

**Two Types of Consequentialism, Two Types of Formalism:
Reconsidering *Bordenkircher* in Light of *Apprendi***

Michael M. O’Hear*

*While the Supreme Court approved of the use of charging threats nearly thirty years ago in *Bordenkircher v. Hayes*, a more recent line of cases has subtly undermined key premises of that landmark decision. In order to induce guilty pleas, prosecutors might use any of a number of different tactics. A prosecutor might, for instance, charge aggressively in the first instance and then promise to drop the most serious charges in return for a guilty plea to a lesser offense. *Bordenkircher* addressed the mirror-image of this tactic: the prosecutor filed relatively minor charges at first, but then threatened to pursue more serious charges if the defendant did not plead guilty. The Supreme Court approved of such charging threats based on two considerations: the efficiency benefits of resolving cases by plea instead of jury trial, and the possibility that prosecutors would evade a ban on threats by charging more aggressively in the first instance. The Court’s reasoning, however, is inconsistent with *Apprendi v. New Jersey* and its progeny. *Apprendi* rejected the use of both efficiency considerations and evasion concerns as grounds for impairing access to juries. *Apprendi* instead emphasized a need for robust checks and balances within the criminal justice system. Because the *Apprendi* line of cases addressed sentencing procedures, not plea bargaining, their relevance to *Bordenkircher* has thus far escaped notice. The Article argues, however, that the Court should now overturn *Bordenkircher* in light of the values it embraced in *Apprendi*. The Article also proposes a new test for evaluating the constitutionality of charging threats.*

I. The <i>Bordenkircher</i> Dilemma	6
A. <i>Bordenkircher</i>	7
B. <i>Goodwin</i>	10
C. Summary	14

* Associate Professor, Marquette University Law School. J.D. Yale Law School, 1996; B.A., Yale College, 1991. Editor, *Federal Sentencing Reporter*. I am grateful for comments on an earlier draft by Daniel Capra, Jason Czarnecki, Scott Moss, Chad Oldfather, Mark Osler, Stephanie Stern, and Sandra Guerra Thompson.

II. A Parallel Dilemma: Elements Versus Sentencing Factors	14
A. <i>McMillan</i> : Taking the Powell Approach Again	17
B. The <i>Apprendi</i> Canon: Stevens’s Triumph	19
C. <i>Apprendi</i> ’s Aims	22
D. Is It Really Just “Doctrinaire Formalism”?	28
III. Inconsistencies Between <i>Apprendi</i> and <i>Bordenkircher</i>	40
A. Common Law Values Versus Bureaucratic Efficiency	40
B. Biased Actors Versus Rational Actors	42
IV. Designing a New Rule for Charging Threats	47
A. Structural Options	47
B. Should Stevens Adhere to <i>Bordenkircher</i> ?	52
C. More “Doctrinaire Formalism”?	54
V. Conclusion	66

It is an iconic case from the era when plea bargaining finally emerged from the shadows and acquired clear constitutional legitimacy.¹ In 1973, a Kentucky grand jury indicted Paul Hayes on a charge of uttering a forged instrument in the amount of \$88.30, an offense punishable by two to ten years in prison.² During plea negotiations, the prosecutor offered to recommend a term of five years if Hayes pled guilty, but, otherwise, threatened to charge Hayes under the Kentucky Habitual Criminal Act, which would subject Hayes to a mandatory life term based on his two prior felony convictions.³ When Hayes rejected the deal, he was charged as threatened.⁴ Eventually, he lost at trial and was sentenced

¹ See *Bordenkircher v. Hayes*, 434 U.S. 357, 365 (1978) (rejecting constitutional rule “that would drive the practice of plea bargaining back into the shadows from which it has so recently emerged”). For a leading history of plea bargaining from its first systematic use in the nineteenth century to the present, see GEORGE FISHER, *PLEA BARGAINING’S TRIUMPH: A HISTORY OF PLEA BARGAINING IN AMERICA* (2003).

² *Bordenkircher*, 434 U.S. at 358.

³ *Id.*

⁴ *Id.* at 359.

pursuant to the Act.⁵ Thus, for his decision to go to trial, Hayes suffered an extraordinary penalty, as his sentence ballooned from about five years to life.

The case, styled *Bordenkircher v. Hayes*, created a dilemma when it reached the Supreme Court. On the one hand, as the Court had established in a recent line of cases, “[t]o punish a person because he has done what the law plainly allows him to do is a due process violation of the most basic sort.”⁶ Because Hayes indisputably had a constitutional right to his jury trial, the prosecutor’s decision to penalize him in such a severe fashion for exercising the right was troubling at the very least.

On the other hand, in another recent line of cases, the Court had endorsed the guilty plea and plea bargain as “important components of this country’s criminal justice system.”⁷ A ruling in favor of Hayes might bring down the whole edifice. The Court could see no principled means of distinguishing among plea inducements: a “threat” (like the prosecutor’s threat to charge Hayes as a recidivist) could easily be restructured as an “offer” (e.g., the prosecutor might have charged Hayes as a recidivist from the outset and then offered to dismiss the charge if Hayes pled guilty to the underlying offense). Thus, if the Court were to prohibit charging *threats*, the Court could not stop there, but would have to regulate *all* plea inducements—a radical step the Court was unwilling to make.⁸ Accordingly, the Court rejected Hayes’ claim.⁹

But is it *really* so hard to distinguish threats from offers? Although its potential connections to *Bordenkircher* have thus far escaped notice, a much more recent line of cases casts doubt on crucial premises of the earlier decision. Since 2000, *Apprendi v. New Jersey*¹⁰ and its progeny¹¹ have rejuvenated the Sixth Amendment right to a jury trial. Even though these decisions dealt with sentencing procedures, not plea bargaining, they have nonetheless undercut *Bordenkircher* in at least two respects. First, and most obviously, *Apprendi*’s vision of a robust role for the jury in the criminal justice system is inconsistent

⁵ *Id.*

⁶ *Id.* at 363 (citing *North Carolina v. Pearce*, 395 U.S. 711, 738 (1969)).

⁷ *Id.* at 361 (quoting *Blackledge v. Allison*, 431 U.S. 63, 71 (1977)).

⁸ *See id.* at 364 (“While confronting a defendant with the risk more severe punishment clearly may have a discouraging effect on the defendant’s assertion of his trial rights, the imposition of these difficult choices is an inevitable—and permissible—attribute of any system which tolerates and encourages the negotiation of pleas.” (quotation marks, alteration, and citation omitted)).

⁹ *Id.* at 365.

¹⁰ 530 U.S. 466 (2000).

¹¹ *See, e.g.*, *United States v. Booker*, 543 U.S. 220 (2005); *Blakely v. Washington*, 542 U.S. 296 (2004).

with *Bordenkircher*'s endorsement of guilty pleas as the most desirable method of resolving criminal cases. Second, and more subtly, the *Apprendi* decisions recognize the significance of the *form* of a government action. Two different government actions producing the same result may be *perceived* quite differently based on differences in their form. The Court's reasoning in this regard, which resonates with recent work in cognitive psychology, points the way to a principled distinction between threats and offers.

Building on these observations, the central thesis of this Article is easily stated: in the interests of jurisprudential consistency, the five Justices constituting the majority in *Apprendi*—all of whom remain on the new Roberts Court—should, if given the opportunity, vote to overturn *Bordenkircher*.¹² The thesis is considerably easier to state than to defend—if for no other reason than that the only member of the “*Apprendi* Five” who was on the Court in 1978 (Justice Stevens) actually voted with the majority in *Bordenkircher*. Yet, as we will see, Stevens subtly backed away from *Bordenkircher* in a later opinion,¹³ and good arguments may be made that he should now reject even his watered-down version of *Bordenkircher* as inconsistent with his *Apprendi* writings.

In order to make these arguments, we shall have to puzzle over one the ongoing points of controversy surrounding the *Apprendi* decisions: whether they constitute, as Justice O'Connor dismissively characterized them, “doctrinaire formalism.”¹⁴ The question has attracted a growing body of scholarly commentary.¹⁵ I will attempt to illuminate the issues by distinguishing between

¹² While I argue that the Court ought to reject one particular form of plea inducement, I do not seek in this Article to join the long-standing academic debate over the merits of plea-bargaining generally. See, e.g., Albert W. Alschuler, *Implementing the Criminal Defendant's Right to Trial: Alternatives to the Plea Bargaining System*, 50 U. CHI. L. REV. 931, 932-34 (1983) (summarizing arguments against plea bargaining); Frank H. Easterbrook, *Criminal Procedure as a Market System*, 12 J. LEGAL STUD. 289, 309 (1983) (“[P]lea bargaining is desirable, not just defensible, if the system attempts to maximize deterrence from a given commitment of resources.”); Stephen J. Schulhofer, *Plea Bargaining as Disaster*, 101 YALE L. J. 1979, 1979 (1992) (“[P]lea bargaining seriously impairs the public interest in effective punishment of crime and in accurate separation of the guilty from the innocent.”); Robert E. Scott & William J. Stuntz, *Plea Bargaining as Contract*, 101 YALE L. J. 1909, 1910 (1992) (“Properly understood, classical contract theory supports the freedom to bargain over criminal punishment.”).

¹³ *United States v. Goodwin*, 457 U.S. 368 (1982).

¹⁴ *Blakely*, 542 U.S. at 321 (O'Connor, J., dissenting).

¹⁵ Most commentators on this issue have echoed O'Connor's criticism. See, e.g., Ronald J. Allen & Ethan A. Hastert, *From Winship to Apprendi to Booker: Constitutional Command or Constitutional Blunder?*, 58 STAN. L. REV. 195, 202 (2005) (“To the extent what emerges from the [*Apprendi*] cases is something other than drafting advice, it will most likely redound to the detriment of defendants, a curious result in a line of cases ostensibly designed to protect defendants' rights.”);

two types of formalism and two types of consequentialism. First, I distinguish “rule-structure formalism” from “consequence-indifference formalism.” The former principle indicates that, in establishing the boundaries of legal rights and duties, courts ought to favor bright-line tests over vague standards.¹⁶ The latter indicates that courts ought to decide cases based on some facially neutral principle, like plain textual meaning or original intent, without regard to the real-world consequences of the decision. It is *this* type of formalism that Justice O’Connor had in mind with her criticism of *Apprendi*. However, while the *Apprendi* Five undoubtedly employ rule-structure formalism, and also sometimes don the mantle of originalism, they just as clearly believe they are doing something that materially enhances the fairness and democratic accountability of the criminal justice system. And because these beliefs are at least plausible, the *Apprendi* decisions should not be dismissed as indifferent to consequences.

I further distinguish between “rational-actor consequentialism” and “biased-actor consequentialism.” Any attempt to consider real-world consequences in judicial decision-making must rely on some basic assumptions about the way that people behave in the real world. Law and economics scholars have long employed rational-actor models to predict the social consequences of changes in legal rules. These models assume that people seek to maximize their utility from a stable set of preferences and accumulate an optimal amount of information in order to do so.¹⁷ A competing approach, sometimes labeled behavioral law and economics (“BLE”), relaxes these assumptions and attempts to take into account how “‘real people’ differ from *homo economicus*.”¹⁸ Of particular note, the “real

Stephanos Bibas, *Originalism and Formalism in Criminal Procedure: The Triumph of Justice Scalia, the Unlikely Friend of Criminal Defendants?*, 94 GEO. L.J. 183, 186 (2005) (“[Justice Scalia] has done a disservice in *Blakely* . . . by exalting formalism so far above humility, practicality, and plain old common sense.”); Frank O. Bowman, III, *Function Over Formalism: A Provisional Theory of the Constitutional Law of Crime and Punishment*, 17 FED. SENT’ING RPTR. 1, 2 (2004) (criticizing “narrow formalism” of *Blakely*); Susan N. Herman, *Applying Apprendi to the Federal Sentencing Guidelines: You Say You Want a Revolution?*, 87 IOWA L. REV. 615, 638 (2002) (characterizing *Apprendi* as “highly formalistic”); Jeffrey Standen, *The End of the Era of Sentencing Guidelines*, 87 IOWA L. REV. 775, 783 (2002) (“[T]he Court will not be able, through formalistic attacks from the fringe, to diminish the preeminent discretion of the legislature to define crimes.”). For a more positive view, see Nancy J. King & Susan R. Klein, *Essential Elements*, 54 VAND. L. REV. 1467, 1486 (2001) (“[T]he formalism of *Apprendi* is not pointless after all.”).

¹⁶ See Michael M. O’Hear, *The Original Intent of Uniformity in Federal Sentencing*, 74 U. CIN. L. REV. 749, 755 (2006) (contrasting rules and standards).

¹⁷ Christine Jolls, Cass R. Sunstein & Richard Thaler, *A Behavioral Approach to Law and Economics*, 50 STAN. L. REV. 1471, 1476 (1998) (citing GARY S. BECKER, *THE ECONOMIC APPROACH TO HUMAN BEHAVIOR* 14 (1976)).

¹⁸ *Id.*

people” of BLE exhibit various forms of cognitive bias, such as loss aversion, that is, the tendency to weigh losses more heavily than gains.¹⁹

As we will see, the consequentialism of the *Bordenkircher* majority rests on rational-actor assumptions, while the *Bordenkircher* dissenters implicitly assume biased actors. In the *Apprendi* decisions, these roles are reversed: the majority employs biased-actor consequentialism, while the dissenters adhere to the rational-actor model. These different premises are not acknowledged in any of the cases, but they are readily teased out and help to explain how it is that *all* of the Justices seem to be claiming the consequentialist high ground.

The Article proceeds as follows. Part I elaborates on the *Bordenkircher* dilemma, identifying the competing approaches to resolving the dilemma taken by the various majority and dissenting opinions in *Bordenkircher* and an important follow-up case, *United States v. Goodwin*.²⁰ Part II discusses a parallel dilemma addressed by the *Apprendi* line of cases: to what extent should the Court defer to legislative decisions about which facts in a criminal case are “elements” (the determination of which implicates the full range of criminal procedure protections) and which facts are mere sentencing considerations (as to which procedures may be far more relaxed)? Countering Justice O’Connor’s charge of “doctrinaire formalism,” this Part offers a consequentialist account of the *Apprendi* line of cases, in which they can be seen through the lens of BLE as advancing democratic, libertarian, and fairness values in the criminal justice system. Part III brings together the *Bordenkircher* and *Apprendi* lines of cases, detailing their inconsistencies. Part IV proposes a new rule of constitutional law that the *Apprendi* Five should adopt if given a fresh opportunity to address the issue of prosecutorial charging threats. The test is consistent with the rule-structure formalism of *Apprendi*, while also taking more seriously the jury-trial right than the rules adopted in *Bordenkircher* and *Goodwin*.

I. THE BORDENKIRCHER DILEMMA

Prosecutors often seek to induce guilty pleas by offering defendants a benefit in return for their plea, such as the dismissal of one charge in exchange for a plea to a lesser-included offense. *Bordenkircher*, though, involved a different sort of inducement: a *threat* to bring a *greater* charge. The case presented a dilemma for the Court because it represented a collision between two established principles. On the one hand, the “vindictiveness” principle, adopted by the Court in such

¹⁹ *Id.* at 1484.

²⁰ 457 U.S. 368 (1982).

cases as *Blackledge v. Perry*,²¹ forbade prosecutors and judges from penalizing defendants for the exercise of procedural rights. But, on the other hand, the Court had also recently come to recognize the important social benefits of guilty pleas and plea-bargaining, which facilitated the efficient resolution of cases and created opportunities for defendants to receive more lenient treatment.²²

To be more specific, the *Bordenkircher* dilemma involved at least three difficult, overlapping questions. First, should the law distinguish between threats of harsh treatment and offers of lenient treatment when, at least in principle, threats could be easily restructured as offers? Second, how much weight should be given to the social benefits of the plea-inducement system? (This question was especially pressing if threats were indistinguishable from offers, in which case Paul Hayes’s claim might call into question all forms of plea inducement.) And, third, if defendants *did* have some sort of constitutional protection from unduly coercive plea inducements, how exactly would the right be structured?

The three opinions in *Bordenkircher* (a majority opinion by Justice Stewart and dissents by Justices Blackmun and Powell) suggest three different ways to resolve the tripartite dilemma. This Part considers each of the opinions, then assesses the majority and dissenting opinions in *Goodwin*. Although *Goodwin* did not involve a prosecutorial charging threat, the opinions nonetheless addressed the meaning of *Bordenkircher* at some length. Indeed, the *Goodwin* majority opinion (authored by Justice Stevens, a member of the *Apprendi* Five) characterized *Bordenkircher* in terms that were startlingly similar to Justice Powell’s *dissent* in the earlier case. Meanwhile, Justice Brennan’s dissent in *Goodwin* suggests a fourth approach to resolving the *Bordenkircher* dilemma—the one that turns out to be the closest in spirit to the *Apprendi* decisions.

A. *Bordenkircher*

1. *Majority Opinion: A Rule of Non-Interference With Plea Inducements*

In the majority’s view, *Bordenkircher*’s outcome resulted inevitably from the social desirability of plea inducements. The majority saw no meaningful distinction between threats and offers,²³ and so framed Hayes’ argument as an attack on plea inducements generally.²⁴ Observing that “the guilty plea and the

²¹ 417 U.S. 21 (1974).

²² See *Bordenkircher v. Hayes*, 434 U.S. 357, 361 (1978) (discussing benefits of plea-bargaining).

²³ *Id.* at 363.

²⁴ See *id.* at 364-65 (“To hold that the prosecutors’ desire to induce a guilty plea is an unjustifiable standard, which . . . may play no part in his charging decision, would contradict the very premises that underlie the concept of plea-bargaining itself.”).

often concomitant plea bargain are important components of this country's criminal justice system,"²⁵ the majority could find no fault in a prosecutor's desire to secure a plea using his broad charging discretion,²⁶ so long as the minimal requirements of probable cause were satisfied.²⁷ The majority acknowledged that the threat of more severe punishment "may have a discouraging effect on the defendant's assertion of his trial rights," but nonetheless concluded that "the imposition of these difficult choices [is] an inevitable—and permissible—attribute of any legitimate system which tolerates and encourages the negotiation of pleas."²⁸ The majority thus suggested no circumstances in which a prosecutor's filing of otherwise permissible charges pursuant to a plea-inducement threat would raise constitutional problems.

2. *Blackmun's Dissent: Questioning the Legitimacy of the Plea-Inducement Motive*

Justice Blackmun, joined by Justices Brennan and Marshall, found the vindictiveness principle controlling. Notwithstanding the unquestioned social benefits of guilty pleas, Blackmun argued that it was unconstitutional for a prosecutor purposely to impose a penalty on a defendant as a result of the defendant's decision to contest the prosecutor's case at trial.²⁹ In order to protect defendants from vindictiveness, he proposed a burden-shifting test that focused on the prosecutor's intent:

[W]hen plea negotiations, conducted in the face of the less serious charge under the first indictment, fail, charging by a second indictment a more serious crime for the same conduct creates a 'strong inference' of vindictiveness. . . . I . . . do not understand why . . . due process does not require that the prosecution justify its action on some basis other than discouraging [the defendant] from the exercise of his right to a trial.³⁰

²⁵ *Id.* at 361.

²⁶ *See id.* at 364 ("[B]y tolerating and encouraging the negotiation of pleas, this Court has necessarily accepted as constitutionally legitimate the simple reality that the prosecutor's interest at the bargaining table is to persuade the defendant to forgo his right to plead not guilty.").

²⁷ *See id.* ("[S]o 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.").

²⁸ *Id.* at 364.

²⁹ *Id.* at 367 (Blackmun, J., dissenting).

³⁰ *Id.*

Blackmun recognized that his approach might cause prosecutors to restructure threats as offers.³¹ In other words, rather than employ a “low-to-high” charging strategy, some prosecutors might switch to “high-to-low.” Blackmun further recognized that judges would not be able to detect and sanction such “overcharging.”³² Yet, Blackmun believed it was preferable for prosecutors to employ high-to-low rather than low-to-high; the switch would protect those defendants who were most determined to go to trial from the most severe reprisals, facilitate public oversight of plea-inducement practices, and give defendants a fairer chance to challenge the enhanced charges.³³ Blackmun, in short, concluded that there were meaningful differences between threats and offers. (His analysis will be discussed in more detail in Part III below.)

3. *Powell’s Dissent: Deference With a Difference*

Justice Powell staked out the middle ground. On the one hand, he agreed with the majority that plea inducements are “essential to the functioning of the criminal justice system,” and that, in general, prosecutors should be “accorded the widest discretion” in attempting to secure pleas.³⁴ He was unwilling, however, to endorse the majority’s view that prosecutors had an entirely free hand to enhance charges in response to a defendant’s failure to plead guilty. At *some* point, the new charge might become “unreasonable and not in the public interest,”³⁵ and it was at that point that Powell would join Blackmun in condemning the prosecutor’s motives. Put differently, the prosecutor’s desire to penalize the exercise of trial rights was immaterial as long as the resulting charges were also justified by society’s legitimate interest in punishing the defendant’s underlying criminal conduct; however, “[i]mplementation of a strategy calculated *solely* to deter the exercise of constitutional rights is not a constitutionally permissible exercise of discretion.”³⁶

Powell thus proposed a test that focused not on whether the prosecutor wished to penalize the defendant’s exercise of a right, but on “whether the prosecutor reasonably might have charged the [defendant with the new crime] in the first place.”³⁷ On the unique facts of *Bordenkircher*, where the defendant was subject to a life sentence for an \$88 crime, Powell concluded that the charges

³¹ *Id.* at 368 (Blackmun, J., dissenting).

³² *Id.* at 368 n.2 (Blackmun, J., dissenting).

