will find this third party liable under Rule 10b-5. See Janus Capital Group, Inc. v. First Derivative Traders, 131 S. Ct. 2296, 2303-04 (2011).
\item 15 U.S.C. § 78c(a)(60) (“The term ‘funding portal’ means any person acting as an intermediary in a transaction involving the offer or sale of securities for the account of others, solely pursuant to section 4(6) [1] of the Securities Act of 1933 (15 U.S.C. 77d (6)), that does not—(A) offer investment advice or recommendations; (B) solicit purchases, sales, or offers to buy the securities offered or displayed on its website or portal; (C) compensate employees, agents, or other persons for such solicitation or based on the sale of securities displayed or referenced on its website or portal; (D) hold, manage, possess, or otherwise handle investor funds or securities; or (E) engage in such other activities as the Commission, by rule, determines
\end{enumerate}
\end{footnotesize}
Even if a secondary market were created, this market would not be open enough to grant plaintiffs the reliance presumption.\textsuperscript{166} Under \textit{Basic}, for the fraud-on-the-market presumption to apply, the markets must be “open and developed.”\textsuperscript{167} Cursorily, crowdfunding may seem like the most “open” market. So long as the requirements are met, almost anyone can become a crowdfunding investor.\textsuperscript{168} However, openness is not a matter of who can enter the market. Instead, openness refers to the market’s reaction to information.\textsuperscript{169}

A court looking at the five \textit{Cammer} factors would not apply the “fraud-on-the-market” presumption.\textsuperscript{170} First, it is unlikely that there will be a significant following by analysts of crowdfunding securities.\textsuperscript{171} Second, the one-year restriction makes it unlikely that crowdfunding stocks will be sold at a sizable volume on a weekly basis. Third, because institutional investors, who are typically market makers,\textsuperscript{172} are skeptical of crowdfunding,\textsuperscript{173} it is unlikely that there would be many market makers in the field of crowdfunding securities. Fourth, even if the required disclosures for crowdfunders on an offering were analogized to an S-3, the difference between the information provided between the two is vast.\textsuperscript{174} Finally, as a result of the four previous factors, it is unlikely that

\textsuperscript{166} See \textit{Basic Inc. v. Levinson}, 485 U.S. 224, 250 (1988). The JOBS Act allows some sales, but only to accredited investors and family. 15 U.S.C. § 77d-1(e). Thus, some resale market may exist, but it would be relatively small.

\textsuperscript{167} \textit{Basic}, 485 U.S. at 241 (quoting \textit{Peil v. Speiser}, 806 F.2d 1154, 1160-61 (3rd Cir. 1986)).

\textsuperscript{168} See § 77d(a)(6). So long as the investor goes through required education materials, they only need to conform to the investment limits. \textit{Id.}

\textsuperscript{169} See, \textit{e.g.}, \textit{In re PolyMedica Corp. Sec. Litig.}, 432 F.3d 1, 19 (1st Cir. 2005).


\textsuperscript{171} See Yoanca Ertimur, Volkan Mushu, & Frank Zhang, \textit{Conflicts of Interest or Selection Bias? Evidence from Analysts’}, REV. OF ACCT. STUD. 6 (forthcoming), \textit{available at} https://faculty.fuqua.duke.edu/~yertimur/bio/EMZ_April2009.pdf (“[T]he selection bias explanation posits that analysts follow companies for which they truly have favorable views, giving rise to a higher proportion of favorable recommendations.”).

\textsuperscript{172} A market maker is:

a dealer who, with respect to a particular security, (i) regularly publishes bona fide, competitive bid and offer quotations in a recognized interdealer quotation system; or (ii) furnishes bona fide competitive bid and offer quotations on request; and, (iii) is ready, willing and able to effect transactions in reasonable quantities at his quoted prices with other brokers or dealers.

information in press releases would immediately affect stock price. Further, the JOBS Act requires only that crowdfunding issuers comply with the documents and information needed for the initial offering. Thus, it is very likely that information will not be fully reflected in the stock price to warrant a fraud-on-the-market presumption.

Even if courts decide to grant a different presumption of reliance, there can be no claim of § 10(b) liability unless all elements are established. Importantly, the resale restriction presents a new challenge to litigants who are trying to prove loss causation. The importance of loss causation is further boosted by the fact that loss causation also speaks to § 12(a)(2) liability, and thus the analogous liability created by the JOBS Act.

