

Article

Judicial Fact-Finding and Sentence Enhancements in a World of Guilty Pleas

Stephanos Bibas[†]

CONTENTS

I.	THE SUPREME COURT’S CASE LAW ON ELEMENTS OF CRIMES	1102
A.	<i>Burdens of Proof, Elements, and Affirmative Defenses</i>	1103
B.	McMillan.....	1105
C.	Almendarez-Torres	1107
1.	<i>Facts and Procedural Background</i>	1107
2.	<i>The Supreme Court Majority Opinion</i>	1108
3.	<i>The Dissent</i>	1109
D.	Monge	1110
E.	Jones.....	1111
1.	<i>Facts and Procedural Background</i>	1111
2.	<i>The Supreme Court Majority Opinion</i>	1112
3.	<i>Concurrences</i>	1114
4.	<i>The Dissent</i>	1114
F.	Appendi.....	1115
1.	<i>Facts and Procedural Background</i>	1115
2.	<i>The Supreme Court Majority Opinion</i>	1117
3.	<i>Concurrences</i>	1119

[†] Associate Professor, University of Iowa College of Law (bibas@philo.org). I would like to thank Akhil Amar, Ian Ayres, Steven Duke, George Fisher, Alon Harel, Jerry Israel, Christine Jolls, Dan Kahan, Susan Klein, Russell Korobkin, John Langbein, Renée Lettow Lerner, Andrew Levine, Jon Molot, Carol Rose, Anjan Sahni, Steve Sheppard, Kate Stith, and Eugene Volokh for their advice, input, and comments on earlier drafts. Any errors remain my own.

4. <i>The Principal Dissent</i>	1120
5. <i>Justice Breyer's Dissent</i>	1122
II. EXAMINING THE ELEMENTS RULE ON ITS OWN TERMS	1123
A. <i>Historical Arguments</i>	1123
1. <i>The English and Colonial Evidence</i>	1123
2. <i>Nineteenth-Century American Authorities</i>	1128
B. <i>Protecting Juries from the Slippery Slope</i>	1133
C. <i>Notice</i>	1139
D. <i>Practical Problems at Trial and on Habeas Corpus</i>	1142
III. THE REAL WORLD OF GUILTY PLEAS.....	1148
A. <i>Constitutional Values in a World of Guilty Pleas</i>	1151
1. <i>Opportunity To Be Heard</i>	1152
2. <i>Checking Arbitrariness</i>	1168
3. <i>Conscience of the Community</i>	1170
4. <i>Notice</i>	1173
B. <i>Solutions Tailored to the World of Guilty Pleas</i>	1174
1. <i>Notice of Maximum Sentences</i>	1174
2. <i>Procedural Protections at Sentencing</i>	1177
3. <i>Historic Elements and Facts That Create Felonies</i>	1180
IV. CONCLUSION.....	1183

“Criminal procedure is, basically, a subset of constitutional law,”¹ according to the reigning academic orthodoxy. Preoccupied with these lofty pretensions, criminal procedure scholars do not deign to integrate their field with its siblings, substantive criminal law and especially sentencing. Furthermore, the reigning academic orthodoxy is preoccupied with jury trials, making them the center of attention and devoting countless articles to them.²

This Article challenges and undermines the orthodox approach, exposing its inadequacies. The vehicle for this critique is a case study of an academic proposal that the Supreme Court recently made into law. In recent years, many commentators have proposed variants of what I call “the elements rule.” The elements rule holds that any fact that increases a defendant’s statutory maximum sentence must be an element of the offense. These facts must therefore be charged in an indictment and proved to a jury beyond a reasonable doubt.³

1. William J. Stuntz, *The Uneasy Relationship Between Criminal Procedure and Criminal Justice*, 107 YALE L.J. 1, 6 (1997). In fairness to Professor Stuntz, I should note that his article goes on to reject the artificial separation of criminal procedure from substantive criminal law, sentencing, and the funding of defense counsel. His is one of the very few articles to do so.

2. *Infra* notes 324-329 and accompanying text.