³³ *Id.* (Blackmun, J., dissenting).

³⁴ *Id.* at 672-73 (Powell, J., dissenting).

³⁵ *Id.* at 371 (Powell, J., dissenting).

³⁶ *Id.* at 373 (Powell, J., dissenting) (emphasis added).

³⁷ *Id.* at 370 (Powell, J., dissenting).

were not reasonable.³⁸ Thus, while Powell’s approach was more deferential to prosecutors than Blackmun’s, Powell would have imposed some limitations on prosecutors in “exceptional case[s].”³⁹

B. Goodwin

Following an altercation with a police officer, Learley Reed Goodwin was charged with several misdemeanor and petty offenses.⁴⁰ After plea negotiations failed, the case was transferred to another prosecutor, who obtained a felony indictment, thus exposing Goodwin to more severe punishment.⁴¹ Convicted at trial, Goodwin argued on appeal that prosecutors had unconstitutionally retaliated against him for invoking his right to a jury trial.⁴² The claim, however, was rejected by the Supreme Court.

Although *Goodwin* did not arise from an express threat, the case merits our attention for two reasons. First, because the Court’s analysis turned on the scope of the *Bordenkircher* exception to prior vindictiveness law, *Goodwin* contains much commentary on the earlier decision. Second, the *Goodwin* majority opinion was authored by Justice Stevens, who later became a key figure in the *Apprendi* cases. Before turning to the majority opinion, though, we will begin with the dissent, authored by Justice Brennan and joined by Justice Marshall, which suggests a fourth distinct approach to vindictiveness.

1. Brennan’s Dissent: The Likelihood-of-Deterrence Test

In Brennan’s view, *Bordenkircher* was limited to its particular circumstances, a prosecutor carrying through on a threat made during plea negotiations.⁴³ Outside that context, vindictiveness analysis turned on the questions considered by the pre-*Bordenkircher* vindictiveness cases: “Did the elevation of the charges against [the defendant] pose a realistic likelihood of vindictiveness? Is it possible that the fear of such vindictiveness may unconstitutionally deter a person in [the defendant’s] position from exercising his statutory and constitutional right to a jury trial?”⁴⁴ And, as Brennan saw it, the government’s elevation of charges against Goodwin did indeed pose such “a realistic likelihood of vindictiveness.”⁴⁵

³⁸ *Id.* at 370-71 (Powell, J., dissenting).

³⁹ *Id.* at 372 (Powell, J., dissenting).

⁴⁰ *United States v. Goodwin*, 457 U.S. 368, 370 (1982).

⁴¹ *Id.*

⁴² *Id.* at 372.

⁴³ *Id.* at 391 (Brennan, J., dissenting).

⁴⁴ *Id.* at 389-90 (citations and internal quotation marks omitted).

⁴⁵ *Id.* at 390 (Brennan, J., dissenting).

Thus, under Brennan's test, the analysis depended not on the prosecutor's *actual* motive, but, rather, on how the prosecutor's actions might realistically be viewed by a defendant.⁴⁶ Even though Brennan had joined Blackmun's dissent in *Bordenkircher*, his opinion in *Goodwin* actually went a step further, inasmuch as Blackmun would have given the prosecutor an opportunity to rebut a presumption of vindictiveness by showing an absence of actual vindictive intent. (Indeed, as if to underscore the point, Blackmun did not join Brennan's dissent in *Goodwin*, but wrote a concurring opinion.⁴⁷) In Part IV, I will discuss how Brennan's approach might be translated into the express threat setting.

2. *Majority Opinion: Powell's Triumph?*

The majority relied on *Bordenkircher* in rejecting Goodwin's claim, but, in doing so, confronted an important difficulty: while *Bordenkircher* turned on a perceived need to preserve the plea-inducement system, Goodwin's claim presented no direct challenge to the system. Justice Stevens's solution was to recharacterize *Bordenkircher*, not as a case about the benefits of plea inducements, but as a case about the benefits of broad prosecutorial charging discretion. That maneuver addressed one difficulty, but opened another: how to account for *Blackledge v. Perry*, the leading pre-*Bordenkircher* case on prosecutorial vindictiveness, in which the Court had rejected a prosecutor's enhancement of charges following a defendant's exercise of his right to a retrial.⁴⁸ In light of *Blackledge*, charging discretion was clearly not unlimited. In order to harmonize *Blackledge* with his take on *Bordenkircher*, Stevens borrowed a little from Blackmun's dissent in *Bordenkircher* and a lot from Powell's. Lending support to this view of *Goodwin*, Powell (alone among the four *Bordenkircher* dissenters) actually joined the majority opinion.

Stevens endorsed the constitutional principle on which *Blackledge* was based (an individual "may not be punished for exercising a protected statutory or constitutional right"⁴⁹), but saw the real difficulty as one of proof. How could the courts distinguish between "governmental action that is fully justified as a legitimate response to perceived criminal conduct" and "governmental action that

⁴⁶ *See id.* at 389 (arguing that *Blackledge* "focused upon the accused's apprehension of . . . retaliatory motivation").

⁴⁷ Blackmun disagreed with the majority's approach, but would have nonetheless affirmed the conviction on the basis that the prosecutor had "dispelled the appearance of vindictiveness" in the case. *Id.* at 386 (Blackmun, J., concurring in judgment).

⁴⁸ 417 U.S. 21, 30-31 (1974).

⁴⁹ *Goodwin*, 457 U.S. at 372.

is an impermissible response to noncriminal, protected activity”⁵⁰? This is where Stevens borrowed from Blackmun, specifically, by structuring the regulation of prosecutorial discretion through a system of presumptions and burden-shifting. Thus, *Blackledge*, in Stevens’s account, had established a rebuttable presumption of vindictiveness based on increased charges in its particular *post-trial* setting.⁵¹

In *Bordenkircher*, by contrast, “the Court for the first time considered an allegation of vindictiveness that arose in a *pretrial* setting,”⁵² and “made clear that the mere fact that a defendant refuses to plead guilty and forces the government to prove its case is insufficient to warrant a presumption that subsequent changes in the charging decision are unjustified.”⁵³ Stevens felt that prosecutors could generally be trusted to charge reasonably and without an improper motive to punish or deter the exercise of procedural rights:

There is good reason to be cautious before adopting an inflexible presumption of prosecutorial vindictiveness in a pretrial setting. In the course of preparing a case for trial, the prosecutor may uncover additional information that suggests a basis for further prosecution or he simply may come to realize that information possessed by the State has a broader significance. . . .

. . . A prosecutor should remain free before trial to exercise the broad discretion entrusted to him to determine the extent of the societal interest in prosecution. . . .⁵⁴

Stevens thus indicated that the key difference between *Blackledge* and *Bordenkircher* was not the plea-inducement context (as the dissenters asserted), but their post-trial versus pretrial settings. Vindictiveness might be presumed in at least some post-trial contexts, but it would never be presumed where charges were increased pretrial.

Stevens’s reluctance, like Powell’s, to presume that prosecutors act with improper motives differentiates his approach from Blackmun’s. (And the focus of all three on prosecutorial motivation, as opposed to defendant apprehension, differentiates their approaches from Brennan’s.) At the same time, there are real differences (albeit subtle and unacknowledged) between Stevens’s opinion for the Court in *Goodwin* and Stewart’s opinion for the Court in *Bordenkircher*. In

⁵⁰ *Id.*

⁵¹ *Id.* at 376.

⁵² *Id.* at 377 (emphasis added).

⁵³ *Id.* at 382-83.

⁵⁴ *Id.* at 381-82.

particular, *Bordenkircher* recognized no apparent limitation on the prosecutor's discretion in pressuring defendants to plead guilty; as long as the charge is supported by probable cause, the prosecutor can use it. (Indeed, as Powell had suggested, it is hard to imagine a much clearer case of actual vindictiveness than *Bordenkircher* itself, in which the prosecutor sought a life sentence for an \$88 crime and made quite clear that he did so as a result of Hayes' refusal to plead guilty to the lesser crime.) By contrast, *Goodwin* stated,

In declining to apply a presumption of vindictiveness, we of course do not foreclose the possibility that a defendant in an appropriate case might prove objectively that the prosecutor's charging decision was motivated by a desire to punish him for doing something that the law plainly allows him to do.⁵⁵

How might this improper motivation be proven in the absence of a presumption? According to *Goodwin*, *Bordenkircher* established that the prosecutor's stated intent to induce a guilty plea was not enough:

The fact that the prosecutor threatened the defendant did not prove that the action threatened was not permissible; the prosecutor's conduct did not establish that the additional charges were brought solely to "penalize" the defendant and could not be justified as a proper exercise of prosecutorial discretion.⁵⁶

Put differently, "mixed motives" were inadequate: "A charging decision does not levy an improper 'penalty' unless it results *solely* from the defendant's exercise of a protected legal right, rather than the prosecutor's normal assessment of the societal interest in prosecution."⁵⁷ Thus, the burden on the defendant was not to show an intent to induce a guilty plea, but to show an intent to induce a guilty plea *using charges that "could not be justified as a proper exercise of prosecutorial discretion."*⁵⁸ *Goodwin* thus contemplated that vindictiveness claims would be built around a showing that charges exceeded "the extent of the societal interest in prosecution."⁵⁹ This is, of course, precisely the inquiry that Powell advocated in his *Bordenkircher* dissent. Little wonder that he parted ways with his fellow *Bordenkircher* dissenters to join the *Goodwin* majority!

⁵⁵ *Id.* at 384.

⁵⁶ *Id.* at 380 n.12.

⁵⁷ *Id.* at 380 n.11 (emphasis added; citation omitted).

⁵⁸ *Id.* at 380 n.12 (emphasis added).

⁵⁹ *Id.* at 382.

C. Summary

While *Goodwin* may have muddied the waters a bit, it is still perfectly clear that a defendant who is penalized, like Paul Hayes, pursuant to a plea-inducement threat has little or no chance of prevailing on a vindictiveness claim. Nothing in *Goodwin* suggested that *Bordenkircher* was incorrectly decided on its facts, and few defendants will have facts nearly as compelling as those of Hayes himself. Moreover, to whatever extent that Powell's slightly more flexible approach ultimately prevailed, recall that Powell made clear that judicial interference in the plea-inducement process should occur "[o]nly in the most exceptional case."⁶⁰

Bordenkircher and *Goodwin* both relied on consequentialist reasoning. Neither, for instance, engaged in any textual exegesis or assessment of original intent. Rather, *Bordenkircher* worried about preserving the benefits offered by the plea-inducement system to defendants and the public at large, while *Goodwin* worried about preserving the prosecutor's "free[dom] before trial to exercise the broad discretion entrusted to him to determine the extent of the societal interest in prosecution."⁶¹ In light of their respective consequentialist concerns, both cases mandated a high degree of deference to prosecutorial charging decisions.

Bordenkircher, adopting a somewhat stronger form of deference, employed a bright-line rule: as long as the charge was supported by probable cause, there was no vindictiveness problem. *Goodwin*, by contrast, seemed less consistent with the ideal of rule-structure formalism, suggesting an inquiry (could the charges "be justified as a proper exercise of prosecutorial discretion?") that implied prudential, case-by-case balancing of interests. As we will see, the Court initially adopted, but then decisively rejected, just this sort of rule as it wrestled with the sentencing factor problem.

II. A PARALLEL DILEMMA: ELEMENTS VERSUS SENTENCING FACTORS

Prior to the 1980's, sentencing in the United States was largely discretionary: after a defendant was convicted of a crime, the judge was free to select a sentence anywhere within a wide statutory range prescribed for the crime.⁶² The two- to

⁶⁰ *Bordenkircher*, 434 U.S. at 372 (Powell, J. dissenting).

⁶¹ 457 U.S. at 382.

⁶² See Kate Stith & Steve Y. Koh, *The Politics of Sentencing Reform: The Legislative History of the Federal Sentencing Guidelines*, 28 WAKE FOREST L. REV. 223, 225-26 (1993) (describing traditional federal sentencing practices). I use the term "discretionary sentencing" throughout this Article in order to avoid confusion surrounding the more common term "indeterminate sentencing." See Steven L. Chanenson, *The Next Era of Sentencing Reform*, 54 EMORY L. J. 377, 381-86 (2005) (discussing terminological difficulties in this area). The latter term is sometimes applied broadly to any sentencing regime in which judges select sentences within wide ranges, and sometimes more

ten-year range Paul Hayes initially faced for false uttering in 1973 was emblematic of such a discretionary system. By contrast, the Kentucky Habitual Criminal Act, with its mandatory life sentence for certain recidivists, was aberrational; indeed, Kentucky actually softened the Act while Hayes' case was pending⁶³—too late to do Hayes any good, but indicative of the times.

Since the 1970's, however, American legislatures have adopted a host of mandatory minimum statutes, binding sentencing guidelines, and other presumptive sentencing schemes in order to curtail the discretion of sentencing judges.⁶⁴ In such regimes, the sentence is largely dictated by the presence of some fact or group of facts that go beyond what is necessary to establish the defendant's legal guilt. For instance, under the Kentucky Habitual Criminal Act, the presence of two prior felony convictions mandated a life sentence. In *some* mandatory sentencing regimes, the critical facts must be found by a jury beyond a reasonable doubt,⁶⁵ but in others the legislature contemplates judicial fact-finding using the preponderance-of-the-evidence standard.

Although the latter procedures may be more efficient and reliable, they have also spawned important constitutional objections. These objections parallel the

narrowly to apply only to systems in which a parole board determines the actual release date from prison. "Discretionary sentencing" in this Article corresponds to the broader understanding of "indeterminate sentencing." The term encompasses not only traditional unguided sentencing systems, but also systems with "advisory" (i.e., nonbinding) sentencing guidelines. See Kim S. Hunt & Michael Connelly, *Advisory Guidelines in the Post-Blakely Era*, 17 FED. SENT'ING RPTR. 233, 233 (2005) (noting that nine states and the District of Columbia employ advisory guidelines). By contrast, six states and the federal government had "mandatory" sentencing guidelines prior to the recent *Apprendi* line of cases. See *id.* at 239 n.3.

⁶³ *Bordenkircher*, 434 U.S. at 359 n.2.

⁶⁴ For a more detailed description of this history and explanation of the trend, see O'Hear, *supra* note 16, at 756-91. Presumptive sentencing schemes mandate a particular sentence or narrow sentencing range, often based on detailed sentencing guidelines that take into account a variety of offense and offender characteristics; judges must impose the presumptive sentence (or sentence within the presumptive range) unless particular aggravating or mitigating facts warrant a different result. Prior to the *Apprendi* decisions, thirteen states and the federal government employed presumptive sentencing. Don Stemen & Daniel F. Wilhelm, *Finding the Jury: State Legislative Responses to Blakely v. Washington*, 18 FED. SENT'ING RPTR. 7, 7 (2005). Presumptive sentencing schemes (with or without guidelines) may be contrasted with discretionary schemes. See *supra* note 62. Note one final distinction: if a legislature wishes to single out a particular type of offense or offender for special condemnation (in either a discretionary or presumptive system), the legislature might increase the applicable *maximum* sentence, in lieu of (or in addition to) increasing the *minimum*. Both sorts of increases will be referred to here as "sentence enhancements," although their operation differs in at least one important respect: an increased maximum enhances the scope of the sentencer's discretion, while an increased minimum diminishes discretion.

⁶⁵ *Bordenkircher*, 434 U.S. at 359.

issues raised by the use of charging threats to induce guilty pleas. On the one hand, both practices curtail the availability to defendants of basic jury trial protections in the determination of the facts on which their punishment depends. On the other hand, judicial regulation of these practices would run counter to traditional doctrines of deference to coordinate branches of government. At issue in *Bordenkircher* was the tradition of judicial deference to prosecutorial charging decisions. In the sentencing context, the issue was judicial deference to the legislative determination of crime elements.

The Court confronted these sentencing issues for the first time in 1986 in *McMillan v. Pennsylvania*.⁶⁶ Although *Apprendi* later repudiated much of *McMillan*, the earlier case remains a good starting point for this Part, both because it illuminates the significance of *Apprendi* and because, with the same line-up of Justices as in *Goodwin*, it also sheds light on the relationship between the vindictiveness cases and the sentencing cases. (The line-up of Justices is set forth in Table 1.) After considering *McMillan*, I summarize the later *Apprendi* cases and describe how exactly they are formalist and how consequentialist.

Table 1. Break-Down of Votes in *Goodwin*, *McMillan*, and *Apprendi*
(Current sitting Justices are indicated in bold; M=majority; D=dissent; C=concurrence in judgment only; N/A=not on Court).

	<i>Goodwin</i>	<i>McMillan</i>	<i>Apprendi</i>
Stevens	M	D	M
Burger	M	M	N/A
O'Connor	M	M	D
Powell	M	M	N/A
Rehnquist	M	M	D
White	M	M	N/A
Blackmun	C	D	N/A
Brennan	D	D	N/A
Marshall	D	D	N/A
Ginsburg	N/A	N/A	M
Scalia	N/A	N/A	M
Souter	N/A	N/A	M
Thomas	N/A	N/A	M
Breyer	N/A	N/A	D
Kennedy	N/A	N/A	D

⁶⁶ 477 U.S. 79 (1986).

A. *McMillan: Taking the Powell Approach Again*

In 1982, Pennsylvania adopted its Mandatory Minimum Sentencing Act, which provided that anyone who committed one of a list of serious felonies would be subject to a mandatory minimum prison sentence of five years if the sentencing judge found, by a preponderance of the evidence, that the person “visibly possessed a firearm during the commission of the offense.”⁶⁷ Dynel McMillan and other defendants challenged the Act’s constitutionality, arguing that it violated the requirement that all elements of a crime be proven beyond a reasonable doubt,⁶⁸ as mandated by *In re Winship*.⁶⁹ These claims were ultimately rejected by the Supreme Court.⁷⁰

Decided just four years after *Goodwin* by a Court that was comprised of the same nine justices, it should not be surprising that *McMillan* reflected a similar jurisprudential dynamic as the earlier case. Indeed, each of the five members of the *McMillan* majority was also part of the *Goodwin* majority, while both of the *Goodwin* dissenters also dissented in *McMillan*. Thus, only two justices “switched” in *McMillan*, and one of those (Blackmun) made just the small step from a concurrence in judgment to outright dissent. That leaves just one justice (Stevens) making a sharp break in alignment from *Goodwin* to *McMillan*. Fortunately, Stevens authored a lengthy dissent in the later case, which offers some clues as to why he “flipped.”

First, however, consider Justice Rehnquist’s majority opinion. Powell joined the opinion without comment, and it does indeed echo major themes from both his *Bordenkircher* dissent and the *Goodwin* majority opinion. Much as the earlier opinions emphasized judicial deference towards prosecutors, *McMillan* emphasized deference towards state legislatures.⁷¹ But just as the earlier opinions declined to make deference absolute, *McMillan* also indicated “there are constitutional limits to the State’s power in this regard.”⁷² *McMillan* sought a middle ground—just like Powell in *Bordenkircher*—between rejecting judicial

⁶⁷ *Id.* at 81.

⁶⁸ *Id.* at 82-83.

⁶⁹ 397 U.S. 358 (1970).

⁷⁰ *McMillan*, 477 U.S. at 82.

⁷¹ The Court put it this way: “[I]n determining what facts must be proved beyond a reasonable doubt the state legislature’s definition of the elements of the offense is usually dispositive [W]e should not lightly construe the Constitution so as intrude upon the administration of justice by the individual States.” *Id.* at 85.

⁷² *Id.* at 86.

review altogether and imposing rigid, formal constraints that might unduly burden state crime-control efforts.

The Court thus declined to articulate a “bright line test,”⁷³ but, instead, identified several potentially relevant considerations. Most notably, the Court suggested the tail-wagging-the-dog test, asking whether the sentencing factor (here, visible possession of a firearm) really dominated the sentencing calculus.⁷⁴ For instance, as to the Pennsylvania Act, because the predicate felonies all involved maximum sentences in excess of five years, the Act’s new five-year mandatory minimum did not amount to a tail wagging a dog: the Act “operates solely to limit the sentencing court’s discretion in selecting a penalty already available to it without the special finding of visible possession of a firearm.”⁷⁵ This *McMillan* test was reminiscent of Powell’s proposed test in *Bordenkircher*, which could be recast in similar terms: Was the state’s effort to induce a guilty plea (the tail) “wagging” the state’s legitimate interest in punishment (the dog)?

With the majority opinion echoing the earlier Powell dissent, Blackmun, Brennan, and Marshall reprised *their* dissenting position from *Bordenkircher*. Marshall, writing for the three, rejected the majority’s approach as overly deferential and basically endorsed the bright-line test proposed by Stevens.⁷⁶

Stevens, in his dissent, likewise rejected unlimited deference to the legislature. “It would demean the importance of the reasonable-doubt standard,” he wrote, “if the substance of the standard could be avoided by nothing more than a legislative declaration that prohibited conduct is not an ‘element’ of a crime.”⁷⁷ Fair enough, but no one in *McMillan* was advocating absolute deference. The real question was the test to be used when a defendant asserted that a legislature had gone too far. Stevens proposed the following: “[I]f a State provides that a specific component of a prohibited transaction shall give rise to a special stigma and a special punishment, that component must be treated as a ‘fact necessary to constitute the crime’ within the meaning of our holding in *Winship*.”⁷⁸ And Stevens had little difficulty concluding that the Pennsylvania

⁷³ *Id.* at 91.

⁷⁴ *See id.* at 88 (“The statute gives no impression of having been tailored to permit the visible possession finding to be a tail which wags the dog of the substantive offense.”).

⁷⁵ *Id.*

⁷⁶ *Id.* at 93-94 (Marshall, J., dissenting). Marshall declined to join Stevens’ opinion because he was not ready to commit to Stevens’ approach to mitigating facts. *Id.* at 94.