C. The Resale Provision Significantly Hinders Proving Loss Causation


A typical loss causation fact pattern should be examined to better understand the impact of intervening causes. In 2001, investor plaintiffs bought stock from defendant, Scientific-Atlanta, Inc. (SA). On January 18, 2001, SA issued a press release stating that it had “record financial results” along with another “press release announcing an increase in manufacturing capacity” in response to consumer demands. SA, through its various agents, contended to make such statements to various financial media. Six months later, in July, SA announced a

---

175 Inefficient markets generally do not respond as predictably to information as efficient ones. See, e.g., Krogman v. Sterritt, 202 F.R.D. 467, 473, 478 (N.D. Tex. 2001).

176 See § 77d-1.

177 See, e.g., In re PolyMedica Corp. Sec. Litig., 432 F.3d 1, 19 (1st Cir. 2005).

178 Some courts adopt another theory of reliance dubbed “fraud-created-the-market.” See, e.g., Regents of Univ. of Cal. v. Credit Suisse First Boston (USA), Inc., 482 F.3d 372, 391-92 (5th Cir. 2005). This theory “assumes [that] investors relied on the market itself to prevent the entry of ‘unmarketable’ securities [to the market]”. Michael J. Kaufman & John M. Wunderlich, Fraud Created the Market, 63 A.L.A. L. REV. 275, 281 (2012). But this theory only applies to securities transacted between the issuer and the investor. See id. at 281-82.


180 See id.; see also In re Britannia Bulk Holdings, Inc. Sec. Litig., 665 F. Supp. 2d 404, 418 (S.D.N.Y. 2009); cf. §§ 77d-1(e)(1)–(2).


182 Id. at 340-41.

183 Id.
sales decrease.\textsuperscript{184} SA attributed this decrease to economic uncertainty and a decline in new orders of its cable equipment.\textsuperscript{185} SA then filed its form 10K with the SEC for the year of 2001.\textsuperscript{186} SA’s Chief Financial Officer stated that it anticipated some damage for the next fiscal year, attributable to the economic decline.\textsuperscript{187} SA’s stock price dropped dramatically.\textsuperscript{188} Plaintiffs then brought suit for violations of Rule 10b-5.\textsuperscript{189}

The district court granted the defendant’s motion for summary judgment, finding that the plaintiffs had “presented genuine issues of material fact on all the required elements of their claim, except for loss causation.”\textsuperscript{190} The Eleventh Circuit affirmed the district court’s ruling finding that the plaintiffs failed to adequately plead loss causation.\textsuperscript{191} The \textit{Phillips} court found that the press releases contained pieces of non-fraudulent information that could explain the loss.\textsuperscript{192} The court cited “uncertain economic climate, reduced marketing by SA customers, [and] unexpectedly slow deployment of interactive cable services” as other potential causal factors.\textsuperscript{193} The plaintiffs used expert testimony to attempt to single out the fraudulent press releases from all the other explanatory factors.\textsuperscript{194} However, they were unable to separate the non-fraudulent statements’ effects (e.g., economic downturn) that affected SA specifically.\textsuperscript{195} Thus, there was no basis to determine loss causation in connection with the fraudulent statements.\textsuperscript{196}


Crowdfunding is risky. Establishing loss causation for a small business using the aforementioned analysis will be difficult. Small businesses are generally known to fail.\textsuperscript{197} A bad

\textsuperscript{184} \textit{Id.}
\textsuperscript{185} \textit{Id.} at 341-42.
\textsuperscript{186} \textit{Id.}
\textsuperscript{187} \textit{Id.}
\textsuperscript{188} \textit{Id.}
\textsuperscript{189} \textit{Id.}
\textsuperscript{190} \textit{Id.} (emphasis added).
\textsuperscript{191} \textit{Id.} at 342.
\textsuperscript{192} \textit{Id.}
\textsuperscript{193} \textit{Id.} at 342.
\textsuperscript{194} \textit{Id.} at 342-43.
\textsuperscript{195} \textit{Id.} at 343.
\textsuperscript{196} \textit{Id.} at 343.
\textsuperscript{197} See, \textit{e.g.}, Scott A. Shane, \textit{Failure is a Constant in Entrepreneurship}, N.Y. TIMES, http://boss.blogs.nytimes.com/2009/07/15/failure-is-a-constant-in-entrepreneurship/ (last updated July 17, 2009)(“According to U.S. Census data, only 48.8 percent of the new establishments started between 1977 and 2000 were alive at age five.”).
economy can certainly affect these small businesses individually. Further, there are other factors that might diminish a plaintiff’s loss causation argument. For example, even if the economy at large is healthy, there are local conditions that may affect a small business. Phillips involved securities traded on a public market, yet the defrauded plaintiffs were still unable to show loss causation. Crowdfunding securities are riskier and less liquid.