3. E.g., Nancy J. King & Susan R. Klein, *Essential Elements*, 54 VAND. L. REV. (forthcoming May 2001) (manuscript at 4-23, on file with author) (endorsing the elements rule); Mark D. Knoll & Richard G. Singer, *Searching for the “Tail of the Dog”: Finding “Elements” of Crimes in the Wake of McMillan v. Pennsylvania*, 22 SEATTLE U. L. REV. 1057, 1112, 1118 (1999) (same); Benjamin J. Priester, *Sentenced for a “Crime” the Government Did Not Prove: Jones v. United States and the Constitutional Limitations on Factfinding by Sentencing Factors Rather than Elements of the Offense*, 61 LAW & CONTEMP. PROBS. 249, 297 (1998) (same); Benjamin E. Rosenberg, *Criminal Acts and Sentencing Facts: Two Constitutional Limits on Criminal Sentencing*, 23 SETON HALL L. REV. 459, 502 (1993) (proposing a version of the elements rule limited to offense characteristics); Richard G. Singer & Mark D. Knoll, *Elements and Sentencing Factors: A Reassessment of the Alleged Distinction*, 12 FED. SENTENCING REP. 203, 206 (2000) (endorsing the elements rule and suggesting that it should also extend to facts that increase stigma, such as racial bias); Todd Meadow, Note, *Almendarez-Torres v. United States: Constitutional Limitations on Government’s Power To Define Crimes*, 31 CONN. L. REV. 1583, 1604-05 (1999) (endorsing the elements rule); Note, *Awaiting the Mikado: Limiting Legislative Discretion To Define Criminal Elements and Sentencing Factors*, 112 HARV. L. REV. 1349, 1361-62 (1999) (proposing the elements rule with exceptions for recidivism and other previously litigated facts); see also Susan N. Herman, *The Tail That Wagged the Dog: Bifurcated Fact-Finding Under the Federal Sentencing Guidelines and the Limits of Due Process*, 66 S. CAL. L. REV. 289, 337 (1992) (proposing procedural protections of trial, including the right to a jury, for all offense characteristics raising the maximum sentence by six months or more); Colleen P. Murphy, *Jury Factfinding of Offense-Related Sentencing Factors*, 5 FED. SENTENCING REP. 41 (1992) (proposing that juries find enhancement facts even if they are labeled sentencing factors). *Contra* Jacqueline E. Ross, *Unanticipated Consequences of Turning Sentencing Factors into Offense Elements: The Apprendi Debate*, 12 FED. SENTENCING REP. 197, 198-202 (2000) (arguing that the elements rule will prejudice defendants at trial, lead to collusive bargaining over sentencing facts, encourage legislative circumvention, and breed uncertainty and litigation); see also Roberta Sue Alexander, Note, *Dueling Views of Statutory Interpretation and the Canon of Constitutional Doubt: Almendarez-Torres v. United States*, 118 S. Ct. 1219 (1998), 24 DAYTON L. REV. 375, 395-96 (1999) (suggesting, in passing, that the elements rule for recidivism enhancements was foreclosed by precedent and would be unfair to defendants).

Many jurisdictions had allowed judges to find these facts by a preponderance of the evidence. But in a series of cases culminating last June, the Supreme Court overturned this settled practice. In *Apprendi v. New Jersey*, the Court adopted the elements rule.⁴ It carved out an exception for recidivism, and it did not apply the indictment requirement to state cases. With these exceptions, all facts that increase statutory maxima must now be charged in indictments and proved to juries beyond a reasonable doubt.⁵

At first blush, the elements rule seems favorable to defendants. It limits sentence enhancements, safeguards jury fact-finding against judicial encroachment, and warns defendants of the sentence enhancements they will face. Thus, one might applaud this rule as a counterweight to harsh sentence enhancements, a victory for defendants' constitutional rights. For similar reasons, almost every commentator who has considered the elements rule has endorsed some form of it.⁶ On the other hand, the dissent pointed out that the rule will hurt defendants at trial by putting inflammatory evidence and inconsistent defenses before juries.⁷

These ways of looking at the elements rule, however, assume that jury trials are the norm. The thesis of this Article is that the Court and commentators reached the wrong answer because they asked the wrong question. They looked only at the few defendants who go to trial, instead of considering that the vast majority plead guilty. We now live in a world of guilty pleas, not trials, and in this world the elements rule will have counterintuitive and perverse results. Ultimately, this Article concludes, *Apprendi* is symptomatic of criminal procedure's preoccupation with jury trials at the expense of the real world of guilty pleas and sentencing.