⁷⁷ *Id.* at 102 (Stevens, J., dissenting).

⁷⁸ *Id.* at 103.

statute did, indeed, impose “special stigma” and “special punishment” based on the “specific component” of visible possession.⁷⁹

Stevens’s approach might be criticized as not merely insufficiently deferential to the state legislature, but also futile. The problem lay in his *per se* distinction between aggravating and mitigating facts: while aggravating facts would always have to be proven beyond a reasonable doubt, no such standard applied to the disproof of mitigating.⁸⁰ The rule invited legislative evasion: a state that was intent on minimizing prosecutorial burdens might simply convert aggravating factors into mitigating. Stevens himself offered as an example

a statute making presence in any private or public place a felony punishable by up to five years imprisonment and yet allowing an affirmative defense for the defendant to prove, to a preponderance of the evidence, that he was not robbing a bank.⁸¹

This hypothetical echoed the prosecutorial evasion scenario that so troubled the Court in *Bordenkircher* and *Goodwin* (i.e., adopting an aggressive high-to-low charging strategy).

Stevens dismissed the evasion possibilities in *this* context, however, as unlikely in a democratic system. “No democratically elected legislature would enact [the hypothetical] law, and if it did, a broad-based coalition of bankers and bank customers would soon see the legislation repealed.”⁸² The aggravating/mitigating distinction was thus explained by reference to the view that “constitutional law appropriately exists for those situations where representative government cannot be trusted, not those where we know it can.”⁸³

B. The Apprendi Canon: Stevens’s Triumph

Beginning with *Jones v. United States*⁸⁴ in 1999, a majority of the Court signaled first its doubts about, and then its rejection of, the approach taken in *McMillan*. Five cases constitute what we may think of as the “*Apprendi* canon”: *Jones*, *Apprendi*, *Ring v. Arizona*,⁸⁵ *Blakely v. Washington*,⁸⁶ and *United States v.*

⁷⁹ *Id.*

⁸⁰ *Id.* at 99.

⁸¹ *Id.* at 100.

⁸² *McMillan*, 477 U.S. at 101 (Stevens, J., dissenting).

⁸³ *Id.* at 101 n.5 (quoting JOHN HART ELY, *DEMOCRACY AND DISTRUST* 183 (1980)).

⁸⁴ 526 U.S. 227 (1999).

⁸⁵ 536 U.S. 584 (2002).

⁸⁶ 542 U.S. 296 (2004).

Booker.⁸⁷ In these cases—all but one vigorously contested 5-4 decisions—the Court adopted and extended the rule proposed by Justice Stevens in his *McMillan* dissent. Indeed, Stevens himself authored the majority opinions in *Apprendi* and *Booker*. He was joined by Justices Ginsburg, Scalia, Souter, and Thomas.

Technically, *Jones* presented a question of statutory, not constitutional, interpretation. The federal carjacking statute at issue in the case provided a standard 15-year maximum sentence, but an enhanced 25-year maximum sentence if there had been “serious bodily injury.”⁸⁸ The Court determined that this “serious bodily injury” prong of the statute should be regarded as an element of a separate, aggravated carjacking offense—thus triggering rights to jury fact-finding beyond a reasonable doubt—rather than merely an additional factor for the judge to find at sentencing.⁸⁹ The Court reached this interpretation, in part, because a contrary result would raise “grave and doubtful constitutional questions,” specifically, by increasing the defendant’s sentencing exposure on the basis of a fact not proven to the jury beyond a reasonable doubt.⁹⁰ As the dissenters recognized, the majority here embraced a more robust conception of jury-trial rights than had been apparent in the Court’s recent jurisprudence.⁹¹

What was merely suggested as a constitutional question in *Jones* became a constitutional holding the following year in *Apprendi*. Charles C. Apprendi, Jr., pled guilty in New Jersey state court to three weapons violations.⁹² At sentencing, the judge determined that Apprendi had acted with a racially biased purpose, which, under the state hate crimes statute, increased the maximum possible sentence from twenty to thirty years.⁹³ In ruling that this process had violated Apprendi’s constitutional rights, the Court limited the *McMillan* wag-the-dog analysis to sentence enhancements at the bottom end of the sentencing range, i.e., fact-finding that triggers a mandatory minimum sentence.⁹⁴ As to enhancements at the top end of the range, the Court adopted this bright-line rule: “Other than the fact of a prior conviction, any fact that increases the penalty for a crime beyond the prescribed statutory maximum must be submitted to a jury, and proved beyond a reasonable doubt.”⁹⁵ The Court thus held that New Jersey’s

⁸⁷ 543 U.S. 220 (2005).

⁸⁸ 526 U.S. at 230.

⁸⁹ *Id.* at 232.

⁹⁰ *Id.* at 239.

⁹¹ *Id.* at 272 (Kennedy, J., dissenting).

⁹² 530 U.S. at 469-70.

⁹³ *Id.* at 470.

⁹⁴ *Id.* at 487 n.13.

⁹⁵ *Id.* at 490.

sentence enhancement procedure was “an unacceptable departure from the jury tradition that is an indispensable part of our criminal justice system.”⁹⁶

In *Ring*, the Court extended *Apprendi* to capital punishment, overturning a sentencing scheme in which a judge found the aggravating factors required to make a murder defendant eligible for death.⁹⁷ The Court saw no good reason to distinguish *Apprendi*; if the jury-trial right encompassed the fact-finding necessary to increase a defendant’s sentencing exposure by ten years, then it must also encompass the fact-finding necessary to impose the death penalty.⁹⁸

In *Blakely*, the Court considered *Apprendi*’s applicability to sentencing guidelines.⁹⁹ Sentencing guidelines prescribe a narrow sentencing range, within a broader statutory range, based on the presence or absence of specified factors. For instance, when Ralph Howard Blakely, Jr., pled guilty in Washington state court to second-degree kidnapping, he faced a broad statutory range of zero to ten years.¹⁰⁰ Pursuant to state sentencing guidelines, however, Blakely’s offense triggered a “standard range” of 49 to 53 months,¹⁰¹ with the possibility of an above-range sentence if the judge found “substantial and compelling reasons justifying an exceptional sentence.”¹⁰² The judge made such a finding on the basis of his determination that Blakely acted with “deliberate cruelty,” and imposed a sentence of 90 months¹⁰³—well above the high end of the standard range, but still well below the ten-year statutory maximum.

The Court nonetheless held that this process violated *Apprendi*, refining the *Apprendi* rule as follows: “When a judge inflicts punishment that the jury’s verdict alone does not allow, the jury has not found all the facts which the law makes essential to the punishment, and the judge exceeds his proper authority.”¹⁰⁴ Thus, because the judge in *Blakely* had imposed a sentence that would not have been legally permissible absent *his* finding of deliberate cruelty, the judge had violated the defendant’s right to a jury trial.¹⁰⁵

⁹⁶ *Id.* at 497.

⁹⁷ 536 U.S. 584, 589 (2002).

⁹⁸ *Id.* at 609.

⁹⁹ 542 U.S. 296, 301 (2004).

¹⁰⁰ *Id.* at 299.

¹⁰¹ *Id.* at 300.

¹⁰² *Id.* at 299.

¹⁰³ *Id.* at 300.

¹⁰⁴ *Id.* at 303-04.

¹⁰⁵ *Id.* at 305.

In overturning Washington’s sentencing regime, *Blakely* also cast into doubt the mandatory guidelines systems employed in a number of other jurisdictions, including, perhaps most importantly, the federal sentencing guidelines. In *United States v. Booker*, the Court finally addressed the federal system, and held that it, too, ran afoul of the *Apprendi* rule to the extent that it relied on judicial fact-finding.¹⁰⁶ Notably, however, the *Apprendi* Five fractured over the remedy for this constitutional violation. Four of the five would have retained the mandatory character of the federal guidelines but required jury fact-finding for sentence enhancements.¹⁰⁷ Ginsburg switched sides, though, to create a separate majority in favor of a different remedy: conversion of the guidelines from mandatory to advisory.¹⁰⁸ *Booker* thus boasted two majority opinions, the “merits opinion” (joined by all members of the *Apprendi* Five) and the “remedy opinion” (authored by Justice Breyer and joined by Ginsburg and the merits dissenters).

The *Booker* remedy has not been the only occasion on which the Five have parted company. Of greatest significance, Scalia joined the *Apprendi* dissenters in *Harris v. United States* to exempt mandatory minimums from the *Apprendi* rule.¹⁰⁹ Thus, a judge may still find the facts necessary to trigger a mandatory minimum (as long as the minimum does not exceed the otherwise applicable maximum). In effect, then, the *Apprendi* rule deals just with the defendant’s worst-case scenario: the rule is violated only if a judge’s fact-finding results in an increase in the *maximum* sentence that may be imposed.

C. *Apprendi’s Aims*

In order to appreciate the inconsistencies between *Bordenkircher* and the *Apprendi* cases, we should begin by considering what exactly the *Apprendi* Five said they were up to: what vision purports to animate the extraordinary doctrinal changes they have embraced? (To be sure, critics contend that the Five cannot

¹⁰⁶ 543 U.S. 220, 226-27 (2005).

¹⁰⁷ *Id.* at 272-73 (Stevens, J., dissenting in part); *id.* at 313 (Thomas, J., dissenting in part).

¹⁰⁸ *Id.* at 245. It remains arguable whether the formal switch to “advisory” guidelines has rendered the federal system materially more discretionary, as appellate courts, exercising “reasonableness review” over post-*Booker* sentences, have proven quite reluctant to approve some of the more adventuresome departures from guidelines norms. See, e.g., *United States v. Clark*, 434 F.3d 684 (4th Cir. 2006) (overturning sentence where district court took into account federal-state sentencing disparity in imposing sentence below guidelines range); *United States v. Eura*, 440 F.3d 625 (4th Cir. 2006) (overturning sentence where district court attempted to mitigate the disparate treatment of crack and powder cocaine in guidelines); *United States v. Galicia-Cardenas*, 443 F.3d 553 (7th Cir. 2006) (overturning sentence where district court adjusted sentence so as to take into account lack of “fast-track” program in prosecuting United States Attorney’s Office).

¹⁰⁹ 536 U.S. 545, 551-52 (2002).

really be serious about the vision they have articulated because there is a seemingly wide disconnect between that vision and the reality of the *Apprendi* rule they adopted. I will address this criticism in the next Section.)

The *Apprendi* decisions represent a self-conscious choice in favor of the “common-law ideal” of jury trial over the “civil-law ideal of administrative perfection.”¹¹⁰ The decisions thus rejected a model of bureaucratized criminal justice, embodied by judicially administered sentencing guidelines, despite the model’s conceded advantages of efficiency and consistency.¹¹¹ In lieu of these values, the *Apprendi* Five prioritized the following: (1) democratic control over the criminal justice system, (2) libertarian checks on state power, and (3) fairness to defendants.¹¹² Let us consider each of these values in turn.

1. *Democratic Control*

The *Apprendi* cases linked the jury to democratic values: “Just as suffrage ensures the people’s ultimate control in the legislative and executive branches, jury trial is meant to ensure their control in the judiciary.”¹¹³ At the same time, it might seem a bit odd for an (unelected) Supreme Court to strike down a host of legislatively adopted sentencing schemes in the name of democratic values. Why not permit a politically accountable legislature to employ the most efficient means (i.e., judicial fact-finding) to implement its choices about how much weight to give to particular sentencing factors?

The *Apprendi* decisions suggest two responses. First, legislative decisions regarding sentencing factors are necessarily somewhat crude; a legislature cannot hope to design a sentencing system that takes into account all of the complexity of the real world. As a result, the literal application of any sentencing rule will inevitably produce some unduly harsh outcomes that are not truly consistent with public preferences. Moreover, this disconnect between global legislative judgments and case-specific community preferences is likely exacerbated by the

¹¹⁰ *Id.* at 313.

¹¹¹ *See, e.g., Ring*, 536 U.S. at 607 (“Entrusting to a judge the finding of facts necessary to support a death sentence might be an admirably fair and efficient scheme of criminal justice designed for a society that is prepared to leave criminal justice to the State.”); *Booker*, 543 U.S. at 244 (acknowledging that jury fact-finding “may impair the most expedient and efficient sentencing of defendants”).

¹¹² Jury fact-finding is also sometimes said to advance accuracy in adjudication. I do not emphasize the point here because it does not play an important role in the *Apprendi* decisions. The accuracy hypothesis, moreover, has at best weak support in the empirical scholarship. Paul H. Robinson & Barbara A. Spellman, *Sentencing Decisions: Matching the Decisionmaker to the Decision Nature*, 105 COLUM. L. REV. 1124, 1145 (2005).

¹¹³ *Blakely*, 542 U.S. at 306.

tendency, demonstrated in the social science literature, for people to make harsher judgments about crime in the abstract than when confronted with the facts of a particular case.¹¹⁴ Preserving a role for the jury at sentencing thus offers a second level of democratic control that may mitigate the structural weaknesses of the first. Indeed, both *Jones* and *Apprendi* made note of the common-law tradition of “pious perjury,” whereby a jury might circumvent a harsh mandatory penalty that seemed disproportionate to the gravity of the offense.¹¹⁵ In a similar vein, a modern sentencing jury might decline to find an aggravating sentencing factor that was literally present in a case, but that would result in a sentence enhancement that, in the jury’s view, was unjust.¹¹⁶

Second, a sentencing jury might also advance democratic values in cases in which the legislature has structured a sentencing factor in a manner that requires a discretionary exercise of judgment. Sentence enhancements are sometimes structured as bright-line rules; think, for instance, of the mandatory minimum in *McMillan* that was triggered by the visible possession of a firearm. But sometimes legislatures employ more open-ended standards; think here of the system overturned in *Blakely*, in which the sentence might be enhanced if the judge found “substantial and compelling reasons justifying an exceptional sentence.” An enhancement of this nature effectively delegates considerable discretionary authority to the sentencer. In such circumstances, where the legislature has left much undecided, we might find it particularly important that implementation of the standard include the participation of another actor with

¹¹⁴ Jenia Iontcheva Turner, *Implementing Blakely*, 17 FED. SENT’ING RPTR. 106, 111 (2004).

¹¹⁵ *Jones*, 526 U.S. at 245; *Apprendi*, 530 U.S. at 479 n.5.

¹¹⁶ As discussed in the next Section, a sentencing *judge* might also function as a check on a prosecutor’s overly aggressive use (relative to public preferences) of sentence enhancements. It is at least plausible, however, that a judge, schooled in rule-of-law norms and subject to appellate review, would be less likely than a jury to function as a check in the circumstances contemplated here, i.e., *de facto* nullification of the enhancement. This view resonates with *Ring*’s characterization of the original intent of the jury-trial right: “If the defendant preferred the common-sense judgment of the of a jury to the more tutored but perhaps less sympathetic reaction of the single judge, he was to have it.” 536 U.S. at 609.

Moreover, there are reasons to view judicial nullification as more problematic than jury nullification from the standpoint of democratic legitimacy. To be sure, in some jurisdictions (though not all), judges are elected and thus also have a measure of democratic legitimacy. At the same time, one can readily discern a number of reasons why the democratic credentials of even an elected judge might be viewed as unsatisfactory: judges are drawn from only a very small subset of the community’s population (i.e., those with law degrees); elected judges are often intentionally shielded from the same level of democratic accountability as elected legislators (e.g., through longer terms in office, retention votes, and other special nonpartisan election processes); and, in the great run of routine cases, judges operate without any meaningful public scrutiny.

democratic legitimacy. And, as *Jones* pointed out, the common-law tradition is that the jury is not merely a finder of facts, but also an applier of law to facts.¹¹⁷ Thus, the *Apprendi* decisions do not necessarily contemplate that the jury's role in sentencing will be limited to fact-finding, in the narrowest sense of the term, but may also extend to determining the applicability of more subjectively-defined sentence enhancements.¹¹⁸ On this view, *Apprendi* advances democratic values by channeling the discretionary authority that is subtly embedded in mandatory sentencing regimes away from judges and towards juries.

Note the complementary nature of these two considerations. The more particularized and rule-like the sentencing regime, the greater will be the need for an equitable mechanism, like the jury's "pious perjury," in order to address unanticipated instances of undue severity. As the sentencing regime grows more flexible and standard-like, the need for this form of law-correction will diminish, but, with an increase in the sentencer's discretionary authority, concerns may grow as to the sentencer's democratic legitimacy. Either way, the jury may have a useful role to play alongside the legislature in helping to ensure that sentencing outcomes conform to public preferences.¹¹⁹

¹¹⁷ *Jones*, 526 U.S. at 247.

¹¹⁸ See Douglas A. Berman, *Conceptualizing Blakely*, 17 FED. SENT'ING RPTR. 89, 92 (2004) (discussing uncertainty as to whether determinations involving "value judgments," not just findings of historical fact, trigger jury trial right under *Blakely*). For an argument that juries are better suited than judges to determine blameworthiness for sentencing purposes, see Robinson & Spellman, *supra* note 112, at 1146-47.

¹¹⁹ In addition to the reasons suggested by the *Apprendi* decisions, there is at least one other reason to regard the sentencing jury as an institution that might further democratic values: the jury brings to bear *local* views, as against the preferences of a distant legislature that may represent a host of communities with quite different views. The federal system offers particularly dramatic illustrations of the potential disconnects between the criminal justice preferences of a local community and a more encompassing polity, *see, e.g.*, Michael M. O'Hear, *National Uniformity/Local Uniformity*, 87 IOWA L. REV. 721, 731 (2002) ("Perhaps most striking are the federal death penalty cases in states that do not authorize capital punishment."), but the same issue can also arise at the state level, *see, e.g.*, Michael M. O'Hear, *Federalism and Drug Control*, 57 VAND L. REV. 783, 860-61 (2004) (discussing wide variations in county-level support for California's Proposition 36, which requires diversion of drug offenders from incarceration to treatment). For an argument that local preferences should generally prevail, at least with respect to crime that occurs on a local scale, *see* Michael M. O'Hear, *Localization and Transparency in Sentencing: Reflections on the New Early Disposition Departure*, 27 HAMLINE L. REV. 358, 360-63 (2004). More controversially, localization through the jury may also provide a vehicle for predominantly African-American communities to address racial disparities in the operation of the criminal justice system. *See* Paul Butler, *Racially Based Jury Nullification: Black Power in the Criminal Justice System*, 105 YALE L. J. 677, 679 (1995). For a leading critique of Professor

2. *Containing State Power*

The *Apprendi* decisions repeatedly claimed to protect individual liberty and contain state power.¹²⁰ How so? One possibility, of course, is that democratic control works to restrict state power. That would imply, in the present context, a tendency for juries to sentence more leniently than judges, thereby diminishing the coercive power of the state. *Ring*, for instance, suggested that defendants might prefer sentencing juries to “less sympathetic” judges.¹²¹ This is doubtlessly true in some cases, but democratic control does not guarantee generous outcomes. Indeed, some empirical studies suggest that jury sentencing is often harsher than judicial sentencing.¹²² Thus, enhanced democratic control, in and of itself, would not seem the most effective way to limit state power.¹²³

But jury sentencing may serve to contain state power in another respect: as a sort of procedural tax on prosecutions, that is, a burden on limited state law enforcement resources that might diminish the number and intensity of prosecutions and give defendants greater leverage in plea negotiations. Jury sentencing might be implemented in one of two ways, either as a proceeding that is distinct from and subsequent to the determination of guilt, or as a connected adjunct to the basic criminal trial. If implemented through the former

Butler’s proposal, see Andrew D. Leipold, *The Dangers of Race-Based Jury Nullification*, 44 U.C.L.A. L. REV. 109 (1996).

Even if one does not accept the view that juries are an effective medium for expressing public preferences, providing a more robust role for the jury may nonetheless serve the parallel end of enhancing public *perception* of democratic control. Robinson & Spellman, *supra* note 112, at 1148. In other circumstances, the Court has recognized the importance of such perceptions. See, e.g., *Nixon v. Shrink Missouri Government PAC*, 528 U.S. 377 (2000) (“[W]e spoke in *Buckley* of the perception of corruption inherent in a regime of large individual financial contributions, as a source of concern almost equal to quid pro quo improbity.” (citation and internal quotation marks omitted)).

¹²⁰ See, e.g., *Jones*, 526 U.S. at 246 (referring to jury as “the grand bulwark” of liberty); *Apprendi*, 530 U.S. at 477 (characterizing purpose of right to jury trial as “guard[ing] against a spirit of oppression and tyranny on the part of rulers”); *Ring*, 536 U.S. at 607 (“The founders of the American republic were not prepared to leave [criminal justice] to the State, which is why the jury-trial guarantee was one of the least controversial provisions of the Bill of Rights. It has never been efficient; but it has always been free.”).

¹²¹ *Ring*, 536 U.S. at 609.

¹²² See, e.g., Nancy J. King & Rosevelt L. Noble, *Jury Sentencing in Practice: A Three-State Study*, 57 VAND. L. REV. 883, 898 (2004) (discussing jury sentencing in Kentucky).