The resale restriction allows for other mitigating factors to come into play. In fact, it is questionable whether the plaintiff can even be said to have suffered any damage until after the one-year period is over. Evaluating loss causation requires determining “how much . . . the price [would] have been at the time of the challenged transaction had there been full disclosure available.” Thus, when the fraudulent statement is made, if the plaintiff could not sell his shares, how could the statement have been said to damage the plaintiff at all? Economic loss is more attenuated from the misrepresentation, since any loss the plaintiff may suffer will only really occur after the one-year resale restriction is over.

3. Proof Problems

The Supreme Court distinguished loss causation from efficient market theory in Halliburton. However, subsequent court cases show the connection between loss causation and efficient markets. Loss causation is a result of the “market’s realization of the circumstances concealed by the [misrepresentation].” Plaintiffs’ experts conduct “event studies”

---

198 See, e.g., David Conn, Going Bust, GUARDIAN (Dec. 11, 2008), http://www.guardian.co.uk/business/2008/dec/12/recession-small-business.
200 Phillips, 489 F. App’x at 342-43.
201 See 15 U.S.C. § 77d-1(e) (2012); see also Arora, supra note 156.
202 The one-year restriction in sales allows any decent defense attorney to find some sort of fact, economic or otherwise, that may explain a loss. Phillips demonstrates that proving loss causation requires isolating the misrepresentation from all other possible explanations. See Phillips, 489 F. App’x at 343.
203 Thomas L. Hazen, 4 LAW SEC. REG. § 12.11, Causation in Actions Under Rule 10b-5.
205 See, e.g., In re Vivendi Universal, 765 F. Supp. 2d 512, 555 (S.D.N.Y. 2011) (“[Loss Causation] is typically shown by the reaction of the market to a ‘corrective disclosure’ which reveals a prior misleading statement.”).
that inquire as to what the “market price of a stock [would] be but for the fraud.”\textsuperscript{207} “Event studies” require efficient capital markets to be reliable.\textsuperscript{208} As discussed earlier, the one-year resale restriction effectively handicaps the secondary market, making the market for crowdfunding securities inefficient.\textsuperscript{209} Thus, the one-year resale restriction presents proof problems for plaintiffs.

IV. THE ONE-YEAR SALE RESTRICTION SHOULD BE MODIFIED

The SEC should exercise its authority to add to or modify the one-year sale restriction.\textsuperscript{210} The SEC can guard against fraud in the secondary markets, without going beyond its statutory authority, by allowing resales to other members who are registered with complying funding portals.\textsuperscript{211} By doing this, the SEC would ensure that the investor pool in crowdfunding resale markets is educated.\textsuperscript{212} This would protect consumers who are unable to guard against such fraud.\textsuperscript{213} A secondary market for crowdfunding shares might be troublesome in some respects,\textsuperscript{214} but it may also allow for information disseminated by issuers to affect the price more efficiently.\textsuperscript{215} This Part will explore the implications of loosening the one-year sale restriction on secondary markets and proving loss causation.\textsuperscript{216}

\begin{itemize}
\item \textsuperscript{207} Id. at 903.
\item \textsuperscript{208} Id.
\item \textsuperscript{209} Supra Part III (applying Cammer factors to crowdfunding market).
\item \textsuperscript{210} 15 U.S.C. § 77d-1(e)(2) (2012) (“[Resales] shall be subject to such other limitations as the Commission shall, by rule, establish.”).
\item \textsuperscript{211} The following resale provision explicitly grants the SEC authority to add additional rules to the resale provision: “[resale] shall be subject to such other limitations as the Commission shall, by rule, establish.” § 77d-1(e)(2) (emphasis added).
\item \textsuperscript{212} See § 77d-1(a)(4).
\item \textsuperscript{213} The courts have interpreted securities laws loosely when the class of investors can protect themselves. See SEC v. Ralston Purina Co., 346 U.S. 119, 125 (1953) (“[T]he applicability of § 4(1) should turn on whether the particular class of persons affected needs the protection of the Act. An offering to those who are shown to be able to fend for themselves is a transaction ‘not involving any public offering.’”).
\item \textsuperscript{214} See Heminway & Hoffman, supra note 37, at 954.
\item \textsuperscript{215} See Schleicher v. Wendt, 618 F.3d 679, 684 (7th Cir. 2010) (finding that since defendant’s stock was traded “in a liquid market,” in addition to other factors, it met fraud-on-the-market requirements).
\item \textsuperscript{216} Additionally, there are other challenges that the resale restriction pose that need to be addressed by the SEC. Bradford argues that there are other reasons why restrictions on resale are harmful. One detrimental effect of resale restrictions are that they may “cause issuers to lose their exemptions.” Bradford, supra note 4, at 144-45. Furthermore, resale restrictions could become a liability trap for unsophisticated investors. Id. at 144.
\end{itemize}
A. The Impact of a Loosened Resale Restriction on a Secondary Crowdfunding Securities Market.