In the real world, the elements rule will hurt many of the defendants it purports to help by undermining the procedural values it purports to safeguard. First, the elements rule in effect deprives many defendants of sentencing hearings, the only hearings they were likely to have. By making important factual disputes elements of crimes, it forces defendants to surrender sentencing issues such as drug quantity when they plead guilty. Because for many defendants going to trial is not a desirable option, they are left without any real hearings at all.⁸ Second, the elements rule chops up crimes and creates more statutory maxima, which permits more arbitrariness. It gives prosecutors more opportunities in plea bargaining to cap the sentences of favored defendants by charging them with offenses

4. *Apprendi v. New Jersey*, 120 S. Ct. 2348, 2365 & n.19 (2000).

5. *Id.* at 2355. For an explanation of why *Apprendi*'s holding technically does not reach the indictment question, though its reasoning does, see *infra* note 150. See generally *infra* Sections I.E, I.F (discussing *Jones* and *Apprendi*).

6. *Supra* note 3.

7. *Infra* text accompanying notes 295-300.

8. *Infra* Subsection III.A.1.

with low statutory maxima. Defendants whom prosecutors dislike, in contrast, will face unequal, harsher treatment.⁹ Third, the elements rule reduces the role of the community's conscience. The Jury Clauses were meant to ensure a democratic, populist check on the administration of justice. The elements rule, however, leaves the power of sentencing commissions in place while making it much harder for democratic legislatures to guide this power. Thus, the rule will give unchecked, unguided, undemocratic power to unelected sentencing commissions.¹⁰ Fourth, the rule is not tailored to give defendants the notice they need before they plead guilty, namely, notice of the maximum sentences they face.¹¹

What the Court should have done instead is to translate the Constitution's values into the world of guilty pleas. The Due Process and Equal Protection Clauses check arbitrariness and promote equal treatment. The Due Process Clauses also guarantee notice and an opportunity to be heard. The Confrontation and Compulsory Process Clauses prescribe procedures for these hearings. The Grand Jury and Notice Clauses give notice to defendants of the charges they face. And the Grand and Petit Jury Clauses check arbitrariness and give voice to the conscience of the community. We must implement these procedural values within our current legal landscape of guilty pleas and sentencing. Asking the question this way leads to different solutions, such as notice of sentence enhancements before defendants plead guilty and procedural protections at sentencing.¹²

Part I of this Article traces the evolution of the Supreme Court's doctrine on elements of crimes, culminating in *Apprendi*. Part II critiques the Court's and commentators' reasoning on their own terms. Section II.A contends that the history on which the Court relies is inadequate to support its sweeping rule. Section II.B discusses the Court's and commentators' concerns with preventing the erosion of jury trials. It shows that the elements rule does little to fix this slippery slope because the rule is so easy to evade. Section II.C argues that while notice to defendants is a worthy goal, the rule is not well tailored to that goal. Section II.D concludes that the elements rule will cause problems at trial and on habeas corpus. The law's obsession with trials, however, caused the majority, dissenters, and commentators to overlook the much bigger problems with guilty pleas and sentencing.

The casual reader may wish to skip Parts I and II and go directly to Part III. Part III analyzes the elements rule in the real world of guilty pleas. Section III.A argues that, in practice, the rule is likely to undercut many of

9. *Infra* Subsection III.A.2.

10. *Infra* Subsection III.A.3.

11. *Infra* Subsection III.A.4.

12. *Infra* Section III.B.

the procedural values it purports to uphold. It will deprive defendants of important sentencing issues, promote arbitrariness by prosecutors, and reduce the role of the conscience of the community. While the rule addresses legitimate concerns, such as the need for notice, it is not tailored to the real world of guilty pleas. Section III.B builds an alternative approach for this real world. Subsection III.B.1 proposes informing defendants of their enhanced maximum sentences before they plead guilty, thus extending the Court's guilty-plea case law. Subsection III.B.2 suggests more procedural protections at sentencing, including compulsory process and some confrontation and cross-examination. Subsection III.B.3 proposes limiting the rigid elements rule to (1) historic elements of common-law crimes, and (2) facts that turn misdemeanors into felonies. These measures would address the Court's and commentators' core concerns without wreaking havoc in the real world of guilty pleas and sentencing. This Article concludes by discussing how criminal procedure scholarship should move beyond its preoccupation with trials at the expense of considering guilty pleas and sentencing.