¹²³ The goal of limiting state power is easier to square with a jury sentencing right to the extent that defendants can unilaterally choose the finder of fact, selecting judge or jury according to which is expected to be more sympathetic. Some jurisdictions, however, require jury-trial waivers from both defendant and prosecutor. See *Singer v. United States*, 380 U.S. 24, 34-36 (1965) (upholding constitutionality of federal rule to this effect).

mechanism, the procedural costs (relative to judicial sentencing) are obvious: a jury must be assembled,¹²⁴ managed, instructed, and argued to, all with the enhanced formality that is attendant to jury, as opposed to bench, trials.¹²⁵ If jury sentencing is simply folded into the trial on the merits, the whole process will be far more efficient, but this may not be appealing or practically feasible in a complex sentencing guidelines system, where sentencing may require fact-finding on a dozen or more discrete factors beyond the elements of the offense.¹²⁶ Moreover, even if this latter approach is selected, there is still the problem of guilty plea cases, in which the *Apprendi* decisions may necessitate the impaneling of a sentencing jury where no jury would otherwise be required.¹²⁷

The procedural tax theory squares with how the *Apprendi* decisions described what they were doing. There can be no question but that the decisions self-consciously rejected efficiency in favor of limited state power. As *Ring* put it, “[The jury-trial guarantee] has never been efficient; but it has always been free.”¹²⁸ The procedural tax theory would go just one step further: “*Because* the jury trial has never been efficient, it has always been free.”

The procedural tax theory also squares with the Court’s pairing of the jury-trial right with the right to fact-finding beyond a reasonable doubt, which also protects individual liberty by imposing increased procedural burdens on prosecutors. The rule in *Apprendi*, for instance, expressly linked the two rights: “[A]ny fact that increases the penalty for a crime beyond the prescribed statutory

¹²⁴ In cases in which the defendant was convicted by a jury, the same jury might be reconvened for sentencing, which would save the not inconsiderable costs associated with jury selection. Most cases, however, are resolved by guilty plea; in such cases, jury sentencing might necessitate the selection of a jury that otherwise would not have been required.

¹²⁵ See *Goodwin*, 457 U.S. at 390 (Brennan, J., dissenting) (detailing reasons jury trial entails “far more prosecutorial work” than bench trial); Turner, *supra* note 114, at 109 (discussing procedural formalities that might have to be observed in sentencing proceedings under *Blakely*). These costs may vary by jurisdiction, though, depending on how complex the jurisdiction’s sentencing scheme is and how the jury-trial rights are implemented. See *id.* at 110 (noting relatively low costs of implementing *Apprendi* in Kansas).

¹²⁶ See Turner, *supra* note 114, at 108 (arguing that, “in many cases, bifurcation is indispensable to ensuring a fair trial” and noting that Kansas, when “*Blakely*-izing” its sentencing guidelines, “let judges determine on a case-by-case basis when bifurcation would be in the interest of justice”).

¹²⁷ To be sure, a plea agreement *might* include a waiver of jury rights as to sentencing; however, not all guilty pleas are rendered pursuant to an agreement, and (depending on the priorities and relative bargaining leverage of the parties) not all agreements need necessarily include a waiver of such rights. But see *id.* at 110 (noting “double-waiver” requirement in some jurisdictions, which requires that sentencing jury right be waived if trial jury right is waived).

¹²⁸ 536 U.S. at 607.

maximum must be submitted to a jury, *and proved beyond a reasonable doubt.*¹²⁹

3. *Fairness to Defendants*

Finally, the *Apprendi* decisions also purport to enhance fairness to defendants. Fairness, here, is used in the sense suggested by some BLE scholars: consistency of outcomes with even imperfectly rational expectations.¹³⁰ Fairness to defendants would thus imply that there are no nasty surprises at the end of the sentencing process; to the extent that the system fosters sentencing expectations, the system should not disappoint those expectations. *Blakely*, in particular, emphasized these considerations:

Any evaluation of *Apprendi*'s "fairness" to criminal defendants must compare it with the regime it replaced, in which a defendant, with no warning in either his indictment or plea, would routinely see his maximum potential sentence balloon from as little as five years to as much as life imprisonment, based not on facts proved to his peers beyond a reasonable doubt, but on facts extracted after trial from a report compiled by a probation officer who the judge thinks more likely got it right than wrong.¹³¹

The *Apprendi* decisions are thus presented as an effort to save defendants from nasty surprises delivered "without warning" at sentencing.

D. Is It Really Just "Doctrinaire Formalism"?

The *Apprendi* Five may claim that they are advancing democratic, libertarian, and fairness values, but are they really? Justice O'Connor, characterizing their handiwork as "doctrinaire formalism," thinks otherwise.¹³² Her contention, along with that of the other *Apprendi* dissenters, is that the *Apprendi* rule will be evaded, such that there will be no step *forward* as to the basic jury-trial values, but, if anything, a step *back*. The dissenters have focused on three potential forms of evasion: (1) legislative inversion of sentence enhancements (i.e., increasing standard sentences and converting aggravating factors into mitigating), (2) a return to discretionary sentencing, and (3)

¹²⁹ 530 U.S. at 466 (emphasis added).

¹³⁰ *Infra* Part II.D.2.

¹³¹ 542 U.S. at 311-12.

¹³² *Id.* at 321 (O'Connor, J., dissenting). O'Connor doubtlessly has in mind here the sort of rigid formalism that dominated American legal thinking a century ago and that was the subject of relentless attacks by the Legal Realists. See Erik Luna, *Gridland: An Allegorical Critique of Federal Sentencing*, 96 J. CRIM. L. & CRIMINOLOGY 25, 89-93 (2005).

prosecutorial use of plea inducements to force defendants to surrender their jury rights at sentencing. (A fourth type of evasion, expanded use of mandatory minimums, is discussed separately in the Conclusion.¹³³) These evasion arguments, which mirror analogous concerns raised by the *Bordenkircher* and *Goodwin* majorities, are discussed in turn below. I will show that there are at least plausible responses to each of these concerns, and that accordingly the *Apprendi* decisions cannot fairly be accused of indifference to consequences. Before reaching the evasion arguments, however, this Section begins with two prefatory considerations: a description of the *real* formalism of the *Apprendi* decisions and a summary of key principles of BLE that will aid our understanding of *Apprendi*'s consequentialism.

1. *Apprendi*'s Formalism

There is at least one sense in which the *Apprendi* decisions are unabashedly formalist: rule structure. Recall that the *McMillan* test, which was largely displaced by *Apprendi* and its progeny, employed a highly indeterminate standard, i.e., whether the tail (the sentencing factor) was wagging the dog (the underlying offense). *Blakely*, in particular, heaped scorn on the test for its indeterminacy.¹³⁴ The *Apprendi* bright-line rule was preferable for the more robust protection it provided for jury rights; it would be perverse, *Blakely* suggested, to adopt a discretionary test that effectively endowed judges with the authority to decide the scope of their own power relative to juries.¹³⁵

¹³³ The form of evasion stands on a different footing than the other three because the *Apprendi* Five have no unified position on it; four would have closed the "loophole" in *Harris*, while Justice Scalia (without explanation) provided the key fifth vote to keep it open. *Supra* Part II.B. Thus, discussion of this form of evasion would not contribute much to the immediate objective of teasing out the shared values and assumptions of the Five.

¹³⁴ 542 U.S. at 312 n.13.

¹³⁵ 542 U.S. at 308 ("Whether the Sixth Amendment incorporates this manipulable [*McMillan*] standard rather than *Apprendi*'s bright-line rule depends on the plausibility of the claim that the Framers would have left definition of the scope of jury power up to judges' intuitive sense of how far [in the legislative creation of new sentencing factors] is *too far*. We think that claim not plausible at all . . ."). A few state courts have recently reached the conclusion that *Booker* represented an implicit retreat from the bright-line rule of *Apprendi* and *Blakely*. See Jonathan D. Soglin & J. Bradley O'Connell, *Blakely, Booker, & Black: Beyond the Bright Line*, 18 FED. SENT'ING RPTR. 46, 46 (discussing decisions by courts in California, Tennessee, and New Mexico). The Supreme Court has agreed to hear one of these cases, *Cunningham v. California*, with oral arguments scheduled in the fall of 2006. Michael M. O'Hear, *Cunningham: Why Federal Practitioners Should Pay Attention*, 18 FED. SENT'ING RPTR. ___ (2006). While the recent addition of two new Justices may change the Court's dynamic on *Apprendi* issues, a retreat from the bright-

Adopting a formalist test, however, is not the same thing as “doctrinaire formalism,” which implies indifference to consequences. Indeed, as *Blakely* suggested, one might adopt a bright-line rule specifically in order to produce more meaningful consequences.¹³⁶ Underscoring this point, the *Apprendi* decisions repeatedly rejected arguments that the constitutional analysis should turn on statutory labels.¹³⁷ As O’Connor herself recognized, these rejected tests were no less “bright-line” than the *Apprendi* rule.¹³⁸ What drew the *Apprendi* Five to their rule was not merely its formalist character, but also its anticipated consequences.¹³⁹

2. *Key Principles of Behavioral Law and Economics*

A central insight of BLE is that “people evaluate outcomes based on the change they represent from an initial reference point, rather than based on the nature of the outcome itself.”¹⁴⁰ This basic insight informs four salient, overlapping principles, each of which has been demonstrated in a host of empirical studies. First, people tend to be *loss-averse*, i.e., “they weigh losses more heavily than gains of equal magnitude.”¹⁴¹ Second, while loss-avoidance may be preferred to gains, the concepts of “gain” and “loss” are malleable.¹⁴² Thus, the *framing* of a transaction as a gain or a loss may play an important role in determining whether or not it is viewed as desirable. Third, the starting point in a transaction tends to condition the outcome, even if the starting point is

line rule approach—which was so ringingly endorsed in *Blakely* by five Justices who remain on the Roberts Court—would be surprising.

¹³⁶ See 542 U.S. at 308 (rejecting indeterminate standard as unlikely to prevent erosion of jury’s power).

¹³⁷ See, e.g., *Apprendi*, 530 U.S. at 494 (“The relevant inquiry is not one of form, but of effect”); *Ring*, 536 U.S. at 604 (“If Arizona prevailed on its . . . argument, *Apprendi* would be reduced to a meaningless and formalistic rule of statutory drafting.”).

¹³⁸ *Blakely*, 542 U.S. at 321 (O’Connor, J., dissenting).

¹³⁹ To be sure, the *Apprendi* decisions might be defended on the basis of other formalist values. See King & Klein, *supra* note 15, at 1485 (arguing that *Apprendi* “maintains fidelity to historical practice and prior decisions”). But see Bibas, *supra* note 15, at 196 (arguing that relevant history is unclear). The point here is not that the *Apprendi* decisions would have come out differently if the Five had been oblivious to consequences, but, rather, that the decisions do plausibly purport to attend to consequences. Indeed, despite frequent citations in the decisions to historical sources, I think Professor Bibas has it right when he argues that originalism was not the “driving force” behind the decisions. *Id.* at 201.

¹⁴⁰ Jolls, et al., *supra* note 17, at 1535.

¹⁴¹ Stephanos Bibas, *Plea Bargaining Outside the Shadow of Trial*, 117 HARV. L. REV. 2463, 2508 (2004).

¹⁴² *Id.* at 2512.

random or irrelevant.¹⁴³ An initial reference effectively becomes an *anchor*, and “people usually do not adjust away from their anchors enough.”¹⁴⁴ Thus, for instance, one study shows that “the asking price of a house strongly influences appraisals of its value, even for experts who consider the asking price completely uninformative and who have plenty of other information.”¹⁴⁵

Fourth, the losses that people try to avoid include not only material losses, but also losses to reputation and self-image.¹⁴⁶ As a result of these considerations, human behavior often conforms to the principle of *reciprocal fairness*: people are willing to sacrifice their own material well-being to help those who are acting fairly, but are also willing to sacrifice so as to punish those who are acting unfairly.¹⁴⁷ (Fairness, as suggested above, can be understood by reference to deviations from expectations; for instance, consumers will perceive as unfair a firm that takes advantage of the short-term scarcity of a good by increasing established prices.¹⁴⁸) People, in other words, like to have a reputation for, and a self-image of, decency; but they also seek to avoid the appearance of being a dupe or a doormat.¹⁴⁹ These tendencies can be observed empirically in experimental variations on the classic Prisoner’s Dilemma.¹⁵⁰

3. *The Evasion Arguments*

Having now identified some of the ways in which actual human behavior might differ from the model of *homo economicus*, let us consider how these principles might inform the debate over *Apprendi*’s consequentialism. This debate largely revolves around three potential strategies for evading *Apprendi*.

a. Legislative Inversion

As Justice O’Connor observed, the *Apprendi* rule could, in principle, be defeated by a seemingly straightforward legislative reform: standard penalties could be increased, with maximum sentences then reduced based on the *absence*

¹⁴³ *Id.* at 2515.

¹⁴⁴ *Id.* at 2516.

¹⁴⁵ *Id.* (citing Gregory B. Northcraft & Margaret A. Neale, *Experts, Amateurs, and Real Estate: An Anchoring-and-Adjustment Perspective on Property Pricing Decisions*, 39 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 84 (1987))

¹⁴⁶ Jolls et al., *supra* note 17, at 1494.

¹⁴⁷ *Id.* at 1494 (citing Matthew Rabin, *Incorporating Fairness Into Game Theory and Economics*, 83 AM. ECON. REV. 1281 (1993)).

¹⁴⁸ *Id.* at 1511-12.

¹⁴⁹ *Id.* at 1495.

¹⁵⁰ *Id.* at 1494.

of specified aggravating circumstances.¹⁵¹ In this inversion process, sentence “aggravators” would effectively be converted into “mitigators,” but, at least in principle, the same sentencing outcomes should be obtained—and without a need ever to impanel a sentencing jury. (Recall that the *Apprendi* rule only requires a jury when the maximum sentence is *increased* beyond the standard range for the offense.) If legislatures were, in fact, to respond to *Apprendi* by systematically inverting their sentencing regimes, then O’Connor’s charge of doctrinaire formalism might have some appeal: the inverted regimes, still relying on judicial fact-finding, would do nothing to advance *Apprendi*’s purported objectives of democratic control and limited state power.¹⁵²

In *Apprendi*, however, Justice Stevens repeated the response he made to this argument in *McMillan*, characterizing the likelihood of inversion as “remote.”¹⁵³ He relied on “democratic constraints” that would “discourage legislatures from enacting penal statutes that exposed *every* defendant convicted of, for example, weapons possession, to a maximum sentence exceeding that which is, in the legislature’s judgment, generally proportional to the crime.”¹⁵⁴ Stevens, in other words, doubted that politically sensitive legislators operating in the real world would actually do as the dissenters predicted.¹⁵⁵

Why should this be so? Why would voters find any more objectionable a high-to-low than a low-to-high sentencing scheme? To use the facts of *Apprendi* itself as an example, why would voters *support* a ten-year maximum for a simple firearms possession offense with a possible additional ten years if there were racial bias, but *oppose* a standard twenty-year maximum for the firearms offense with a reduction to ten years in the absence of racial bias? In principle, the

¹⁵¹ O’Connor put it this way in *Apprendi*:

New Jersey could cure its sentencing scheme, and achieve virtually the same results, by drafting its weapons possession statute in the following manner: First, New Jersey could prescribe, in the weapon possession statute itself, a range of 5 to 20 years’ imprisonment for one who commits that criminal offense. Second, New Jersey could provide that a defendant convicted under the statute whom a judge finds, by a preponderance of the evidence, *not* to have acted with a purpose to intimidate an individual on the basis of race may receive a sentence no greater than 10 years’ imprisonment.

530 U.S. at 541-42 (O’Connor, J., dissenting).

¹⁵² On the other hand, inversion might advance the third objective, fairness to defendants, inasmuch as an increase in the standard maximum sentence might diminish the likelihood of defendants receiving a nasty surprise at sentencing.

¹⁵³ 530 U.S. at 466 n.16.

¹⁵⁴ *Id.*

¹⁵⁵ In case this were not enough, the *Apprendi* Court also reserved some “wobble room” for itself to deal with inversion through constitutional adjudication. Chanenson, *supra* note 62, at 415.

consequences of both laws should be the same.¹⁵⁶ Yet, there is something intuitively plausible in Stevens’s claim that voters will *perceive* the two laws quite differently, and, indeed, precisely as Stevens predicted, legislatures have not rushed to adopt top-down schemes in the wake of *Apprendi*.¹⁵⁷

While not fully articulated, Stevens’s intuition must rest on assumptions of cognitive bias.¹⁵⁸ In order to see why, we first need to consider what it is that legislators want from sentencing law. In his influential work on the politics of criminal law, Professor Stuntz has identified two key objectives: (1) ensuring conviction and punishment of people “who commit the kinds of offenses that voters fear,” and (2) taking symbolic stands against the latest crime dominating the headlines¹⁵⁹ (the proverbial “offense du jour”). Legislators, however, do *not* want indiscriminate across-the-board sentence increases. For one thing, they must recognize that voters do not necessarily support severe sentences for low-level criminals.¹⁶⁰ For another, sentence increases cost money, and legislators (at the state level at least) must be mindful of fiscal pressures.¹⁶¹ Indeed, if

¹⁵⁶ On the margins, a few cases might come out differently based on the different allocation of the burden of proof. It seems unlikely, however, that a technical legal difference of this nature would engender the sort of dramatically different voter response that Stevens contemplated.

¹⁵⁷ The states most affected by the *Apprendi* line of cases have instead either engrafted jury fact-finding onto their presumptive sentencing regimes, converted from presumptive to discretionary, or ignored the problem. Stemen & Wilhelm, *supra* note 64, at 8-9.

¹⁵⁸ The “democratic constraints” argument has been subject to much criticism on the ground that, among other things, it rests on the implausible assumption that voters understand statutory maxima but not sentence enhancements. See, e.g., Stephanos Bibas, *Judicial Fact-Finding and Sentence Enhancements in a World of Guilty Pleas*, 110 YALE L.J. 1097, 1137 (2001). I offer here an original defense of Stevens’s hypothesis that does not rest on the assumption that voters have a patchwork knowledge of how sentencing works, but, rather, on the assumption that voters and legislators evaluate what they “know” in ways that are shaped by framing and anchoring effects.

¹⁵⁹ William J. Stuntz, *The Pathological Politics of Criminal Law*, 100 MICHIGAN L. REV. 505, 530-32 (2001).

¹⁶⁰ See Vincent Schiraldi & Judith Greene, *Reducing Correctional Costs in an Era of Tightening Budgets and Shifting Public Opinion*, 14 FED. SENT’ING RPTR. 332, 332-33 (2002) (summarizing evidence of recent softening in public attitudes towards sentencing and growing interest in rehabilitative alternatives for nonviolent criminals). In light of recent public corruption prosecutions, there may also be a measure of self-interest on the part of legislators in avoiding indiscriminate sentence increases. For a discussion of this phenomenon, see Craig S. Lerner, *Legislators as the “American Criminal Class”: Why Congress (Sometimes) Protects the Rights of Defendants*, 2004 U. ILL. L. REV. 599, 603-04 (2004).

¹⁶¹ See Ronald F. Wright, *The Power of Bureaucracy in the Response to Blakely and Booker*, 43 HOUSTON L. REV. 389, 412-13 (2006) (arguing, based on record of state responses to *Blakely* and *Booker*, that state sentencing commissions can successfully influence legislative reforms with appeals to fiscal responsibility).

expenditures on incarcerating low-level criminals are too high, legislators may compromise their own future capacity to respond to new offenses du jour.

It is the undesirability of indiscriminate sentence increases that makes inversion so politically unattractive. To be sure, in a world of rational actors, it would be easy to invert without causing a broad increase in sentence lengths. But two forms of cognitive bias, framing and anchoring, complicate matters. First, because aggravators are framed as losses for convicted defendants at sentencing, while mitigators are framed as benefits, legislators will find it unappealing to enact a new sentencing scheme that is rich in mitigators but poor in aggravators; doing so is likely to appear to the public as “soft on crime.” (What politician would tout her vote for reducing the sentences of criminals who don’t happen to be racists?) Yet, without the simultaneous conversion of aggravators to mitigators, a broad increase in standard sentence ranges is likely to result in increased actual sentence lengths.

Second, research on anchoring effects teaches that starting points can condition outcomes, even if the starting points are irrelevant or arbitrary. This suggests that, in practice, simple inversion of a sentencing system (even assuming full conversion of aggravators to mitigators) will not produce the same outcomes that were achieved pre-inversion; instead, the newly increased standard sentence that is triggered by a conviction, even though technically irrelevant to defendants who qualify for mitigated sentences, will likely push sentences up across the board. This provides further support for Stevens’s intuition that legislatures will avoid the inversion option.

b. Discretionary Sentencing

If inversion is politically unappealing, a legislature might instead evade *Apprendi* by adopting a discretionary sentencing system. Nothing in the *Apprendi* line of decisions casts doubt on the constitutionality of judicial fact-finding in connection with the selection of a sentence within a broad range; indeed, *Booker* itself makes clear that a discretionary system employing *advisory* guidelines with judicial fact-finding is perfectly constitutional.¹⁶² Such a switch would plainly undermine the goal of democratic control, as well as lift the procedural tax on the exercise of state prosecutorial power.

Responding to this argument, *Blakely* suggested that the switch to discretion, like inversion, might prove to be more a theoretical than an actual concern,

¹⁶² 543 U.S. at 233 (“We have never doubted the authority of a judge to exercise broad discretion in imposing a sentence within a statutory range.”).

noting that Kansas, the first state to modify its sentencing system as a result of *Apprendi*, had opted for jury fact-finding in lieu of judicial discretion.¹⁶³ Still, the possibility that some states might abandon presumptive sentencing schemes was not dismissed out of hand, as inversion had been. And the fact that the *Booker* remedy majority subsequently adopted this very approach for the federal system suggests that such dismissal would have been mistaken.