Secondary securities markets allow investors to transact with each other instead of the issuer.\textsuperscript{217} Many such transactions take place on markets where large public corporations’ shares are traded.\textsuperscript{218} Yet, given the rate of failure of small businesses, crowdfunding issuers, who are small businesses, will likely not go public.\textsuperscript{219} On the other hand, the crowdfunding issuers, unlike private corporations, are in little control of who can buy their shares.\textsuperscript{220} It is in this respect that they resemble private corporations. Yet there is a rising trend for privately held securities to be traded in secondary markets.\textsuperscript{221} Thus, crowdfunding issuers may welcome, or at least tolerate, their securities being traded on secondary markets.\textsuperscript{222}

\textit{Dura} and Basic deal with two distinct elements of a 10b-5 claim, but they turn on the same issue.\textsuperscript{223} In \textit{Dura}, the Court stated that loss causation turned on the degree and time it takes for information to impact the market.\textsuperscript{224} In \textit{Basic}, the Court stated that the presumption of reliance turned on the market’s ability to accurately reflect information.\textsuperscript{225} By making the shares transferable, with some limits, the shares gain a

\begin{itemize}
  \item \textsuperscript{217} See Secondary Market definition, INVESTOPEDIA, http://www.investopedia.com/terms/v/venturecapital.asp (last visited Nov. 9, 2012) (example, is such a marketplace).
  \item \textsuperscript{218} Basic Inc. v. Levinson, 485 U.S. 224, 243-44 (1988) (“[M]odern securities markets, literally involve[d] millions of shares changing hands daily . . . .”).
  \item \textsuperscript{219} U.S. SMALL BUSINESS ADMINISTRATION, Frequently Asked Questions: Advocacy Small Business Statistics and Research, at 1, available at http://www.sba.gov/sites/default/files/sbfaq.pdf (last updated Jan. 2011) (“[H]alf [of all new businesses survive] 5 or more years”; cf. Carney, supra note 46, at 151 (“Aggregate compliance costs are likely to be staggering.”)).
  \item \textsuperscript{220} 15 U.S.C. § 77d-1(a) (2012) (the intermediary is largely responsible for a majority of the transaction).
  \item \textsuperscript{221} See Elizabeth Pollman, Information Issues on Wall Street 2.0, 161 U. PA. L. REV. 179, 193 (Dec. 2012).
  \item \textsuperscript{222} If the crowdfunding resales were limited to other crowdfunding investors, issuers likely will have no problem with their shares ending up in hands that could have bought shares when they were first issued.
  \item \textsuperscript{223} Compare Dura Pharm., Inc. v. Broudo, 544 U.S. 336, 343 (2005) (“I]lower price may reflect, not the earlier misrepresentation, but changed economic circumstances, changed investor expectations, new industry-specific or firm-specific facts, conditions, or other events, which taken separately or together account for some or all of that lower price.”), with Basic, 485 U.S. at 246 (“[M]arket price of shares traded on well-developed markets reflects all publicly available information, and, hence, any material misrepresentations.”).
  \item \textsuperscript{224} Dura, 544 U.S. at 342.
  \item \textsuperscript{225} Basic, 485 U.S. at 243-44.
\end{itemize}
higher price.\textsuperscript{226} Liquidity also helps to ensure that any information impacts the price of said stock.\textsuperscript{227} Thus, determining whether the market price of crowdfunding securities reflects information accurately requires evaluating the crowdfunding secondary market.

Secondary market performance plays a pivotal role in securities fraud litigation.\textsuperscript{228} The defrauded crowdfunder, through his or her class action, will inevitably need to show that the market for his or her crowdfunding shares is “open and developed.”\textsuperscript{229} The five \textit{Cammer} factors show that, with the one-year restriction, it is unlikely that courts will find that the market for crowdfunding shares would be open and developed.\textsuperscript{230} However, having a looser resale restriction changes that analysis. Specifically, if crowdfunding secondary markets resemble “Over The Counter Bulletin Boards” (OTCBB), then they are unlikely to be open and developed.\textsuperscript{231} The key is to distinguish crowdfunding markets from OTCBB’s.