I. THE SUPREME COURT'S CASE LAW ON ELEMENTS OF CRIMES

Over the last three decades, the Supreme Court has struggled to explain which facts are elements of crimes and which are sentencing factors. Elements must be charged in an indictment and proved beyond a reasonable doubt to a jury.¹³ Sentencing factors, in contrast, are entrusted to the sentencing judge under a lower standard of proof.¹⁴

The Court's case law in this field, however, has hardly been a model of clarity. On the one hand, the Court has repeatedly recognized that legislatures have historically had broad latitude to define crimes and punishments.¹⁵ It has further stressed that judges have historically had broad latitude to find facts and exercise discretion at sentencing.¹⁶ On the other hand, it has said there must be some constitutional limits on the legislative prerogative to define elements.¹⁷ Until *Apprendi*, the Court had never clearly enunciated what those limits were.

13. *Apprendi v. New Jersey*, 120 S. Ct. 2348, 2356 (2000); *Jones v. United States*, 526 U.S. 227, 243 n.6 (1999).

14. *McMillan v. Pennsylvania*, 477 U.S. 79, 84-86, 91 (1986).

15. *E.g.*, *Staples v. United States*, 511 U.S. 600, 604 (1994); *Payne v. Tennessee*, 501 U.S. 808, 824 (1991); *McMillan*, 477 U.S. at 85; *Liparota v. United States*, 471 U.S. 419, 424 (1985).

16. *E.g.*, *Witte v. United States*, 515 U.S. 389, 397-98 (1995); *Nichols v. United States*, 511 U.S. 738, 747 (1994); *Payne*, 501 U.S. at 820-21; *Williams v. New York*, 337 U.S. 241, 246 (1949).

17. *E.g.*, *Patterson v. New York*, 432 U.S. 197, 210 (1977) (stating that "there are obviously constitutional limits beyond which the States may not go" in redefining elements as affirmative defenses); *McMillan*, 477 U.S. at 85-88 (quoting this language from *Patterson* in the context of sentencing factors).

Section I.A discusses the Court's early cases on burdens of proof and elements in the context of affirmative defenses: *In re Winship*,¹⁸ *Mullaney v. Wilbur*,¹⁹ and *Patterson v. New York*.²⁰ Section I.B addresses *McMillan v. Pennsylvania*, the first case in which the Court considered sentence enhancements.²¹ Section I.C analyzes *Almendarez-Torres v. United States*,²² in which the Court rejected the elements rule three years ago. Section I.D notes *Monge v. California*, which again rejected the elements rule.²³ Section I.E describes the Court's sea change in *Jones v. United States*,²⁴ and Section I.F surveys *Apprendi*.

A. *Burdens of Proof, Elements, and Affirmative Defenses*

The Court's first cases on elements of crimes mentioned them in the context of burdens of proof. In *Winship*, the Court held that the Constitution requires proof beyond a reasonable doubt of "every fact necessary to constitute the crime charged."²⁵ Because convictions bring such enormous losses of liberty and stigma, the Court reasoned, the government should bear the risk of error.²⁶ Moreover, the criminal law's moral force depends on the community's respect and confidence that it is not condemning the innocent. To prevent doubts about convicting the innocent, the law insists upon proof beyond a reasonable doubt.²⁷ The Court said nothing, however, about how to determine what is a "fact necessary to constitute the crime charged."²⁸

The Court first tried to define elements in *Mullaney v. Wilbur*.²⁹ Maine law defined unlawful killings with malice aforethought as murder and those without malice aforethought as manslaughter.³⁰ But it also presumed malice aforethought.³¹ Defendants bore the burden of disproving malice aforethought by showing, by a preponderance of the evidence, that they had killed in the heat of passion.³²

18. 397 U.S. 358 (1970).

19. 421 U.S. 684 (1975).

20. 432 U.S. 197.

21. 477 U.S. 79.

22. 523 U.S. 224 (1998).

23. 524 U.S. 721 (1998).

24. 526 U.S. 227 (1999).

25. *In re Winship*, 397 U.S. 358, 363 (1970) (quoting *Davis v. United States*, 160 U.S. 469, 493 (1895)); see also *id.* at 362-64 (noting that many earlier cases had assumed that the Constitution requires proof beyond a reasonable doubt). Thus, the Court held, states may not convict juveniles of delinquency by only a preponderance of the evidence. *Id.* at 365-68.

26. *Id.* at 363-64.

27. *Id.* at 364.