But the possibility of greater reliance on discretionary sentencing does not really undermine *Apprendi*'s aims as much as might first appear. First, as the example of Kansas and other states suggest, the contemporary unpopularity of discretionary sentencing means that many presumptive jurisdictions will likely choose to remain presumptive (with the addition, of course, of jury fact-finding).¹⁶⁴ Even if democratic control is not uniformly advanced across the country, it will be in at least some jurisdictions.¹⁶⁵

Second, while a switch to discretion may lift the procedure tax, a system reformed along these lines may nonetheless offer greater constraints on state power than a mandatory system. In a discretionary system, the judge may serve as a meaningful check on prosecutorial overreaching,¹⁶⁶ while in a mandatory system the judge may be reduced to a much less significant "bean-counting" role.¹⁶⁷ Prosecutors can dominate a mandatory system by controlling which sentence enhancements are sought and which grounds for leniency are supported or contested.¹⁶⁸ While prosecutors certainly do not always use their power to obtain the longest possible sentence, there are good reasons to believe that prosecutors, by and large, tend to take a more favorable view of their own cases

¹⁶³ 542 U.S. at 390-10.

¹⁶⁴ Additionally, some states, like Minnesota, have chosen to retain their presumptive systems because they have come to rely on the ability of presumptive sentencing to predict and control corrections resource needs. Dale G. Parent & Richard S. Frase, *Why Minnesota Will Weather Blakely's Blast*, 18 FED. SENT'ING RPTR. 12, 17 (2005).

¹⁶⁵ In addition to Kansas, at least six more states have recently adopted jury fact-finding for sentencing purposes in the wake of the *Apprendi* decisions. Stemen & Wilhelm, *supra* note 64, at 8. At least two states have switched from mandatory to discretionary, while several others with mandatory systems that seem to violate *Blakely* have yet to respond. *Id.* at 8-9.

¹⁶⁶ See Jeffrey Standen, *Plea Bargaining in the Shadow of the Guidelines*, 81 CAL. L. REV. 1471, 1474 (1993) ("Historically, the prosecutor's extraordinary bargaining power over defendants was constrained by independent judicial sentencing.").

¹⁶⁷ See *id.* at 1475 ("Judicial sentencing no longer limits prosecutorial power in federal courts."); Kate Stith & José A. Cabranes, *Judging Under the Federal Sentencing Guidelines*, 91 NORTHWESTERN U. L. REV. 1247, 1255 (1997) ("One judge has recently likened his role in sentencing to that of a 'notary public' . . .").

¹⁶⁸ O'Hear, *supra* note 16, at 808.

than do judges.¹⁶⁹ And, indeed, federal judges, with their newly enhanced discretion, have been imposing a higher rate of sentences below guidelines-mandated levels than pre-*Booker*.¹⁷⁰

Finally, *Blakely* argued for the fairness of a discretionary system relative to a mandatory one.¹⁷¹ Fairness, as BLE scholars suggest, should take into account baseline expectations, and a discretionary sentencing system will satisfy expectations insofar as it reliably results in a sentence within the maximum established by the conviction. As Justice Scalia put it, “I think it not unfair to tell a prospective felon that if he commits his contemplated crime he is exposing himself to a jail sentence of 30 years—and that if, upon conviction, he gets anything less than that he may thank the mercy of a tenderhearted judge.”¹⁷² By contrast, the problem with a mandatory system is that it might, “with no warning,” result in a dramatically higher sentence than expected on the basis of the offense of conviction.¹⁷³

¹⁶⁹ See Alafair S. Burke, *Improving Prosecutorial Decision Making: Some Lessons of Cognitive Science*, 47 WM. & MARY L. REV. 1587, 1590-91 (2006) (describing how cognitive bias helps to explain “the failure of prosecutors always to make just decisions”). There are limits to this line of reasoning; as Professor Stuntz observes, “[M]ost of the judges are elected by the same voters who elect district attorneys.” William J. Stuntz, *Plea Bargaining and Criminal Law’s Disappearing Shadow*, 117 HARV. L. REV. 2548, 2561 (2002). Many judges, moreover, are former prosecutors or are otherwise philosophically inclined to give prosecutors the benefit of the doubt. At the same time, it is important to recognize that democratic accountability is often structured quite differently—and less robustly—for judges than for prosecutors, *supra* note 116, and judges are largely protected from the sorts of case-specific cognitive bias described by Professor Burke. Judges’ separation from the institutional culture of law enforcement may also enhance their ability to take an appropriately skeptical view of some prosecutions. See GARY T. LOWENTHAL, *DOWN AND DIRTY JUSTICE: A CHILLING JOURNEY INTO THE DARK WORLD OF CRIME AND THE CRIMINAL COURTS* 111 (2003) (describing anti-defendant culture within which prosecutors work).

¹⁷⁰ Douglas A. Berman, *Tweaking Booker: Advisory Guidelines in the Federal System*, 43 HOUSTON L. REV. 341, 350 (2006) For a more detailed argument in favor of judicial discretion at sentencing as a means of protecting individual liberty, see Luna, *supra* note 132, at 95-100.

¹⁷¹ *Supra* text accompanying note 131.

¹⁷² *Apprendi*, 530 U.S. at 498 (Scalia, J., concurring).

¹⁷³ *Blakely*, 542 U.S. at 311-12. The argument here is colorable, but admittedly not compelling on its face, for it assumes that defendant expectations in a discretionary system will be shaped by the maximum possible sentence. It is possible, however, that in some cases the “anchor” will be a much smaller number, for instance, the sentencing prediction of the defendant’s lawyer, the recommendation made by the prosecutor, or a sentence recently imposed in a similar case. Indeed, this is almost certain to be the case in a discretionary system with robust advisory guidelines, such as the post-*Booker* federal system. If the anchor is low, then a defendant’s expectations may be as severely disappointed as in a mandatory system. *Cf.* United States v. Vampire Nation, 451 F.3d 189, 197 (3d Cir. 2006) (holding that judge need not give advance notice to defendant of intent to impose sentence above the federal guidelines’ advisory range). Indeed, if a mandatory system

In sum, while I do not mean to argue that discretionary sentencing is necessarily better than mandatory, there are at least plausible reasons, taking into account libertarian and fairness values, to prefer a discretionary system over a mandatory system that uses judicial fact-finding.¹⁷⁴

c. Plea Bargaining

Given the option of discretionary sentencing, legislatures might or might not choose to give juries a more robust role to play in the sentencing process. To whatever extent juries were formally made available, however, the *Apprendi* dissenters predicted that they would be seldom used: “[T]he greater expense attached to trials and their greater complexity, taken together in the context of an overworked criminal justice system, will likely mean, other things being equal, fewer trials and a greater reliance upon plea-bargaining—a system in which punishment is set not by judges or juries but by advocates acting under bargaining constraints.”¹⁷⁵ Moreover, not only would *Apprendi* produce fewer trials, but, the dissenters argued, the resulting system would be less fair, in the sense that it would be less uniform. Prosecutors would control the punishment by deciding which sentencing factors to charge and then bargain away.¹⁷⁶ These prosecutorial processes, however, “lack transparency and too often mean nonuniform, sometimes arbitrary, sentencing practices.”¹⁷⁷

Since at least the time of *Bordenkircher*, increasing numbers of defendants have been surrendering their right to a jury trial on the basic issue of guilt or innocence,¹⁷⁸ and there is no reason to doubt the assumption of the *Apprendi* dissenters that the right to a jury trial on *sentencing factors* will be treated any differently. This does not necessarily mean, however, that *Apprendi* is a futile gesture. Again, consider *Apprendi*’s three aims in turn.

First, while most defendants might trade away their *Apprendi* rights, that does not necessarily mean that all will. Moreover, those defendants most likely

operates in a transparent and predictable fashion, it may do an even better job of fulfilling expectations than a discretionary system. A definitive resolution of the fairness question, then, depends on a number of uncertain empirical issues.

¹⁷⁴ Professor Huigens also defends discretionary systems as better able to achieve “fine-grainedness” in sentencing, which is “important for maintaining the moral credibility and public standing of the criminal justice system.” Kyron Huigens, *Solving the Williams Puzzle*, 105 COLUM. L. REV. 1048, 1069 (2005).

¹⁷⁵ *Blakely*, 542 U.S. at 338 (Breyer, J., dissenting).

¹⁷⁶ *Id.*

¹⁷⁷ *Id.* at 345.

¹⁷⁸ See Ronald F. Wright, *Trial Distortion and the End of Innocence in Federal Criminal Justice*, 154 U. PENN. L. REV. 79, 91 (2005) (showing long-term guilty plea trends in federal cases).

to insist on a jury trial will include those who most stubbornly believe that the application of a particular sentence enhancement sought by a prosecutor would be unjust. It is precisely in these sorts of cases, in which there are legitimate disputes surrounding the application of a sentencing factor, that a jury might have the most to contribute as the voice of the community. It is the hard cases in which democratic control is the most important in the implementation of sentencing enhancements, and *Apprendi* does at least ensure the availability of a jury in those cases. As *Blakely* observed, “That more defendants elect to waive that right [to a jury] (because, for example, government at the moment is not particularly oppressive) does not prove that a constitutional provision guaranteeing *availability* of that option is disserved.”¹⁷⁹ The flipside is that when government *does* act oppressively, as by seeking morally or legally dubious sentencing enhancements, the defendant may seek protection from his or her peers on the jury.

Second, to the extent that defendants do bargain away their *Apprendi* rights, the procedural tax will indeed be lifted, but in a manner that is consistent with the goal of limiting state power, for defendants will likely receive concessions from prosecutors in the process. “Every new element that a prosecutor can threaten to charge is also an element that a defendant can threaten to contest at trial and make the prosecutor prove beyond a reasonable doubt.”¹⁸⁰

Finally, we reach the nub of the dissenters’ objection to *Apprendi*, the fairness point. A mandatory guidelines system that rests on jury fact-finding will leave prosecutors with gatekeeping authority over the application of sentencing enhancements, and prosecutors will not ensure uniform results across cases.¹⁸¹ The *Apprendi* Five responded, however, that the system of “judicial” sentencing was plagued by its own uniformity problems: prosecutors had plenty of tools available to influence the sentence even in the world of judicial fact-finding, as

¹⁷⁹ *Blakely*, 542 U.S. at 312.

¹⁸⁰ *Id.* at 311. While not all would agree with this proposition, see Bibas, *supra* note 15, at 198-99, Professors King and Klein have persuasively argued that the circumstances in which defendants would be worse off under *Apprendi* are narrow and unusual. Nancy J. King & Susan R. Klein, *Apprendi and Plea Bargaining*, 54 STAN. L. REV. 295, 306-07 (2001).

¹⁸¹ Note that the same uniformity concerns might be raised as to discretionary sentencing systems. It is important, however, to recall that discretionary sentencing can be implemented in a variety of ways other than through the traditional system of completely unguided discretion. Such alternative approaches may, in fact, be effective in achieving uniformity goals. See Hunt & Connelly, *supra* note 62, at 235 (discussing higher compliance rates with advisory guidelines in some states than with presumptive guidelines in other states).

by making sentencing recommendations.¹⁸² Such tools could easily be used as bargaining chips, thereby creating the same sorts of problems of prosecutor-created disparity that the dissenters feared would arise post-*Apprendi*.¹⁸³

More fundamentally, though, there was the issue of *perceived* unfairness discussed above: in the world of judicial fact-finding, “a defendant, with no warning in either his indictment or plea, would routinely see his maximum potential sentence balloon from as little as five years to as much as life imprisonment.”¹⁸⁴ Note the very different views of what constitutes fairness from the perspective of the dissenters (whose objective is the “systematic fairness” of consistent results across the run of cases¹⁸⁵) and the majority (who concentrate on the perceptions of individual defendant in light of the expectations established by earlier proceedings in their own cases). Seen through the latter lens, the *Apprendi* canon can indeed be viewed as providing a fairness benefit to defendants, even in the absence of a greater number of jury trials.

d. Summary

The *Apprendi* Five repeatedly rejected the charge of doctrinaire formalism and instead asserted that they were advancing democratic, libertarian, and fairness values in a meaningful way. One does not have to agree fully with all of their contentions to acknowledge that they could, in good faith, lay claim the consequentialist high ground. In any event, what is important for present purposes is not whether the Five ultimately had the better of the argument with the dissenters, but what the Five should think of *Bordenkircher*, assuming that they really meant what they said in the *Apprendi* decisions.

The Five expected that their decisions would cause some jurisdictions to enhance judicial discretion and others to create a formal role for juries in the sentencing process. Their alignment in the *Booker* remedy opinions suggests that one of the Five (Justice Ginsburg) preferred the discretion option, while the other four preferred the jury option. In any event, they plainly viewed both options as

¹⁸² 542 U.S. at 311.

¹⁸³ This is one among several reasons why I have argued that the dissenters greatly overstated the uniformity in federal sentencing prior to *Booker*. Michael M. O’Hear, *The Myth of Uniformity*, 17 FED. SENT’ING RPTR. 249 (2005).

¹⁸⁴ *Blakely*, 542 U.S. at 311.

¹⁸⁵ *Blakely*, 542 U.S. at 339 (Breyer, J., dissenting). Interestingly, while the dissenters framed the uniformity issue as one of abstract, system-wide fairness, they might have reframed the issue in ways that were more in line with the majority’s subjective understanding of fairness: sentencing reformers in the 1970’s and 1980’s frequently argued that sentencing disparities were a source of great unhappiness among prisoners. O’Hear, *supra* note 16, at 760, 772-73.

constitutionally acceptable and preferable to the pre-*Apprendi* world of sentence enhancements triggered by judicial fact-finding.

Finally, *Apprendi*'s consequentialism seems influenced by the view that (as BLE scholars put it) "people evaluate outcomes based on the change they represent from an initial reference point, rather than based on the nature of the outcome itself."¹⁸⁶ This insight helps to explain *Apprendi*'s dismissal of the inversion argument, as well as the contrasting definitions of fairness employed by the majority and dissenters. For this reason, we might usefully think about *Apprendi*'s consequentialism as "biased-actor" consequentialism, differentiating it from a "rational-actor" consequentialism that predicts and evaluates consequences on the traditional *homo economicus* model, in which people think about outcomes without regard to the path by which the outcomes are reached.

III. INCONSISTENCIES BETWEEN *APPRENDI* AND *BORDENKIRCHER*

This Part elaborates on the inconsistencies between the *Apprendi* decisions and *Bordenkircher*. These inconsistencies have two dimensions: (1) the relative value ascribed to bureaucratic efficiency in the criminal justice system, and (2) the use of rational-actor assumptions.

A. *Common Law Values Versus Bureaucratic Efficiency*

The *Apprendi* decisions were not framed as narrowly addressed to technical questions of law, but rather purported to select one "paradigm of criminal justice" over another: "the common-law ideal of limited state power" over "the civil-law ideal of administrative perfection,"¹⁸⁷ the "common-sense judgment" of the jury over the "more tutored" judgment of the legal professionals.¹⁸⁸ *Bordenkircher* also reflected an underlying choice of paradigms, but the paradigm chosen was the one that *lost* in the *Apprendi* cases. The *Bordenkircher* Court concluded that guilty pleas were an "important[] component of this country's criminal justice system"¹⁸⁹; characterized its prior decisions as not merely tolerating, but "encouraging," the inducement of pleas¹⁹⁰; and indicated that its "acceptance of the basic legitimacy" of the practice necessitated deference to the prosecutor's choice of tactics.¹⁹¹ *Bordenkircher*'s vision of unconstrained plea-inducement is profoundly at odds with *Apprendi*'s paradigm.

¹⁸⁶ Jolls, et al., *supra* note 17, at 1535.

¹⁸⁷ *Blakely*, 542 U.S. at 313.

¹⁸⁸ *Ring*, 536 U.S. at 609.

¹⁸⁹ *Bordenkircher*, 434 U.S. at 362 (citation omitted).

¹⁹⁰ *Id.* at 364.

¹⁹¹ *Id.* at 363-64.

It is a system dominated by legal professionals, with no room for the “common-sense judgment” of jurors. While the system may boast considerable speed and efficiency, the *Apprendi* decisions repeatedly rejected the centrality of these virtues.¹⁹²

To be sure, the *Apprendi* Five did not anticipate that their decisions would turn back the clock to a time when jury trials were the norm in the criminal justice system.¹⁹³ But they suggested that, even in the absence of common-law *practices*, the Constitution should be interpreted so as to safeguard common-law *values*, most notably, the value of maintaining a robust check on the ability of prosecutors to impose disproportionate punishments on an arbitrary or vindictive basis. The repeated reference in the *Apprendi* decisions to the common-law tradition of pious perjury is telling.

Heavy-handed charging threats, a la *Bordenkircher*, embody contempt for the common-law values of checks and balances in the criminal justice system. They send defendants a message that only the prosecutor’s view of the case counts, and that the system accords little actual value to their formal rights to be heard.¹⁹⁴ Moreover, the inconsistencies between *Apprendi* and *Bordenkircher* go beyond such symbolic considerations.¹⁹⁵ If prosecutors have a free hand to raise the

¹⁹² See, e.g., *Ring*, 536 U.S. at 607 (“It has never been efficient; but it has always been free.”). As the *Apprendi* decisions suggest, efficiency did not become a central preoccupation of the criminal justice until well after the framing of the Sixth Amendment. See Nancy Jean King, *Priceless Process: Nonnegotiable Features of Criminal Litigation*, 47 U.C.L.A. L. REV. 113, 121-23 (1999) (describing emergence of efficiency concerns in late 1800’s and early 1900’s).

¹⁹³ See King, *supra* note 192, at 119-20 (“Courts of the nineteenth century would have ridiculed the idea that an accused and a prosecutor could dicker over what kind of break the defendant deserves for waiving a piece of the criminal process.”).

¹⁹⁴ Professors Scott and Stuntz have a wonderful analogy for *Bordenkircher* that nicely captures this aspect of the prosecutor’s actions:

The . . . analogy is . . . the lone gas station in the middle of the desert that charges fifty dollars for a gallon of gas. Like the prosecutor in *Bordenkircher*, the gas station usually gets its asking price, because the difference between that price and the cost of going without (death in the desert) is so high. . . . Imagine, however, that the gas station owner goes further. Figuring that the buyer will kick and scream and haggle for an hour, but will eventually agree to the seller’s price, the seller decides to cut the negotiation short by letting the air out of the buyer’s tires and offering to refill the tires if, but only if, the buyer pays the seller’s asking price for gas.

Scott & Stuntz, *supra* note 12, at 1964.

¹⁹⁵ As a growing body of social psychology research suggests, however, symbolism itself can play an important role in the ability of the criminal justice system to achieve its objective of enhancing public compliance with the mandates of the law and legal institutions. See, e.g., Larry Heuer, *What’s Just About the Criminal Justice System? A Psychological Perspective*, 13 J. L. & POL’Y

stakes in a jury trial as high as they wish, then only the most impetuous defendants will dare to invoke the protection from government overreaching that is potentially afforded by a jury. The availability of the jury is thereby effectively curtailed. And those who dare resist the pressure, like Paul Hayes, may pay an extraordinary price for doing so. Ironically, those who ask a jury to determine whether the government is overreaching may thereby *assure* that they become victims of overreaching.

One response to all of this, of course, is that, however troubling charging threats might be, it is futile to regulate them. Prosecutors will simply file more aggressive initial charges, switching from “threats” to “offers,” with equally coercive effects. This argument, analogous to the inversion argument made by the *Apprendi* dissenters, is considered in the next Section.

B. Biased Actors Versus Rational Actors

Bordenkircher's consequentialism parallels that of the *Apprendi* dissenters. This becomes most apparent by reconsidering a portion of Justice Blackmun's dissent in *Bordenkircher*, specifically, his response to the majority's view that regulating the low-to-high strategy would merely cause prosecutors to switch to the seemingly equivalent high-to-low approach. Blackmun identified three reasons why high-to-low was actually preferable to low-to-high. These three reasons resonate with the thinking behind the *Apprendi* decisions and indicate why the *Apprendi* Five should find *Bordenkircher*'s analysis unpersuasive.

First, Blackmun argued, under the high-to-low approach,

the prosecutor is required to reach a charging decision without any knowledge of the particular defendant's willingness to plead guilty; hence the defendant who truly believes himself to be innocent, and wishes for that reason to go to trial, is not likely to be subject to quite such a devastating gamble since the prosecutor has fixed the incentives for the average case.¹⁹⁶

209, 226 (2005) (“Fair treatment enhances satisfaction with the criminal justice system and improves perceptions of the legitimacy of the law and of the authorities who enforce it. . . . Perceptions of fair treatment are heavily influenced by symbolic criteria, such as politeness and respect.”); TOM R. TYLER & YUEN J. HUO, TRUST IN THE LAW: ENCOURAGING PUBLIC COOPERATION WITH THE POLICE AND COURTS 83-84 (2002) (discussing study showing that people's views of procedural justice are heavily influenced by such considerations as whether decisionmakers are polite and show concern for people's rights).

¹⁹⁶ *Bordenkircher*, 434 U.S. at 368 n.2 (Blackmun, J., dissenting).

The reasoning here parallels Stevens’s argument in *Apprendi* that legislatures won’t really increase penalties across the board and convert aggravators into mitigators: both arguments assume a subtle, practical check on the ability of politically responsive state actors to inflate penalties—even if the inflated penalties merely operate, at a formal level, as a starting point in the analysis.