An OTCBB is an “electronic quotation system that displays real-time quotes, last-sale prices and volume information for many over-the-counter securities that are not listed on a national securities exchange.”\textsuperscript{232} Instead, these markets rely on the traders to supply this information themselves.\textsuperscript{233} This is what makes them “over the counter.” Some courts have decided, as a matter of law, that these markets are undeveloped.\textsuperscript{234} Other courts use the five \textit{Cammer} factors on a case by case basis.\textsuperscript{235} Even applying those factors, most courts conclude that these

\begin{thebibliography}{9}
\bibitem{228} See Basic, 485 U.S. at 243-44.
\bibitem{229} \textit{Id.} at 241-42 (quoting \textit{Peil}, 806 F.2d at 1160-61).
\bibitem{230} See supra Part III.
\bibitem{233} OTCBB.COM, \textit{Investor Information}, http://www.otcbb.com/investorinformation/investorinfo.stm (last visited Jan. 17, 2012) (“Market Makers will be required to provide the periodic financial reports filed by OTCBB issuers with the SEC or other regulatory authorities pursuant to the Eligibility Rule.”).
\bibitem{234} See \textit{In re Data Access Sys. Sec. Litig.}, 103 F.R.D. 130, 138 (D.N.J. 1984), rev’d on other grounds by 848 F.2d 1537 (3d Cir. 1988).
\bibitem{235} See, e.g., Unger v. Amedisys Inc., 401 F.3d 316, 323 (5th Cir. 2005) (listing eight factors, the first five of which are the \textit{Cammer} factors).
\end{thebibliography}
OTCBB markets are inefficient when compared with more traditional secondary markets. 236

OTCBBs are not the only avenue for reselling stock. A new type of secondary market has been on the rise that caters to private placements. 237 Second Market is one example. 238 These markets rely on two factors—(1) the investor’s knowledge and sophistication and (2) the investor’s access to information—to exempt transactions from reporting requirements. 239 Using these factors, these marketplaces “match[] buyers and sellers of [private securities].” 240 Second Market is different from OTCBBs because it provides information for valuing stock. 241 Second Market also allows for investors to create profiles for themselves. 242 This decreases the chances of the investor being an anonymous backroom dealer. 243 Second Market has also “integrated aspects of social media” into its interface. 244 And finally, Second Market has begun to “offer[] and pay for analyst coverage of certain companies.” 245

The funding portals created by the JOBS Act have the potential to resemble Second Market. Funding portals cater to a special class of investors because they are required to ensure that investors are educated. 246 Further, these funding portals already interact with crowdfunding issuers by virtue of the fact that the primary sale takes place through these portals. 247 Resale markets for crowdfunding shares may also benefit from social networking. 248

236 See id. (“[S]uch holdings are indicative of the wide gulf between the type of markets for stocks that trade millions of shares daily, [citations omitted], and the much less active [OTCBB] market for stocks.”).
237 See Pollman, supra note 221, at 193.
238 See id. Eighty percent of the sellers in Second Market are former employees. Id. at 196.
239 See id. at 189.
240 See id. at 195.
241 See id. at 203. This is done by involving the issuer. See id.
242 See id. at 195.
243 In Basic the court, in evaluating whether or not a fraud-on-the-market presumption was warranted, took note of the fact that face-to-face transactions differed from faceless ones. Basic v. Levinson, 485 U.S. 224, 243-44 (1988). It is no jump in logic to conclude that allowing investors to create profiles allows for a higher degree of personal interaction between seller and buyer. Furthermore, many of the websites pitching themselves as crowdfunding intermediaries already resemble social media sites. See, e.g., Taylor, supra note 82.
244 See Pollman, supra note 221, at 199.
245 See id. at 198-99.
248 Social networking could be the equivalent of press for these crowdfunding issuers. In fact, many companies now advocate social networking as a way to spread
Crowdfunding shares may exceed OTCBB shares in numerosity, increasing the trade volume and analyst coverage of those shares. Thus, having a special class of investors in the crowdfunding context would help to guard against fraud while encouraging smarter investments.

It may be early to conclude that a resale market for crowdfunding securities will look like Second Market. But such a market would make for a more open market under the Cammer factors. The resale of crowdfunding securities entails a greater trading volume of such stock. Further, if Second Market is any indication, a resale market in crowdfunding shares would incentivize more analysts in the crowdfunding arena. A less restrictive resale provision would result in greater incentive for Qualified Institutional Buyers to transact in crowdfunding securities.

The informational issues mentioned earlier in this note would still be present. However, this is likely not fatal if, as some scholars have persuasively argued, the SEC mandates the disclosure of additional information. Whether or not this would warrant a fraud-on-the-market presumption is a matter for courts. But, as the next sub-section will show, the efficiency of the secondary market reverberates beyond fraud-on-the-market.

---


250 Heminway & Hoffman, supra note 37, at 936 (“Better investor education and stronger enforcement efforts should make the increase in fraud bearable, however.”) (quoting Dale A. Oesterle, 1 ENTREPREN. BUS. L. J. 369, 379 (2006)).


252 See id. at 1286.


254 Darian M. Ibrahim, The New Exit in Venture Capital, 65 VAND. L. REV. 1, 22-23 (Jan. 2012) (“[B]y improving liquidity for individual investors ex post, the direct market has the potential to increase the number of start-ups that will receive [venture capital] funding ex ante.”).

255 See supra Part I (discussing the JOBS Act and fraud).


257 Furthermore, whether or not there should be a fraud-on-the-market presumption in the crowdfunding context is an open question. On the one hand, as mentioned earlier, it is likely that most of litigation in the crowdfunding fraud area will be done in a class action setting. On the other hand, unlike securities that are traded on NASDAQ, usually involving larger corporations that can afford litigation, crowdfunding issuers are small businesses that are unlikely to be able to afford defending a class action suit. See Ian Simmons & Charles E. Borden, The Defense Perspective: The Class Action
B. How an Active Secondary Market Affects Loss Causation

Even if the crowdfunding secondary market is not developed enough to warrant a fraud-on-the-market presumption, a rigorous resale market makes it easier to establish loss causation. Loosening the resale restriction accomplishes two things. First, it reduces the chances of intervening events affecting market price. The looser resale restriction narrows the gap between the fraudulent misrepresentation and the economic effect on the investor’s shares. Narrowing that gap further reduces the likelihood that other intervening causes will touch upon the stock price loss. Per Dura, this makes it more likely that the misrepresentation impacted the shares as opposed to other factors.

Second, loosening the resale restriction also addresses a crucial proof problem. Greater share liquidity entails a more efficient market. An efficient crowdfunding securities market helps plaintiffs prove loss causation. This is because a more efficient crowdfunding securities market means that the plaintiff will be able to isolate the effect of the fraudulent statement from other factors. Thus, an event study may be more reliable.

Secondary markets may be vulnerable to informational issues that would lead to “pump and dump” schemes. But what will happen once the one-year restriction on transfers is over? The SEC surely cannot restrict resale ad infinitum. By adopting the aforementioned rule, the SEC can ensure that the
resale market is continuously informed and educated. A regulated secondary market with informed investors that more accurately reflects price strikes a proper balance. Honest crowdfunding businesses gain the added benefit of easier access to capital, while loosening the resale restriction would also make it easier to pursue claims against fraudsters. The benefit of added investment should well outweigh the fraud that occurs.

CONCLUSION

The JOBS Act’s crowdfunding provisions may not please everyone. However, there are no signs of Congress legislating otherwise. Critics and proponents of the JOBS Act’s crowdfunding provisions have made important points that the SEC should address as it makes crowdfunding rules. Thus, the SEC should keep litigation hurdles in mind when it makes rules regarding the resale of crowdfunding shares. A looser resale restriction rule that guides the transition of crowdfunding securities into the secondary market serves anti-fraud purposes while providing an incentive for investors to join the crowdfunding market. This would allow crowdfunding laws to better fit the mosaic of federal securities regulation.

Sherief Morsy

---

268 See supra Part IV (discussing the impact of loosening the one-year restriction on the secondary market).
269 See Ibrahim, supra note 254, at 22-23.
270 See supra Part III.C.3 (discussing the event studies).
271 In Stoneridge Inv. Partners, LLC v. Scientific-Atlanta Justice Stevens, in his dissent, noted:

The success of the U.S. securities markets is largely the result of a high level of investor confidence in the integrity and efficiency of our markets. The SEC enforcement program and the availability of private rights of action together provide a means for defrauded investors to recover damages and a powerful deterrent against violations of the securities laws.

† J.D. Candidate, Brooklyn Law School, 2014; B.A., Rutgers University, 2010. I want to thank the Brooklyn Law Review staff for their assistance. I would like to thank Professor Roberta Karmel for providing guidance for my note. Importantly, I would like to thank my family for their love and support. Lastly, I want to dedicate this to Isis Issa, my grandmother, for always believing in me and inspiring me to pursue this career.
BROOKLYN LAW SCHOOL FACULTY 2013-2014

Faculty Emeriti

RICHARD ALLAN, B.A., LL.B., Professor of Law Emeritus
JOSEPH CREA, B.A., J.D., LL.M, LL.D., Professor of Law Emeritus
MARY FALK, B.A., J.D., Associate Professor of Legal Writing Emerita
MARTIN HAUPTMAN, LL.B., LL.M., Professor of Law Emeritus
WILLIAM HELLERSTEIN, B.A., J.D., Professor of Law Emeritus
NORMAN POSEY, A.B., LL.B., Professor of Law Emeritus
GARY SCHULTZKE, B.A., J.D., Professor of Law Emeritus