In a world of unbounded rationality, it is not clear why a prosecutor going from high to low would base her initial charge on the “average case,” rather than charging the legally permissible maximum.¹⁹⁷ In a world of cognitive bias, however, a number of justifications for the practice are apparent. The prosecutor’s selection of a disproportionately serious charge may have powerful anchoring effects. The prosecutor may appreciate the difficulty that she would have in moving a great distance down from the initial charge, and avoid gross overcharging on that basis; she will understand that if she cannot move far enough down from the initial charge, she risks wrecking the plea negotiations.¹⁹⁸ Moreover, an eventual conviction of a disproportionately high charge, whether by trial or plea, may be undesirable in and of itself, as the prosecutor may fear the exhaustion of limited criminal justice resources on relatively minor cases or an adverse public reaction to undue prosecutorial harshness.¹⁹⁹

¹⁹⁷ To be sure, ethical standards for prosecutors prohibit using charges “only as a leverage device in obtaining guilty pleas to lesser charges.” NAT’L DISTRICT ATTORNEYS ASS’N, NATIONAL PROSECUTION STANDARDS 130 (2d.ed. 1991). Such standards, however, are not enforceable. *See id.* at 5 (noting that standards not intended to be basis for sanctions). Our present concern, moreover, is not with prosecutors who rigorously adhere to the standards, but with prosecutors who already use the threat of enhanced charges to obtain plea bargaining leverage, and the extent to which they would simply switch to a different leveraging tactic if prohibited from using threats.

¹⁹⁸ As BLE scholars have observed, people in a variety of contexts recognize and take steps to address the boundedness of their willpower, as by joining “Christmas Clubs” in order to avoid undersaving. Jolls et al., *supra* note 17, at 1479.

¹⁹⁹ Prosecutors do not always seek the harshest sentence available in every case; indeed, there is ample evidence that prosecutors will sometimes even act so as to subvert mandatory sentencing guidelines in order to achieve what they believe to be a more appropriate set of sentences. *See, e.g.*, Frank O. Bowman, III, & Michael Heise, *Quiet Rebellion II: An Empirical Analysis of Declining Federal Drug Sentences Including Data From the District Level*, 87 IOWA L. REV. 477, 559 (2002) (“Our work demonstrates that prosecutors . . . use their discretion liberally, but irregularly, to reduce drug sentences.”). *See generally* Stuntz, *supra* note 169, at 2554 (“Once the defendant’s sentence has reached the level the prosecutor prefers—or, if you like, the level that the local voters who elect her boss demand—adding more time offers no benefit to the prosecutor. . . . Voters’ preferences, courthouse customs, the prosecutor’s reputation as a tough or lenient bargainer, her own views about what is a proper sentence for the crime in question—all these things play a role in defining the sentences that prosecutors are likely to seek in plea bargains.”). *See also id.* at 2257 (“[W]ho bears the blame if something goes wrong—if sympathetic defendants are punished or if unsympathetic defendants are punished more harshly than the public thinks just? In the United

Adding weight to the anchoring effects, there are also framing effects. Movement up in charges will be seen by the public as a loss for defendants, but movement down represents a win. In light of tough-on-crime political pressures, the prosecutor will want to be careful about routinely handing out big wins to defendants. This consideration, too, would militate against indiscriminately maximizing initial charges; it is better to set initial charges with some modesty so that defendants will usually get only small wins in plea negotiations.²⁰⁰

Finally, there is the principle of reciprocal fairness: a prosecutor who significantly overcharges risks being perceived as unfair by opposing counsel, which may prompt spiteful responses that make the prosecutor's life more difficult, such as more filing of pretrial motions.²⁰¹ Similarly, unfair overcharging may leave some defendants feeling more reluctant to plead guilty or otherwise cooperate with prosecutors. In short, Blackmun's assumptions about prosecutorial charging restraint seem at least plausible in a world of bounded rationality and willpower.

Second, Blackmun reasoned,

It is healthful to keep charging practices visible to the general public, so that political bodies can judge whether the policy being followed is a fair one. Visibility is enhanced if the prosecutor is required to lay his cards on the table with an indictment of public record at the beginning of the bargaining process, rather than making use of unrecorded verbal warnings of more serious indictments yet to come.²⁰²

This argument, like the first, also rests on an assumption of prosecutorial restraint in determining the initial charges. Consider an example: a defendant is initially charged with second-degree murder, but, pursuant to an agreement, ultimately

States, the answer is almost always the overzealous prosecutor, not the overcriminalizing or oversentencing legislator.”).

²⁰⁰ Not only may prosecutors arouse a negative public reaction by handing out big wins, but they may also run into difficulties in having their plea deals approved by judges. See King, *supra* note 192, at 136 (noting that all jurisdictions judges retain authority to review and reject plea deals). For an example, see Dan Christensen, *Florida Judge Complains U.S. Prosecutor Is “Weak-Kneed”*, MIAMI DAILY BUSINESS REV., March 18, 2004, available at, <http://www.law.com/jsp/article.jsp?id=1079144426509> (discussing case in which judge threatened to reject plea deal in which prosecutor agreed to dismiss two felony drug distribution counts in return for guilty plea to one misdemeanor possession charge).

²⁰¹ For a discussion of social science research on spiteful responses to perceived unfairness, see Jolls et al., *supra* note 17, at 1494-96.

²⁰² *Bordenkircher*, 434 U.S. at 368 n.2 (Blackmun, J., dissenting).

pleads guilty only to reckless endangerment. What is the public to make of this deal? At first blush, it appears that the prosecutor paid a high price for the defendant's guilty plea. In the world of the *Bordenkircher* majority, however, the initial murder charge may have been far more than the prosecutor actually thought appropriate, a massive overcharge intended to ensure that only the most impetuous defendant would actually go to trial. In the dissenters' world, however, prosecutors act in a more restrained fashion, only overcharging to the relatively modest extent necessary to secure a guilty plea from the average defendant. In the dissenters' world, the public can more confidently conclude that the prosecutor has indeed moved a long distance from what she initially judged to be an appropriate disposition of the case.

Third, Blackmun asserted, "I would question whether it is fair to pressure defendants to plead guilty by threat of reindictment on an enhanced charge for the same conduct when the defendant has no way of knowing whether the prosecutor would indeed be entitled to bring him to trial on the enhanced charge."²⁰³ The objection, of course, is that defendants always bargain in the face of uncertain trial outcomes²⁰⁴; it is not clear why the modest additional layer of uncertainty in *Bordenkircher* should be accorded particular significance.

Blackmun is nonetheless onto something by raising fairness concerns. A threat is perceived differently than a offer: the former is framed as a loss, and the latter as a gain. Powerful loss-aversion instincts may lend considerably greater coercive power to a threat than a offer. BLE scholars have shown that people will go to irrational extremes in order to avoid a loss,²⁰⁵ which ought to raise concerns about whether guilty pleas are as well-considered when made in response to a threat as they are when made in response to an offer.²⁰⁶ The defendant, moreover, is likely at least to *perceive* himself treated less fairly in the threat scenario: his expectations are shaped by the charges initially filed, but then

²⁰³ *Id.*

²⁰⁴ See Bibas, *supra* note 141, at 2493-96 (discussing information deficits that plague plea negotiations).

²⁰⁵ *Id.* at 2508. Defense lawyers might, in principle, serve as debiasers for their clients, but, as Professor Bibas has observed, "[L]awyers vary widely in their knowledge, skill, and incentives to debias their clients." *Id.* at 2520.

²⁰⁶ See Wright, *supra* note 178, at 84 (arguing that declining federal acquittal rates demonstrate that prosecutors have accumulated too much power to extract guilty pleas from defendants who might otherwise win acquittals at trial). A number of recent police misconduct scandals provide further anecdotal evidence that even innocent defendants will some times take plea deals in order to avoid the risks of going to trial. . Oren Gazal-Ayal, *Partial Ban on Plea Bargains*, 27 *CARDOZO L. REV.* 2295, 2305 (2006).

the prosecutor suddenly brings the possibility of much worse losses into view.²⁰⁷ This echoes *Blakely*'s fairness objections to charging a defendant in a way that indicates one particular maximum penalty, then punishing him in a way that far exceeds the maximum.²⁰⁸

In sum, the *Apprendi* Five should not find the *Bordenkircher* evasion argument any more persuasive than they found the evasion arguments made by the *Apprendi* dissenters. There are good reasons to doubt that regulation of charging threats would result in dramatic inflation of initial charges. And, even to the extent that some inflation does occur, there may nonetheless be reasons to view high-to-low charging as preferable to low-to-high. High-to-low is more transparent to the public, and is likely perceived by defendants as less coercive and unfair.

²⁰⁷ In contrast to the bait-and-switch character of low-to-high charging, high-to-low negotiations are not necessarily as crass as they are sometimes imagined. See Gerard E. Lynch, *Screening Versus Bargaining: What Exactly Are We Trading Off?*, 55 STAN. L. REV. 1399, 1403 (2003) ("Most plea negotiations, in fact, are primarily discussions of the merits of the case, in which defense attorneys point out legal, evidentiary, or practical weaknesses in the prosecutor's case, or mitigating circumstances that merit mercy, and argue based on these considerations that the defendant is entitled to a more lenient disposition than that originally proposed by the prosecutor's charge."). *But cf.* Ronald Wright & Marc Miller, *Honesty and Opacity in Charge Bargains*, 55 STAN. L. REV. 1409, 1413-14 (2003) (asserting the Lynch's description of plea bargaining, while perhaps accurate as to the federal system, does not match the reality of state court practice); MILTON HEUMANN, PLEA BARGAINING: THE EXPERIENCES OF PROSECUTORS, JUDGES, AND DEFENSE ATTORNEYS 41 (1977, paperback ed. 1981) (finding within single state system substantial differences in length of plea discussions and thoroughness with which facts of case were covered).

²⁰⁸ This approach, which focuses on changes to baseline expectations, also resonates with the abundant philosophical literature on coercion. See Mark A. Godsey, *Rethinking the Involuntary Confession Rule: Toward a Workable Test for Identifying Compelled Self-Incrimination*, 93 CAL. L. REV. 465, 525-26, 530 (2005) (summarizing literature); George C. Thomas III, *A Philosophical Account of Coerced Self-Incrimination*, 5 YALE J. L. & HUMAN. 79, 86 (1993) ("[A] proposal is a threat only when it makes [a person's] condition worse; and it is an offer when it makes [a person's condition] better."); ALAN WERTHEIMER, COERCION 136 (1987) (discussing difficulty of determining baseline in plea-inducement context). This literature identifies three ways of determining the baseline: empirical (i.e., what is statistically likely to occur absent the proposal), phenomenological (i.e., what is subjectively expected to occur absent the proposal), and normative (i.e., what is expected assuming compliance with moral norms). Thomas, *supra*, at 87-88. Assuming that charges are not usually enhanced after they are formally made, and that defendants do not expect them to be enhanced, then a *Bordenkircher*-style proposal would amount to a threat in light of either the empirical or the phenomenological baseline. (And if different baselines produce different results, then the right baseline to use would be one that reflects the defendant's preference as to the future. *Id.* at 89.) Assuming that the prosecutor intends to deliver a threat and that the threat actually causes the defendant to plead guilty, then the chief conditions of coercion have been satisfied. See *id.* at 83 (outlining basic conditions of coercion).

IV. DESIGNING A NEW RULE FOR CHARGING THREATS

The previous Part demonstrated why the *Apprendi* Five should reject *Bordenkircher*. Assume now that the Court actually grants certiorari in a new case that presents a similar fact pattern, with a prosecutor delivering on a threat of new charges following a break-down in plea negotiations. This Part considers how the Five might address a hypothetical *Bordenkircher II*.²⁰⁹ Section A discusses potential ways that a protective rule might be structured by the Court. Section B considers the possibility that Justice Stevens, the only holdover on the Court from the *Bordenkircher* majority, might defect from the Five in a hypothetical *Bordenkircher II*. Finally, Section C considers the consequences of the proposed rule in more detail, showing that the “doctrinaire formalism” charge would no more fairly apply in this context than in the *Apprendi* cases.

A. Structural Options

Part I above described the contrasting protective rules proposed by the dissenters in *Bordenkircher* and *Goodwin*. This Section considers which of the rules would be most attractive to the *Apprendi* Five, concluding that Justice Brennan’s test squares nicely with the *Apprendi* model.

1. Powell’s Wag-the-Dog Test

Justice Powell’s dissent in *Bordenkircher* suggested a case-by-case analysis of whether the trial penalty imposed by the prosecutor exceeded the State’s legitimate interests in punishing the defendant. Powell’s test offers flexibility and, assuming the sort of generous understanding of the scope of the State’s interests suggested by Powell, only minimal intrusiveness by judges into the charging and plea-inducement process. Indeed, Justice Stevens himself seemingly endorsed the Powell test in *Goodwin*.²¹⁰

More will be said about Stevens in the next Section, but, for now, note the basic jurisprudential inconsistency between Powell’s approach and the *Apprendi* decisions. Powell’s test employs a standard, instead of a rule, in contrast to *Apprendi*’s preference for rule-structure formalism. Powell’s approach, leaving wide discretion in the judiciary, offers little reassurance that jury-trial rights will not be progressively eroded, particularly given the institutional incentives for judges to support prosecutorial practices that result in more guilty pleas and

²⁰⁹ While this Part is written with the possibility of a new United States Supreme Court case in mind, the proposal developed here might alternatively be adopted by state courts using their independent authority to construe due process rights under their own state constitutions.

²¹⁰ *Supra* Part I.B.2.

fewer trials.²¹¹ For that reason, the *Apprendi* Five are unlikely to unite behind a wag-the-dog test.

2. *Blackmun's Actual Motivation Test*

Blackmun's dissent in *Bordenkircher* suggested a somewhat different approach: rather than focusing on the state's interests in punishment, the inquiry should turn on the prosecutor's actual motivation in bringing enhanced charges.²¹² Where a new charge is added after a break-down in plea negotiations, a burden-shifting analysis is triggered, requiring the prosecutor to justify her action on some non-vindictive basis.²¹³ Blackmun seems to have contemplated something like the *Batson* test for racial discrimination in jury selection, which makes use of a similar burden-shifting approach en route to a determination of whether the prosecutor discriminated against a group of prospective jurors on account of their race.²¹⁴

The *Batson* analogy, however, demonstrates the basic objection to Blackmun's approach. *Batson* has proven notoriously ineffectual in practice. Prosecutors have little difficulty in providing race-neutral explanations for their peremptory strikes, and courts have little stomach for probing their actual motivations once a facially neutral explanation is offered.²¹⁵ Likewise, one imagines that reasonably intelligent prosecutors will always be able to point to *something* that has changed in order to justify increased charges after the break-down in plea negotiations: an investigator has come up with a new item of

²¹¹ See Rachel E. Barkow, *Separation of Powers and the Criminal Law*, 58 STAN. L. REV. 989, 1037-38 (2006) (arguing in favor of greater use of formalist rules to preserve separation of powers in criminal justice system); Stuntz, *supra* note 169, at 2561 (discussing evidence that judges "are invested in plea bargaining and try to facilitate it"). Professor Stuntz has suggested a modified version of the wag-the-dog test, in which the government would be required to "point to some reasonable number of factually similar cases in which the threatened sentence had actually imposed, not just threatened." William J. Stuntz, *Bordenkirch v. Hayes: Plea Bargaining and the Decline of the Rule of Law*, in CRIMINAL PROCEDURE STORIES 351, 373 (Carol S. Steiker ed. 2006). This seems a move in the right direction, although, as Stuntz himself acknowledges, not all states collect the sort of data that would make the required showing possible. *Id.* Additionally, the proposed requirement might still be overly indeterminate and deferential for the tastes of the *Apprendi* Five, leaving open the questions of what is a "reasonable number" and how to know when cases are sufficiently "factually similar" to count as suitable comparisons.

²¹² See *supra* Part I.A.2.

²¹³ *Id.*

²¹⁴ See *Batson v. Kentucky*, 476 U.S. 79 (1986).

²¹⁵ For a recent summary and critique of this *Batson* case law along these lines, see Antony Page, *Batson's Blind-Spot: Unconscious Stereotyping on the Peremptory Challenge*, 85 B.U. L. REV. 155, 166-78 (2005).

evidence, a victim's injuries have taken longer than anticipated to heal, the prosecutor has just learned of a favorable appellate decision, etc. Given the ephemeral nature of actual motives, the judge will have considerable "wiggle room" to find in favor of the prosecutor once such a neutral explanation is offered. And, given the institutional incentives to support plea-bargained outcomes, one imagines that judges will liberally make use of this wiggle room, just as they have in the *Batson* context. In short, Blackmun's test offers hardly any firmer protection against the erosion of jury-trial rights than Powell's wag-the-dog test.

3. *Brennan's Likelihood-of-Deterrence Test*

Justice Brennan suggested a third approach in his *Goodwin* dissent. Brennan would ask, "Did the elevation of the charges . . . pose a realistic likelihood of vindictiveness? Is it possible that the fear of such vindictiveness may . . . deter a person in [the defendant's] position from exercising his . . . right to a jury trial?"²¹⁶ This test focuses neither on the propriety of the charge in the abstract (Powell's test) nor on the prosecutor's motive in bringing the charge (Blackmun's), but on the deterrent *effects* of bringing the charge in the manner in which it was brought.

a. Adapting the Test to Charging Threats

Goodwin itself did not involve the specific low-to-high plea-bargaining issue presented by *Bordenkircher*, but it is not hard to imagine how Brennan would have applied his test in such circumstances. Where a prosecutor threatens and then delivers an increased charge in response to a refusal to plead guilty, defendants will undoubtedly perceive vindictiveness, and there is at least a realistic likelihood that some will be deterred from exercising their procedural rights. Brennan's likelihood-of-deterrent-effects approach might thus be adapted to the *Bordenkircher* scenario through the following rule: *The Due Process Clause is violated when, after a defendant has been indicted or otherwise formally charged, the prosecutor (a) makes a plea offer; (b) in connection with the offer, threatens to take an action that exposes the defendant to a longer maximum sentence than would otherwise be possible based on the existing charge(s); and (c) does take such an action following the defendant's rejection of the offer.*

In light of the clarity of the rule, it should be easy for prosecutors to avoid running afoul of it. If a prosecutor has a potential charge that he or she would

²¹⁶ *Goodwin*, 457 U.S. at 389-90 (Brennan, J., dissenting).

like to use as a bargaining chip, then the prosecutor has two choices: either (1) convey the terms of the proposed deal prior to the formalization of charges in an indictment or information, or (2) include the charge in the indictment or information. Once post-charging plea negotiations begin, the charge might still be added—but only until there is an express threat relating to the charge, at which time the prosecutor would lose the ability to add the charge. (And even at that point, the prosecutor would still be able to employ any other form of plea-inducement that did not increase the defendant’s overall sentencing exposure, for instance, a promise to stand silent at sentencing.) In order to maintain flexibility in modifying charges, prosecutors would thus have a strong incentive never to make the post-charging threat.

Using the rule, a defendant’s proof of vindictiveness in the *Bordenkircher*-type setting would turn on a handful of objective facts, which should not involve protracted collateral litigation or undue intrusion into the inner workings of the prosecutor’s office. To be sure, the rule, as framed here, contains a few *legal* ambiguities whose resolution I will leave for another day.²¹⁷ Yet, the presence of such legal questions does not render the rule any less bright-line than the *Apprendi* test—which has also raised its fair share of legal questions.

b. Consistency With *Apprendi* (and Other Precedent)

The *Apprendi* Five should find the proposed rule preferable not only to the open-ended permissiveness of *Bordenkircher*, but also to the Blackmun or Powell approaches. The structure of this rule echoes the formalism of *Apprendi*, avoiding ephemeral considerations, such as state interests and prosecutorial intent, in favor of more objective fact-finding. Moreover, like the *Apprendi* test, the proposed test focuses squarely on the fairness of the *process* to which the defendant was subject, without regard to the fairness of the *outcome* as an abstract proposition.²¹⁸ Finally, like the *Apprendi* rule (as clarified in *Harris*), the proposed test makes the maximum available sentence the touchstone of the analysis; it constitutes a baseline against which subsequent changes are assessed.

²¹⁷ For instance, it would be too easy for prosecutors to evade the rule if they were narrowly prohibited from bringing only a specific charge that was expressly mentioned in a threat. How much broader should the prohibition extend, though? To all charges that would satisfy the Double Jeopardy “same elements” test? Or more broadly to all transactionally related charges? Likewise, should an analog to the Double Jeopardy Dual Sovereignty Doctrine apply in this context? And what about charging threats that do not increase the potential maximum prison term (the standard metric for punishment severity), but that instead increase sentencing exposure in other respects, such as by triggering a forfeiture statute that would not otherwise be applicable?

²¹⁸ For instance, the *Apprendi* Court nowhere asked whether twelve years was, in the abstract, a just sentence for the defendant’s crime.

This test, in short, is what the Court should adopt in our hypothetical *Bordenkircher II*.

Interestingly, parallels to this rule can be found in other areas of criminal procedure law. Consider, for instance, the Court's decisions regarding a defendant's invocation of his right to remain silent at trial: the Court's protective rule prohibits the imposition of any penalty on the defendant for exercising the right, without regard to such subjective matters as severity or state of mind.²¹⁹ *Griffin v. California* offers a classic example.²²⁰ During closing arguments, the prosecutor urged the jury to draw a negative inference against the defendant because the defendant refused to testify.²²¹ The Court subsequently found the argument unconstitutional, observing that an adverse comment on the refusal to testify "is a penalty . . . for exercising a constitutional privilege. It cuts down the privilege by making its assertion costly."²²² Likewise, the proposed *Bordenkircher II* test prohibits prosecutor-imposed penalties that make assertion of the jury-trial right "costly."