Faculty

NICHOLAS ALLARD, B.A., B.A./M.A., J.D., Joseph Crea Dean and Professor of Law
WILLIAM ARAIZA, B.A., M.S., J.D., Professor of Law
CAMERON W. ARNOLD, B.A., M.A., J.D., Instructor of Legal Writing
JONATHAN ASKIN, A.B., J.D., Associate Professor of Clinical Law
MIRIAM BAER, A.B., J.D., Associate Professor of Law
CHRISTOPHER BEAUCHAMP, B.A., M.Phil., Ph.D., Assistant Professor of Law
DEBRA BECHTEL, B.A., J.D., Associate Professor of Clinical Law
URSULA BENTLE, B.A., J.D., Professor of Law
ANITA BERNSTEIN, B.A., J.D., Anita and Stuart Subotnick Professor of Law
BRADLEY BORDEN, B.B.A., M.B.A., J.D., LL.M., Professor of Law
DANA BRALMAN REISER, B.A., J.D., Vice Dean and Professor of Law
MICHAEL CAHILL, B.A., M.P.P., J.D., Vice Dean and Professor of Law
I. BENNETT CAPERS, B.A., J.D., Professor of Law
STACY CAPLOW, B.A., LL.M., Associate Dean for Professional Legal Education and Professor of Law
NEIL B. COHEN, S.B., J.D., Jeffrey D. Forchelli Professor of Law
STEVEN DEAN, B.A., J.D., Professor of Law
ROBIN EFFRON, B.A., J.D., Associate Professor of Law
ELIZABETH FAJANS, B.A., M.A., Ph.D., Associate Professor of Legal Writing
JAMES FANTO, B.A., M.A., Ph.D., J.D., Professor of Law
RICHARD FARRELL, B.A., LL.B., Wilbur A. Levin Distinguished Service Professor of Law
LINDA FELDMAN, B.A., M.A., J.D., Assistant Professor of Legal Writing and Director, Academic Success Program
MARYELLEN FULLERTON, B.A., J.D., Professor of Law
MARSHA GARRISON, B.A., J.D., Suzanne J. and Norman Miles Professor of Law
MICHAEL GERBER, B.A., J.D., Professor of Law
CYNTHIA GODSOE, A.B., J.D., Assistant Professor of Law
JOEL GORA, B.A., LL.B., Professor of Law
SUSAN GRIFFEN, B.A./B.S., J.D., Instructor of Legal Writing
SUSAN HERMAN, B.A., J.D., Centennial Professor of Law
EDWARD JANGER, B.A., M.D., David M. Barse Professor of Law
BRYL JONES-WOODIN, B.A., J.D., Professor of Law
ROBERTA KARMEL, B.A., LL.B., Centennial Professor of Law
ADAM KOLBER A.B., J.D., Professor of Law
MINNA KOTKIN, A.B., J.D., Professor of Law
JAY E. KUKLIN, B.S., J.D., Professor of Law
REBECCA KVSAR, B.A., J.D., Associate Professor of Law
BRIAN LEE, B.A., M.A., J.D., Associate Professor of Law
GREGG MACEY, B.A., M.A., J.D., Ph.D., Associate Professor of Law
ROSIEANN MACKECHNIE, B.A., J.D., Instructor of Legal Writing
GARY MINDA, B.A., M.A., S.J.D., Professor of Law
SAMUEL MURUMBA, LL.B., LL.M., Ph.D., Professor of Law
MINOR MYERS, B.A., J.D., Associate Professor of Law
Oona Peterson, B.A., J.D., Instructor of Legal Writing
ARTHUR PINTO, A.B., J.D., Professor of Law
ROBERT PITLER, B.B.A., LL.B., S.J.D., Professor of Law
KAREN PORTER, B.A., M.S., J.D., Associate Professor of Clinical Law
DAVID REISS, B.A., J.D., Professor of Law
JAYNE RESSLER, B.S., J.D., Associate Professor of Legal Writing
ELIZABETH SCHNEIDER, B.A., M.Sc., J.D., Rose L. Hoffer Professor of Law
JANET SINDER, A.B., J.D., M.S., Director of Library and Associate Professor of Law
LISA SMITH, B.A., J.D., Assistant Professor of Clinical Law
DAN SMULIAN, B.A., M.Phil., J.D., Associate Professor of Clinical Law
LAWRENCE SOLAN, B.A., Ph.D., J.D., Don Forchelli Professor of Law and Director of Graduate Education
WINNIE TAYLOR, B.A., J.D., LL.M., Professor of Law
NELSON TERBE, A.B., J.D., Ph.D., Professor of Law
CARRIE TEITCHER, B.A., J.D., Assistant Professor of Legal Writing
AARON TWINER, B.S., J.D., Irwin and Jill Cohen Professor of Law
MARIANNA WALDER, B.A., J.D., Professor of Law
JOAN G. WEXLER, B.S., M.A.T., J.D., Dean and President Emerita, Professor of Law
Visiting Faculty