That said, the approach advocated here is concededly in tension with two other lines of cases: the pre-*Bordenkircher* line of cases that approve of plea-inducements, and *Goodwin* and its progeny. Neither set of cases, however, must necessarily be overturned. *Blakely* indicates that plea-inducements are not inconsistent with "common-law values" as long as juries are at least available as a meaningful check on government over-reaching. In this spirit, the proposed test does not target plea inducements generally, but only a particular form of plea inducement that raises especially strong coercion and fairness concerns. As to *Goodwin* and its progeny,²²³ the proposed test does indeed jettison *Goodwin*'s system of presumptions and burden-shifting, but only in a narrow set of circumstances; otherwise, the *Goodwin* test still applies. The special carve-out might be justified based on the repeated assertions by the *Apprendi* Five that the jury-trial right, consistent with the Framers' intent, requires particular vigilance to prevent its erosion.²²⁴ Moreover, the proposed rule targets a particular set of

²¹⁹ For a succinct discussion of the cases, see Godsey, *supra* note 208, at 492-95.

²²⁰ 380 U.S. 609 (1965).

²²¹ *Id.* at 610-11.

²²² *Id.* at 614.

²²³ See, e.g., *Alabama v. Smith*, 490 U.S. 794 (1989).

²²⁴ See, e.g., *Jones*, 526 U.S. at 246-47. Thus, the proposed rule might not encompass threats that are not directed to a waiver of the jury-trial right, for instance, if a prosecutor threatened enhanced charges solely in order to obtain a waiver of the right to appeal, to contest a sentence enhancement, or to contest civil remedies. Likewise, charging threats directed to obtaining the defendant's testimony against another person might not be included. In practice, however, it is often impossible

circumstances in which there can be little doubt that the prosecutor's purpose was to discourage exercise of the right, thereby rendering *Goodwin*'s intent-determination inquiry superfluous.

This is not to say that the Five should necessarily avoid a broader reconsideration of the vindictiveness and plea-inducement case law, but, rather, to suggest that they can proceed incrementally, precisely as they have done in the *Apprendi* line of cases. The charging-threat issue can be addressed through a narrow holding that leaves other precedent, besides *Bordenkircher* itself, intact.

B. Should Stevens Adhere to Bordenkircher?

As the only member of the *Apprendi* Five who voted with the majority in *Bordenkircher*, Justice Stevens requires separate consideration. Should he be willing to reconsider that vote today? In fact, Stevens's opinion for the Court in *Goodwin* suggests that his views are no longer wholly in sync with *Bordenkircher*.²²⁵ Where *Bordenkircher* suggested no limitation on plea inducements, *Goodwin* seemed to contemplate some sort of wag-the-dog limitation. Moreover, the fact that Justice Powell (author of the wag-the-dog dissent in *Bordenkircher*) actually joined the *Goodwin* majority lends support to the view that the Court (and Stevens) had moved between the two opinions.

Assuming that Stevens does indeed agree with the propriety of regulating plea inducements, would he accept the (rule-structure) formalist test proposed

to disentangle the various motives behind a charging threat; usually, the prosecutor seeks the defendant's waiver of a bundle of rights together. In order to make the proposed rule meaningful, it should encompass "mixed-motive" scenarios, in which a waiver of the jury trial is part of the package sought by the prosecutor.

²²⁵ Indeed, Stevens actually seemed to move away from the most extreme reading of *Bordenkircher* the very next term with his dissent in *Corbitt v. New Jersey*, 439 U.S. 212 (1978). *Corbitt* challenged the constitutionality of a New Jersey statute that mandated life imprisonment for defendants convicted by a jury of first-degree murder, but permitted a lesser penalty for defendants who pled out. *Id.* at 215. In upholding the statute, the majority relied heavily on *Bordenkircher*. *Id.* at 221-22. Stevens, in dissent, conceded that the New Jersey statute served the same state interest as do plea inducements delivered by prosecutors, *id.* at 231 (Stevens, J., dissenting), but rejected the majority's suggestion that the basic legitimacy of this interest shielded all state action in furtherance of it. Rather, what made the New Jersey statute different than *Bordenkircher*-type threats was that the statute mandated "a different standard of punishment depending *solely* on whether or not a plea is entered." *Id.* at 232 (emphasis added). By contrast, he assumed that prosecutors would consider "individual factors relevant to the particular case," rather than making charging decisions based solely on the defendant's plea. *Id.* at 231-32. Stevens seems to be anticipating *Goodwin* and backing away from *Bordenkircher*'s open-ended deference. See *Goodwin*, 457 U.S. at 380 n.11 ("A charging decision does not levy an improper 'penalty' unless it results *solely* from the defendant's exercise of a protected legal right . . ." (emphasis added)).

here, or would he continue to prefer the wag-the-dog test of *Goodwin*? There are at least two good reasons to believe that Stevens would prefer the proposed test.

First, he embraced an analogous test in his *McMillan* dissent and the subsequent *Apprendi* cases. Recall that, in *McMillan*, Stevens acknowledged that this test would be easy to circumvent *in principle*, but he responded by arguing that political realities would prevent circumvention *in practice*. As suggested in the previous Part, similar arguments might be made in response to concerns about the circumvention of a formalist rule restricting charging threats.

Second, in the years since *Bordenkircher* and *Goodwin*, the plea-inducement playing field has tilted dramatically in favor of prosecutors, creating much more compelling risks that defendants will be effectively coerced into surrendering their trial rights. At the time *Bordenkircher* was decided, draconian sentencing statutes, like Kentucky's Habitual Criminal Act, were comparatively rare. Indeed, Kentucky's Act was actually softened not long after Paul Hayes' sentencing.²²⁶ Subsequent decades have been a different story. Foreshadowing the future direction of criminal law, Pennsylvania enacted the mandatory minimum statute at issue in *McMillan* in 1982²²⁷—the very year Stevens wrote his opinion in *Goodwin*. In the years immediately following *Goodwin*, Congress began its biennial ritual of adopting new mandatory minimums at the federal level.²²⁸ Many states followed suit, with the trend perhaps reaching its apex with the adoption of California's notorious three-strikes law in 1994.²²⁹ The upshot is that prosecutors now routinely have the ability to make the sort of extreme charging threats used against Paul Hayes. While the holding in *Bordenkircher* rested, in part, on the Court's view that prosecution and defense "arguably possess relatively equal bargaining power,"²³⁰ such an assertion would be nearly laughable today.²³¹

²²⁶ *Bordenkircher*, 434 U.S. at 371 n.3 (Powell, J., dissenting).

²²⁷ *McMillan*, 477 U.S. at 81.

²²⁸ See Sandra Guerra Thompson, *The Booker Project: The Future of Federal Sentencing*, 43 HOUSTON L. REV. 269, 271 n.13 (summarizing important federal mandatory minimum statutes).

²²⁹ FRANKLIN E. ZIMRING ET AL., PUNISHMENT AND DEMOCRACY: THREE STRIKES AND YOU'RE OUT IN CALIFORNIA ix (2001).

²³⁰ *Bordenkircher*, 434 U.S. at 667.

²³¹ See, e.g., LOWENTHAL, *supra* note 169, at 112 (describing defense lawyers' "bitter[ness]" regarding "tilt of the playing field"); STANDEN, *supra* note 166, at 1473-74 ("[P]rosecutors possess and have the incentive to exercise substantial power to overwhelm criminal defendants in the plea bargaining process.").

Stevens has been an acute and concerned observer of these trends, and made note of them in his *Apprendi* and *Booker* opinions. Indeed, in *Booker*, he framed the *Apprendi* line of cases this way:

As enhancements became greater [after *McMillan*], the jury’s finding of the underlying crime became less significant. And the enhancements became very serious indeed. . . . [T]he Court was faced with the issue of preserving an ancient guarantee under a new set of circumstances.²³²

Similarly, this same “new set of circumstances” should cause also Stevens to consider the need for modifying prosecutorial vindictiveness law in order to preserve the same “ancient guarantee.” Neither the open-ended permissiveness of *Bordenkircher* nor the nearly-as-deferential approach of *Goodwin* plausibly provide the sort of robust safeguard that Stevens later demanded in *Apprendi*.

C. More “Doctrinaire Formalism?”

Would modifying *Bordenkircher* as proposed be anything more than an exercise in “doctrinaire formalism”? In other words, would a post-*Bordenkircher II* world be any different than the present world, and, if so, would the changes be in any sense appealing ones? The previous Part set forth the basic argument that, taking into account the underlying logic of the *Apprendi* decisions, the Court could indeed regulate charging threats in a meaningful manner. With a specific proposal now on the table, this Section develops the earlier argument in a more detailed fashion.

We should begin with an account of current charging and plea-inducement practices. Characterizing such practices is a difficult task, for every prosecutorial office has its own policies and culture, and many offices leave considerable discretion in the hands of the line prosecutor. That said, one may hazard a few generalizations. First, in run-of-the-mill cases, initial charging decisions are often made by prosecutors based chiefly on information and recommendations supplied by police officers, without a great deal of independent investigation, research, or deliberation.²³³ Thus, the prosecutor may have little information about, for instance, potential defenses in the case, witness credibility issues and

²³² *Booker*, 543 U.S. at 236-37.

²³³ The rigor of prosecutorial charge-screening may vary considerably by jurisdiction. Darryl K. Brown, *The Decline of Defense Counsel and the Rise of Accuracy in Criminal Adjudication*, 93 CAL. L. REV. 1585, 1600 n.47 (2005). For a description of the charging processes of the New Orleans District Attorney’s office, which devotes an extraordinary amount of resources and care to the initial charging decision, see Ronald Wright & Marc Miller, *The Screening/Bargaining Tradeoff*, 55 STAN. L. REV. 29, 62-63 (2002).

similar sources of litigation risk, and (other than criminal history) the background and character of the defendant.²³⁴ Second, given the broad scope and many overlapping provisions of contemporary criminal codes,²³⁵ as well as the low level of proof required to initiate a criminal case,²³⁶ the prosecutor will typically have a range of permissible charging options, which will expose the defendant to greater or lesser degrees of punishment. Third, taking into account informational limitations and the need for later plea-inducement flexibility, initial charges are typically a bit harsher than what the prosecutor actually thinks would be the optimal resolution of the case.²³⁷ Fourth, the prosecutor will not typically pursue every conceivable charge in every case.²³⁸ Fifth, given that prosecutors, defense lawyers, and sometimes defendants themselves are repeat players in the system, prosecutors are usually (but not always) quite successful in setting an initial charge that permits enough downward movement to induce a plea while still producing a final result that lies within the range of what the prosecutor considers an appropriate response to the crime. Sixth, in addition to modifying charges, the prosecutor typically has a range of additional means to induce pleas, which might include, for instance, the promise of a favorable sentencing recommendation, an offer to stipulate to particular sentencing factors, or an agreement not to pursue civil remedies. Because such devices do not involve the formal filing of new charges and related transaction costs (arraignment, preliminary hearing, fresh

²³⁴ See Schulhofer, *supra* note 12, at 1983 (discussing information asymmetry between prosecution and defense). In some cases, prosecutors become actively involved in the investigative process. In such cases, prosecutors are apt to have more information available, but may be less able to assess the information in unbiased manner. Brown, *supra* note 233, at 1600. Either way, the prosecutor may initiate criminal proceedings without a realistic sense of potential weaknesses in her case. One notable exception may be in federal white-collar cases, in which defense lawyers “almost always actively attempt, from a very early stage, to influence the conclusions of the prosecutor.” Gerard E. Lynch, *Our Administrative System of Criminal Justice*, 66 *FORDHAM L. REV.* 2117, 2126 (1998).

²³⁵ See, e.g., LOWENTHAL, *supra* note 169, at 104-05 (describing charging options made possible by overlapping crime definitions in one case).

²³⁶ See *id.* at 98 (discussing probable cause standard used for charging purposes).

²³⁷ See Gazal-Ayal, *supra* note 206, at 2331 (discussing “safety margin” built into charging decisions). Professor Alschuler has provided a helpful taxonomy of “overcharging,” which includes vertical overcharging (“charging a single offense at a higher level than the circumstances of the case seem to warrant”) and different forms of horizontal overcharging (e.g., charging a defendant with a separate offense for each criminal transaction in which he participated). Albert W. Alschuler, *The Prosecutor’s Role in Plea Bargaining*, 36 *U. CHI. L. REV.* 50, 85-87 (1968).

²³⁸ Charging restraint, however, varies among jurisdictions and by type of case. See, e.g., LOWENTHAL, *supra* note 169, at 105 (describing policy of Maricopa County, Arizona, to charge most serious possible crime when gun was used). To the extent that prosecutors already routinely charge the maximum in particular categories of cases, then concerns about prosecutors “upping the ante” in response to our hypothetical *Bordenkircher II* are moot.

discovery demands, and so forth), prosecutors will often find them preferable to the tactic of threatening enhanced charges. And, finally, because plea deals can usually be reached quickly and easily within the standard high-to-low framework, plea-bargaining more frequently follows this pattern than the low-to-high approach exemplified by *Bordenkircher*. (By “standard high-to-low,” I mean that initial charges build in some bargaining room, but do not encompass every conceivable crime that could be charged. I will refer to the latter approach as “super high-to-low.”)

With these assumptions in mind, let us now consider potential prosecutorial responses to our hypothetical *Bordenkircher II*.²³⁹

1. *Option One: Do Nothing*

Prosecutors might respond to our hypothetical holding in *Bordenkircher II* by charging as they always have. This, in fact, seems a likely response in many jurisdictions. Not only is it consistent with the recognizable inertia of all social institutions, but it also reflects the fact that prosecutors are generally quite successful in setting initial charges in such a way as to induce a guilty plea. Given the range of plea-inducement tools otherwise available, prosecutors in many—perhaps most—jurisdictions are unlikely to miss the option of threatening more serious charges except in a small number of unusual cases. For many prosecutors, the potential for these occasional cases will not justify the costs of developing a systematic response to *Bordenkircher II*.

To the extent that “do nothing” is the prosecutorial response of choice, the post-*Bordenkircher II* world would not look dramatically different than the present. Most defendants would experience the criminal justice system precisely as they do now. Some defendants, however, would be spared threats of enhanced charges. In some of their cases, a plea deal will be reached anyway because the prosecutor is willing to make sufficient additional concessions to the defendant. In others, the upshot will be a trial that would not have otherwise occurred. Overall, the effects of *Bordenkircher II* would look like the effects that the *Apprendi* Five anticipated would flow from their earlier decisions: a few more trials and a bit more plea-bargaining leverage for some defendants. While such

²³⁹ While the proposed rule is not, technically, a ban on charge-threats during plea negotiations, I assume that prosecutors would treat the rule in that way. Because defense counsel would recognize that charging threats were ineffectual, such threats would not carry much plea-inducement weight; therefore, prosecutors would have little incentive to make them. Thus, while one potential drawback of the rule is that, once a charging threat was made, a prosecutor would be unable to adjust charges upward based even on legitimate reasons, such as newly discovered evidence, this should not be a serious problem in practice.

results are certainly not cost-free to society,²⁴⁰ they do seem consistent with the overarching vision of a more robust role for the jury, greater limitations on the coercive power of the state, and greater fairness, as perceived by defendants.²⁴¹

2. *Option Two: Up the Ante Across the Board*

The *Bordenkircher* and *Goodwin* majorities feared that prohibiting low-to-high would simply cause prosecutors to switch to a more aggressive version of high-to-low, upping the ante across the board with tougher initial charges. Recall that at present prosecutors do not always charge every conceivable crime in every case. Were prosecutors broadly to adopt a “super high-to-low” strategy, then our hypothetical holding in *Bordenkircher II* could indeed be accused of perverse consequences. At best, cases would simply be resolved on the same terms as they are now, albeit sometimes following a different path (high-to-low instead of the reverse). At worst, many defendants might actually be harmed in a number of respects, such as by facing greater stigma and more onerous bail conditions as a result on the higher initial charges.²⁴²

These concerns, however, should be greeted with some skepticism. First, as suggested in the previous Subsection, there are good reasons to doubt whether prosecutors would assume the transaction costs of broad, systematic responses to *Bordenkircher II*, when prosecutors can generally reach comfortable results within the standard high-to-low framework. Second, to the extent there is a systematic response, prosecutors will be wary of inflating initial charges across the board. As discussed in the previous Part, inflating initial charges may make it much harder, in light of anchoring and framing effects, for prosecutors ultimately to reach negotiated outcomes that are satisfactory to them. Moreover, in light of reciprocal fairness tendencies, prosecutors also risk a backlash of uncooperative behavior from defendants and defense counsel.

²⁴⁰ See Joseph A. Colquitt, *Ad Hoc Plea Bargaining*, 75 TULANE L. REV. 695, 704-05 (2001) (noting social benefits of plea bargains, including reduced costs and reduced time lag between the offense and the punishment, which benefits defendants subject to pretrial detention and potentially enhances deterrent effects).

²⁴¹ Even a small incremental increase in the number of trials may have great significance in a system like ours where trials have become exceedingly rare. See Wright, *supra* note 178, at 83 n.12 (2005) (noting that if acquittals become too rare, then “they cannot serve their market discipline function during plea negotiation”).

²⁴² The Court expressed precisely this concern in *Goodwin*. 457 U.S. at 380 n.10. More onerous bail conditions are perhaps of special concern; they not only increase the likelihood of pretrial detention, but may also thereby shape the ultimate disposition of the case. See Bibas, *supra* note 141, at 2492-93 (noting that pretrial detention impairs the defendant’s ability to mount a defense and, in small cases, increases pressure on the defendant to take a plea deal quickly).

Finally, to the extent that prosecutors do choose to up the ante across the board, we should recognize that this will not necessarily represent an unmitigated loss for defendants.²⁴³ While upping the ante would produce a period of instability in criminal practice, there would ultimately be a new equilibrium in which judges (certainly) and the public (possibly) would be able to recognize that charges have been inflated and discount accordingly. For instance, when judges recognize that today's aggravated assault charge was yesterday's simple assault, and today's attempted murder yesterday's aggravated assault, then judges will likely discount today's bail conditions, at least to some extent, in order to account for the change.²⁴⁴ At the same time, for reasons discussed in the previous Part, even when high-to-low and low-to-high produce the same outcomes, there is good reason to believe that the two tactics are *perceived* differently by defendants. Given the different ways that losses and gains are experienced, defendants might perceive high-to-low as fairer and less coercive. Whatever costs defendants bear, those costs should be considered in light of the benefits of perceived gains in fairness.²⁴⁵

3. *Option Three: Up the Ante Selectively*

If prosecutors do not wish to bear the potential costs of upping the ante across the board, they might instead up the ante selectively, targeting categories of defendants who tend to be particularly resistant to conventional plea inducements within the standard high-to-low framework. It is not entirely clear what categories these would be, but one is readily identifiable: defendants who believe themselves innocent, or otherwise likely to prevail at trial, are likely particularly resistant to standard plea inducements. This tendency may be explained in light of framing effects and loss-aversion; whereas the knowingly

²⁴³ Of course, there are likely some discrete categories of defendants (serious recidivists, gangsters, terrorists, and the like) against whom prosecutors are already filing all plausible charges, either for public relations purposes or out of a genuine desire to obtain the maximum possible sentence. The prosecution of such defendants should be unchanged by *Bordenkircher II*; there was never any room for threatening them with higher charges, so losing the ability to make the threat cannot affect the plea-bargaining or litigation process.

²⁴⁴ Anchoring effects, however, suggest that the discounting will not be complete.

²⁴⁵ Enhanced perceptions of fairness may produce important long-term benefits for both the defendant and the community. For instance, in an important body of theoretical and empirical work, Professor Tyler and various colleagues have identified a link between perceptions of fair treatment, beliefs in the legitimacy of legal institutions, and voluntary deference to the law and legal authorities. See, e.g., TYLER & HO, *supra* note 195, at xiv-xv, 12. Cf. *Morrissey v. Brewer*, 408 U.S. 471, 484 (1972) (“[S]ociety has a further interest in treating the parolee with basic fairness: fair treatment in parole revocations will enhance the chance of rehabilitation by avoiding reactions to arbitrariness.”).

guilty defendant likely welcomes a standard plea deal as a gain, an “innocent” defendant will see the agreement to plead guilty to any charge as a loss.

In light of these considerations, prosecutors in a post-*Bordenkircher II* world would have more of an incentive to determine before the start of plea-bargaining which defendants have a real claim of innocence or other grounds for optimism. However, an early, thorough review of the merits of the case by the prosecutor would be a welcome development.²⁴⁶ For instance, BLE scholarship suggests that early review would help prosecutors better appreciate defendant perspectives before positions harden in the charging and negotiation process and thereby facilitate the declination or voluntary dismissal of inappropriate charges.²⁴⁷

These benefits, however, might be offset by an unfortunate effect of closer prosecutorial scrutiny in the early stages of the case: aggressive prosecutors anxious to ensure adequate plea-inducement leverage might inflate the initial charges against defendants whose optimism is both strong and misplaced.²⁴⁸ In these circumstances, whatever negative consequences flow from the super high-to-low tactic (e.g., tougher bail condition, greater stigma) might effectively be distributed on the basis of the sorts of circumstances that cause defendants to take an unjustifiably optimistic view of their situation, such as incompetent legal counsel, poor cognitive functioning, or simple inexperience with the criminal justice system—all circumstances unrelated to the actual severity of the crime.

While it is possible to identify some categories of defendants as to whom prosecutors may wish to inflate initial charges, note that prosecutors already have incentives to be unusually aggressive in charging these defendants. At present,

²⁴⁶ See Wright & Miller, *supra* note 233, at 95 (noting benefits of pre-charge screening by prosecutors).