HILARY ALLEN, B.A., LL.B. LL.M., Visiting Associate Professor of Law
JEFFREY ATIK, A.B., J.D., Ph.D., Dr. Jur., Visiting Professor of Law
EMILY BERMAN, B.A., J.D., LL.M., Visiting Assistant Professor of Law
NEIL COHEN, B.A., J.D., LL.M., Diploma in Criminology, Visiting Professor of Law
LOUISE CROWLEY, B.C.L., LL.M., Ph.D., Visiting Professor of Law
STEPHAN LANDSMAN, B.A., J.D., Visiting Professor of Law
ROGER MICHALSKI, B.A., Ph.D., J.D., Visiting Assistant Professor of Law
MELISSA MORTAZAVI, B.A., J.D., Visiting Assistant Professor of Law
ANDREW NAPOLITANO, B.A., J.D., Distinguished Visiting Professor of Law
THANK PITTMAN, B.A., M.A., Ph.D., Visiting Professor of Law
GERALD SHARGEL, B.A., J.D., Practitioner-in-Residence
BRAD SNYDER, A.B., J.D., Visiting Associate Professor of Law
ALAN M. TRAMMEL, B.A., M.Sc., M.St., J.D., Visiting Assistant Professor of Law
SHONNA TRINCH, B.A., M.A., Ph.D., Visiting Professor of Law

Librarians

ROSEMARY H. CAMPAGNA, B.A., M.S.L.I.S., Reference/Government Documents Librarian
GILDA CHIU, B.A., M.L.S., Collection Development/Acquisitions Librarian
KATHY DARVIL, A.B., M.S.I., J.D., Access Services/Reference Librarian and Adjunct Assistant Professor of Law
JEAN J. DAVIS, B.A., J.D., M.S.L.I.S., Associate Librarian for International Law and Adjunct Professor of Law
SARA GRAS, B.S., J.D., M.S.L.I.S., Reference Librarian and Adjunct Assistant Professor of Law
LINDA HOLMES, B.A., M.S., Associate Librarian
HAROLD O’GRADY, B.A., J.D., M.L.S., Reference Librarian and Adjunct Professor of Law
JANET SINDER, A.B., J.D., M.S., Director of Library and Associate Professor of Law
HAINAN YU, B.A., M.S., Systems Librarian
BROOKLYN LAW SCHOOL BOARD OF TRUSTEES

STUART SUBOTNICK ’68
Chairman of the Board
President and Chief Executive Officer
Metromedia Company

HON. EDWARD R. KORMAN ’66
Vice Chairman of the Board
Senior Judge
United States District Court for the Eastern District of New York

BURNSIDE E. ANDERSON III ’76
Senior Corporate Counsel, Legal Division
Pfizer Inc. (Ret.)

FRANCIS J. AQUILA ’83
Sullivan & Cromwell LLP

HON. HERMAN BADILLO ’54

DAVID M. BARSE ’87
Chief Executive Officer
Third Avenue Management, LLC

FREDERICK COHEN ’67
Duane Morris LLP

JEFFREY J. FEIL ’73
Chief Executive Officer
The Feil Organization

MARTIN A. FISCHER ’64
West Center Associates

JEFFREY D. FORCHELLI ’69
Forchelli, Curto, Deegan, Schwartz, Mineo & Terrana, LLP

SHANNON M. HALEY ’08
Sullivan & Cromwell LLP

ROBERT M. KAUFMAN ’57
Proskauer Rose LLP

LEO KITTA’ ’08
Fross Zelnick Lehrman & Zissu, P.C.

EILEEN T. NUGENT ’78
Skadden, Arps, Slate, Meagher & Flom LLP
FLORENCE SUBIN ’75
Subin Associates

STEVEN L. ZELKOWITZ
Chief Executive Officer
Sycamore Energy Consulting

Members Emeriti

ROBERT B. CATELL
Chairman
Advanced Energy Research and Technology Center (AERTC) at
New York State University at Stony Brook

FRANCES MARGOLIS FRIEDMAN ’39