²⁴⁷ See Burke, *supra* note 169, at 1614-15 (“Because the theory of guilt triggers sources of cognitive bias, prosecutorial neutrality should be at its peak prior to the prosecutor’s charging decision”); Stephanie Stern, *Cognitive Consistency: Theory Maintenance and Administrative Rulemaking*, 63 U. PITT. L. REV. 589, 602-620 (2002) (discussing empirical evidence supporting theory of cognitive consistency, which predicts tendency to hold on to beliefs in face of disconfirming evidence, particularly where there has been public commitment to belief). A leading criticism of plea-bargaining is that it leads to the conviction of innocent defendants. Gazal-Ayal, *supra* note 206, at 2297. As Professor Gazal-Ayal has recently pointed out, the root cause of the problem is that the strength of plea inducements leave prosecutors with insufficient incentives to screen out weak cases at the charging stage. *Id.* at 2298-99. Any marginal decrease in the prosecutor’s plea-inducement leverage, as by restricting the use of charging threats, should result in some marginal increase in the incentives to screen better.

²⁴⁸ Professor Schulhofer has also argued that, if prosecutors could overcome information barriers to determine which defendants truly believe themselves innocent, they would respond with a more aggressive approach to plea inducement. Schulhofer, *supra* note 12, at 1984.

however, prosecutors may resist these incentives, knowing that they retain the ability to threaten increased charges if the usual plea inducements prove inadequate. Our hypothetical holding in *Bordenkircher II* would cancel the insurance policy, leaving prosecutors somewhat more likely to adopt the super high-to-low strategy. Such a change might indeed impose costs on some defendants. Other groups of defendants, however, may benefit, particularly to the extent that closer pre-charging examination of cases leads prosecutors to decline the more dubious ones up front. Moreover, as discussed in the previous Subsection, costs also need to be weighed in light of perceptions of greater fairness in a system that does not employ explicit threats in order to extract waivers of constitutional rights.

4. *Option Four: Engage in More Pre-Charge Bargaining*

If a prosecutor fears the loss of a *post-charging* plea-inducement tool (the threat of higher charges), then another logical response would be to secure the plea agreement *before* filing charges. At this point, charges are still inchoate, and defendants may not have any clear expectation of what they will be; it is accordingly difficult to speak in terms of a “threat to increase charges,” and our hypothetical *Bordenkircher II* rule would be inapplicable.

There are good reasons, however, to doubt that prosecutors would routinely conclude plea deals before filing charges. For one thing, no right to counsel attaches before the initiation of adversary proceedings,²⁴⁹ and prosecutors will not relish the inefficiencies and potential misconduct claims arising from direct negotiation with an unrepresented lay defendant.²⁵⁰ For another, defendants held in custody are generally entitled to a probable cause determination within 48 hours of arrest.²⁵¹ There are obvious efficiency benefits to combining this determination with the arraignment process, but doing so may leave little time for the prosecutor to negotiate a pre-charge deal. In light of these sorts of constraints, increased pre-charge bargaining seems most likely in two types of cases: (1) white-collar cases in which the defendant is capable of securing counsel prior to the initiation of adversary proceedings,²⁵² and (2) routine, high-

²⁴⁹ *United States v. Gouveia*, 467 U.S. 180 (1984).

²⁵⁰ *See, e.g., Wright & Miller, supra* note 233, at 78-79 (explaining absence of pre-charge bargaining in New Orleans based on lack of public defender availability between bail hearing and arraignment).

²⁵¹ *County of Riverside v. McLaughlin*, 500 U.S. 44 (1991).

²⁵² *See Lynch, supra* note 234, at 2126 (discussing routine pre-indictment contact between prosecutors and defense counsel in white-collar cases).

volume cases for which prosecutors may develop standard, take-it-or-leave-it plea offers, such as those used by many United States Attorneys' Offices for illegal reentry cases.²⁵³

Increased pre-charge bargaining would not necessarily be either unfair to defendants or unwelcome as a matter of policy. An early deal compresses the period of uncertainty faced by the defendant, and may lessen the likelihood that defendants will develop firm expectations as to how the case will be handled that are later dashed. Additionally, because an early deal relieves the prosecutor of various procedural burdens, such as the need to obtain an indictment or show probable cause at a preliminary hearing, the prosecutor may be willing to pay a premium for the defendant's agreement to plead guilty.²⁵⁴ While an early deal may mean that the defendant (or the prosecutor, for that matter) may lack important information that would later come to light, this cost may be offset by the benefit of negotiations that occur outside the shadow of the cognitive bias that arises when the prosecutor publicly commits to a particular set of charges and a particular theory of the case.²⁵⁵ Finally, to the extent that early plea deals reduce the transaction costs borne by prosecutors, judges, and public defenders, more resources will be available for other purposes (including the potential increase in trials resulting from a prohibition on threats of increased charges).

5. *Option Five: Do More Sentence-Bargaining*

If prosecutors lose some flexibility in using their charging power for plea-inducement purposes, some may respond with greater use of sentence-related inducements.²⁵⁶ Depending on the particulars of sentencing law and practice in the jurisdiction, such inducements may take any of a number of different forms. For instance, in order to induce a plea, a prosecutor may offer to stand silent at sentencing or to recommend a particular sentence that the defendant would view as a favorable outcome. Alternatively, in some jurisdictions, a plea agreement may be made contingent on a specific sentence or sentencing range.²⁵⁷ Or, more modestly, a prosecutor might offer to stipulate to a specific sentencing factor,

²⁵³ See O'Hear, *supra* note 16, at 789 (describing federal "early disposition" programs).

²⁵⁴ For instance, in the federal early disposition programs, defendants receive a special "downward departure" under the federal sentencing guidelines. *Id.*

²⁵⁵ See Stern, *supra* note 247, at 640-43 (providing justification based on cognitive consistency theory for analogous practice of "regulatory negotiation," in which stakeholders negotiate over a proposed rule before rule is formally published).

²⁵⁶ To be sure, charging-related inducements affect the ultimate sentence by triggering a particular statutory sentencing range. By "sentence-related inducements," I mean inducements that affect the selection of a sentence within that statutory range.

²⁵⁷ See, e.g., FED. R. CRIM. P. 11(c)(1)(C).

such as amount of drugs or mitigating role in the offense, or agree not to seek a sentence enhancement on the basis of a particular factor. In all such forms of “sentence-bargaining,” the prosecutor makes an offer or threat relating to the selection of a specific sentence within the applicable statutory maximum.

Because sentence-bargaining is already common in many jurisdictions,²⁵⁸ and not a practicable alternative in others,²⁵⁹ it is far from clear that a ban on *Bordenkircher*-style threats would meaningfully increase frequency of the practice. To the extent the practice increased, some defendants might prefer the change, which would potentially leave them with a perception of greater control over the sentencing process and greater certainty as to the outcome. On the other hand, there is cause for legitimate concern over sentence-bargaining. The practice diminishes the visibility, and hence accountability, of the sentencing process, and may lead to unwarranted sentencing disparities between similarly situated offenders. Additionally, threats to seek sentence enhancements, even within a given statutory maximum, may be no less coercive and perceived as no less unfair than threats to seek an increased maximum. It is not clear, then, that defendants would gain from a switch on the margins from *Bordenkircher*-style charge-bargaining to similar threat-based *sentence*-bargaining. At the same time, there seems no compelling reason to conclude that a switch from charge threats to sentence threats would constitute a *loss* for defendants.

6. *Option Six: Make Implicit Threats*

If prosecutors lose the ability to make express charging threats, then some might attempt to make implicit threats, as by routinely pursuing additional charges against defendants who refuse plea deals. The prosecutor would hope to develop a reputation for this practice among defense lawyers, so that defendants would be routinely counseled to take the prosecutor’s offer. This would be an

²⁵⁸ See, e.g., Stephen J. Schulhofer & Ilene H. Nagel, *Plea Negotiations Under the Federal Sentencing Guidelines: Guideline Circumvention and Its Dynamics in the Post-Mistretta Period*, 91 NW. U. L. REV. 1284, 1285 (1997) (estimating that plea bargaining results in circumvention of federal sentencing guidelines in 20 to 35 percent of cases).

²⁵⁹ In jurisdictions without sentencing guidelines and with a strong tradition of judicial control of sentencing, for instance, the prosecutor’s sentencing-related bargaining chips will have little value. See KATE STITH & JOSÉ A. CABRANES, FEAR OF JUDGING: SENTENCING GUIDELINES IN THE FEDERAL COURTS 80 (1998) (“[In the pre-guidelines federal system] in many districts . . . prosecutors generally refrained from rendering specific sentencing recommendations to the judge, and the judge would neither elicit nor condone such recommendations.”); Wright & Miller, *supra* note 233, at 79-80 (describing absence of sentence bargaining in New Orleans); King, *supra* note 192, at 136 (“Judges more readily reject sentence agreements than charge bargains, considering the selection of an appropriate sentence to be within their special domain . . .”).

unfortunate practice, perhaps even worse than an express *Bordenkircher*-style threat, which at least has the virtue of giving the trial-bound defendant a clear opportunity to avoid the “threatened” charge.

There are some natural checks on this practice. The transaction costs of new charges may deter some prosecutors from routine charge enhancements. Others will avoid the practice because it offends their sense of fairness. Still others will be reluctant to subvert a clearly articulated constitutional norm of no threats.

In order to buttress these tendencies, however, courts should be willing to entertain “implicit threat” claims from defendants who were subject to pretrial charge increases. These claims might proceed along two lines. First, the prosecutor might be required to provide a neutral explanation for the charge increase. For reasons described in an earlier Section, however, this sort of requirement is not likely to provide a meaningful check on prosecutors. Thus, defendants should be permitted to present evidence of past practices so as to establish a pattern of routine charge enhancements, either on the individual prosecutor level or the office level. Unfortunately, there is no obvious line to be drawn here, and courts will not have an easy time deciding when the rate of charge enhancements reaches a level sufficient to demonstrate “a realistic likelihood of vindictiveness.”²⁶⁰ Just permitting the claims, however, may have a salutary effect, as prosecutors will doubtless wish to avoid coming close to the line and thereby prompting burdensome collateral litigation.²⁶¹ Prosecutors could easily do so by exercising care in their initial charging decisions and making a habit of standing by those initial decisions. (And, as indicated above, improved initial charging is both viable and desirable on a number of grounds.²⁶²)

7. Summary

In weighing the actual consequences of the hypothetical holding in *Bordenkircher II*, it may be helpful to distinguish among three categories of defendants: (1) those who would not have been subject in any event to *Bordenkircher*-style charging threats; (2) those who would have been subject to such threats and pled guilty; and (3) those who would have been subject to such

²⁶⁰ See *Goodwin*, 457 U.S. at 223 (Brennan, J., dissenting).

²⁶¹ I assume that the defendant would have to make some sort of minimal threshold showing in order to trigger full-blown litigation of the issue. This showing should not, however, disingenuously require the defendant effectively to carry his burden on the merits before getting discovery, as the Court has unfortunately required in an analogous context. See *United States v. Armstrong*, 517 U.S. 456 (1996).

²⁶² See *Wright & Miller*, *supra* note 233, at 32-34 (arguing that early and careful assessment of cases is both viable and attractive).

threats and gone to trial. Those in the first category will not likely be much affected by the holding. There may be some adverse consequences on the margins to the extent that prosecutors respond by upping the initial charging ante. However, these effects are not likely to be profound or widespread, and may be offset by the benefits of, for instance, increased pre-charge investigation and bargaining.

Those in the second category will be spared charging threats. Some will go to trial as a result. Of these, some will win, and some will lose; and, of those who lose, some may actually be worse off, from a sentencing standpoint, than if they had taken advantage of the plea deal offered to them.²⁶³ Still, all who go to trial will have the advantage of a process they preferred to a bargained resolution, and all should perceive greater fairness than in a system that promises a right to a trial but expressly penalizes those who invoke it.

Some in the second category will opt, even in the absence of a charging threat, to plead guilty, perhaps as a result of sentencing threats, enhanced initial charges, or a simple failure of nerve. On the whole, one imagines that these defendants will be better off insofar as *Bordenkircher II* removes one, sometimes quite powerful, device from the prosecutor's plea-inducement toolbox. While some defendants may suffer a net loss (e.g., by facing higher initial charges), there are good reasons to doubt that such results will be widespread; in general, a prosecutor's loss of one source of plea-inducement leverage should strengthen, not weaken, the defendant's negotiating position.

Those in the third category should most clearly benefit from *Bordenkircher II*. These defendants are bound for trial regardless of the prosecutor's best efforts to the contrary; the only question is what charges they will face. Under *Bordenkircher II*, those charges will often be less than they would have been under *Bordenkircher I*. To be sure, some will face the same charges they would have otherwise faced because the prosecutor will have upped the ante in the initial charges. But, again, there are good reasons to doubt that such results will be universal. Many category-three defendants will unambiguously benefit from *Bordenkircher II*, and few will unambiguously suffer.²⁶⁴

²⁶³ The judge, for instance, may impose a "trial penalty" on the defendant at sentencing. See, e.g., Nancy J. King et al., *When Process Affects Punishment: Differences in Sentences After Guilty Plea, Bench Trial, and Jury Trial in Five Guidelines States*, 105 COLUM. L. REV. 959, 961-62 (2005).

²⁶⁴ Some category-three defendants may suffer a net loss if the prosecutor makes sentencing threats that the prosecutor would not have made if charging threats had been available, the defendant loses at trial, and the prosecutor's delivery on the sentencing threats results in a longer sentence than would have otherwise been imposed. One imagines, however, that these cases would be rare. For

These defendants, moreover, are defendants for whom we should feel a particular solicitude. By their determination to go to trial, the category-three defendants are likely signaling one or more of the following conditions: a firm belief in their innocence, unusual optimism regarding their chances of winning at trial, extraordinary bullheadedness, ineffective legal counsel, and/or an expectation that the results of a conviction will be especially onerous. Under *Bordenkircher I*, as the dissenters in the case suggested,²⁶⁵ defendants are effectively penalized for these sorts of circumstances, even though they have no legitimate bearing on the defendants' blameworthiness. Indeed, if anything, the defendant's firm belief in his own innocence should raise serious questions about the defendant's culpability.

In sum, there are good reasons to believe that *Bordenkircher II* will have real effects on the outcomes of some cases; this is not a matter of "doctrinaire formalism" in the sense that any potential consequences of the holding will be swamped by prosecutorial circumvention. Nor is it a matter of doctrinaire formalism in the perverse consequences sense. While one must concede a likelihood that some defendants will suffer worse outcomes as a result of the holding, there are good reasons to doubt that these effects will be widespread. Other defendants should be unambiguously better off, and some of these are among the most vulnerable in the system.

There may also be broader, if more subtle, benefits. The hypothetical holding in *Bordenkircher II* would provide prosecutors with additional incentives to do more investigation and negotiation before filing charges; this may promote more judicious charging decisions, more efficient negotiation and litigation processes, and reduced uncertainty for defendants and victims. *Bordenkircher II* would also invite prosecutors to reconsider whether they overvalue the goal of speedy convictions, and perhaps contribute to stronger due process norms in connection with plea-inducement. Finally, *Bordenkircher II* would signal defendants that the system takes constitutional rights seriously and that even prosecutors operate in a rule-bound fashion; this may strengthen respect for the system and promote compliance with the rules that the system lays down for defendants.²⁶⁶

instance, it is not clear why a prosecutor who was willing to use charging threats under *Bordenkircher I* would withhold sentencing threats, on which the transaction costs of follow-through are often less.

²⁶⁵ *Bordenkircher*, 434 U.S. at 366 n.2 (Blackmun, J., dissenting).

²⁶⁶ See *supra* note 245. While this Article most directly addresses the *Apprendi* Five, the same considerations that would justify their overturning *Bordenkircher* would also justify experimental

V. CONCLUSION

Apprendi's critics can be forgiven for using the "doctrinaire formalism" label: there *is* a real gap in the *Apprendi* decisions between the rhetoric and the rule. The decisions amount to incrementalism disguised as absolutism. They give the appearance of absolutism because of their frequent invocation of a romantic ideal of the jury as a check on government oppression, because of their express rejection of efficiency and uniformity as legitimate grounds on which to limit access to juries, and because their legal test permits no prudential case-by-case balancing of interests. Yet, this apparent absolutism is an illusion. The decisions allow, indeed expressly contemplate, both discretionary sentencing by judges and negotiated guilty pleas. In a world in which these practices are allowed to persist, jury fact-finding for sentencing purposes will always be the exception, not the norm.

The Court's failure to live up to its absolutist rhetoric does not necessarily mean that the Court is failing to advance the basic values embraced by the rhetoric. And the *Apprendi* decisions do indeed provide for at least incremental progress. The decisions limit the range of legislative options in structuring sentencing systems. Denied their most preferred choice, many states will (indeed, many already have) adopt jury fact-finding. And while many defendants will surrender their right to such fact-finding, not all will. *Apprendi* and its progeny should therefore result in at least an incremental increase in the reliance on juries for sentencing purposes. And in those cases in which defendants do bargain away their jury rights, they will often be able to obtain additional concessions in the process, thereby marginally diminishing prosecutorial domination of the system. In jurisdictions that do not adopt jury fact-finding, the likely response will be discretionary sentencing. While a switch to discretionary sentencing does not necessarily enhance democratic control, it also incrementally advances the ideal of checks and balances in the exercise of state power. In short, while the *Apprendi* decisions do not *require* states to adopt sentencing systems that fully embody "common law values," they at least steer states in the right direction.

In this Article, I have suggested how this type of incrementalism, in the service of the same underlying values, might lead to a new approach to the

regulation of charging threats by state and local legislatures and executive authorities. Such experiments would offer the ancillary benefit of empirical data to aid in the assessment of costs and benefits of different regulatory approaches. Such data, in turn, may contribute to the elaboration or modification of the constitutional rule proposed here. *Cf. Booker*, 543 U.S. at 237 (noting Court's need to adjust to "new circumstances" in order to preserve jury-trial right).

regulation of *Bordenkircher*-style charging threats. The goal here is not to steer legislatures in the right direction, but prosecutors. In some cases, the prosecutor's most preferred choice is to charge the defendant one way and then threaten enhanced charges in order to extract a guilty plea. Denied the opportunity to do so, prosecutors may respond in any of a number of different ways, some considerably more attractive than others. On the whole, though, there is good reason to believe that the rule proposed here will result in an incremental increase in the number of jury trials, an incremental decrease in the scope of prosecutorial domination, and an enhanced perception among defendants that the system operates in a fair, predictable, and respectful manner.

One drawback to incrementalism, of course, is that the line-drawing will always seem a bit arbitrary. There is always an argument that some *other* increment best strikes the balance between protecting the right at issue, minimizing disruptions, and preserving flexibility. While a number of such arguments may be made in the present context, one in particular stands out as requiring some commentary. The most troubling limitation to the Court's incrementalism in the *Apprendi* line of cases is the exception for mandatory minimums, as confirmed in *Harris*.²⁶⁷ For purposes of symmetry, I have, with some reluctance, incorporated the same exception into the proposed *Bordenkircher II* rule.

The *Harris* exception is unfortunate, but not such a gaping loophole as to wholly undermine the *Apprendi* rule. While some jurisdictions may adopt more mandatory minimums, or convert sentencing guidelines into mandatory minimums, in response to the *Apprendi* decisions, not all jurisdictions have done so or (in light of prison budget concerns, if nothing else) are likely to do so.²⁶⁸ Moreover, the *Harris* exception may be exploited only to the extent that the mandatory minimum lies within the statutory maximum. The *Harris* exception thus carries its greatest significance in cases in which the offense of conviction provides a generous statutory maximum, but has much less room to operate in the context of less serious offenses. For instance, in *Bordenkircher* itself, although the Habitual Criminal Act was framed as a mandatory minimum, it would not have escaped the proposed charging threats rule because the minimum (life) far exceeded the maximum for the underlying offense (ten years).

²⁶⁷ See *supra* Part II.B.

²⁶⁸ See Wright, *supra* note 161, at 408-09 (discussing reasons that states "have shown relative restraint when it comes to mandatory minimum sentences"). Moreover, evasion efforts based on the conversion of guidelines into mandatory minimums may run into a number of serious constitutional objections. Berman, *Tweaking*, *supra* note 170, at 360-62.

That said, it would be preferable to reject the *Harris* exception in both the sentencing factor and the charging threats context.²⁶⁹ Indeed, this would not really be inconsistent with the views of the *Apprendi* Five, four of whom dissented in *Harris*, but only inconsistent with the views of Justice Scalia. He should reconsider.²⁷⁰ The political constraints on the adoption of mandatory minimums are not nearly as reassuring as the political constraints on the adoption of inverted sentencing. Mandatory minimums further empower prosecutors, especially when they can be triggered through the relaxed procedures available for “sentencing factors,” which is inconsistent with the checks-and-balances ideal. And *Blakely*’s fairness objections to out-of-the-blue sentence enhancements seem no less compelling as to minimums than as to maximums.

The plea-inducement system—what Judge Lynch aptly terms “our administrative system of criminal justice”²⁷¹—is likely with us for the long haul, and that is not necessarily a bad thing. The challenge is to develop legal rules so as to bring some semblance of checks and balances to the system, to dispel the perception (and sometimes the reality) of momentous decisions about human liberty being dictated by prosecutors according to their own whims, biases, and personal convenience.²⁷² With or without a mandatory minimum exception, overturning *Bordenkircher* would be a very good place to start.

²⁶⁹ Fortunately, as other commentators have observed, in light of Justice Breyer’s reluctant concurrence in *Harris* and other recent developments, there is good reason to believe that the decision will not stand for long. *Id.* at 359-60.

²⁷⁰ *Cf.* *Lawrence v. Texas*, 539 U.S. 558, 604 (2003) (Scalia, J, dissenting) (noting obligation of judges to “carry things to their logical conclusion”).

²⁷¹ Lynch, *supra* note 234, at 2118.

²⁷² See O’Hear, *supra* note 16, at 805-811 (discussing unchecked prosecutorial discretion as a form of dignitary harm to defendants); Barkow, *supra* note 211, at 1050 (arguing that analysis of plea bargaining should focus more on balance of power issues).