28. *Id.* at 363 (quoting *Davis*, 160 U.S. at 493).

29. 421 U.S. 684 (1975).

30. *Id.* at 686 n.3 (quoting ME. REV. STAT. ANN. tit. 17, § 2651 (West 1964)).

31. *Id.* at 686.

32. *Id.* at 686-87.

The Supreme Court unanimously held that the absence of heat of passion was an element of the offense. Thus, the state bore the burden of disproving it beyond a reasonable doubt. First, the Court noted that at common law, heat of passion was the most important determinant of a killer's culpability.³³ It also stressed the clear trend toward putting that burden on the prosecution.³⁴

Second, it made no difference that heat of passion went only to the degree of culpability and not to guilt or innocence.³⁵ True, Wilbur already faced stigma and loss of liberty because he was guilty of manslaughter. Even so, due process protected his interest in avoiding the much greater consequences of a murder conviction.³⁶

Third, the Court stated that *Winship* cannot be limited to those facts defined as elements by state law. If it were, states could evade *Winship* by redefining elements as factors affecting only the amount of punishment.³⁷ "*Winship* is concerned with substance rather than this kind of formalism."³⁸

Read literally, this language could have made states bear the burden of proving all facts that affect culpability. Two years later, however, *Patterson v. New York*³⁹ retreated from *Mullaney*'s broad language. There, New York law required murder defendants to prove the affirmative defense of extreme emotional disturbance by a preponderance of the evidence.⁴⁰

The Supreme Court held, five to three, that the prosecution need not disprove extreme emotional disturbance beyond a reasonable doubt. It noted that New York law made death, intent to kill, and causation elements of the offense.⁴¹ The state had proved each of these facts beyond a reasonable doubt, and state law imposed no further burdens.⁴² Nor did state law presume or infer any facts against Patterson, unlike the Maine law in *Mullaney*.⁴³ Rather, it simply created an affirmative defense, a separate issue distinct from the statutory elements.⁴⁴ The Constitution allows states to punish all intentional killings. If a state chooses to allow affirmative defenses, the Court held, it may insist that the defendant prove them.⁴⁵

33. *Id.* at 696.

34. *Id.*

35. Maine courts construed homicide as a single crime, with gradations of murder and manslaughter. *Id.* at 691-92. The Supreme Court held that it was bound to accept Maine's characterization of its homicide statute. *Id.* at 690-91. But this fact made no difference. *Id.* at 697-99.

36. *Id.* at 698.

37. *Id.* at 698-99.

38. *Id.* at 699.

39. 432 U.S. 197 (1977).

40. *Id.* at 200.

41. *Id.* at 205.

42. *Id.* at 205-06.

43. *Id.* at 205-07.

44. *Id.* at 206-07.

45. *Id.* at 207, 209.

Nevertheless, “there are obviously constitutional limits beyond which the States may not go in this regard.”⁴⁶ The Court did not specify these limits, except to say that legislatures may not presume “all the facts essential to guilt.”⁴⁷

Patterson read *Mullaney* as narrowly as possible. The *Mullaney* law, *Patterson* stressed, had treated malice aforethought as an element and then presumed that same element.⁴⁸ The Supreme Court had previously held that defendants may bear the burden of proving the insanity defense, and *Mullaney* had not undercut that holding.⁴⁹ *Patterson* refused to read *Mullaney* as requiring the state to prove all facts that affect the degree of culpability or punishment.⁵⁰

While *Patterson* stressed deference to legislative definitions of elements, it and *Mullaney* both dealt with affirmative defenses to guilt. The Supreme Court did not apply this doctrine to facts that affect sentences until *McMillan v. Pennsylvania*.⁵¹

B. McMillan

In *McMillan*, the Court held that a fact that triggered a mandatory minimum sentence was not an element of the offense. Writing for the Court, then-Justice Rehnquist reiterated several statements from *Patterson*. First, states need not prove beyond a reasonable doubt every fact that

46. *Id.* at 210.

47. *Id.* (quoting *Tot v. United States*, 319 U.S. 463, 469 (1943)).

48. *Id.* at 215-16.

49. See *Leland v. Oregon*, 343 U.S. 790, 798 (1952) (holding that a state may place the burden of proving insanity beyond a reasonable doubt on the defendant); *Patterson*, 432 U.S. at 205 (noting that Chief Justice Burger and Justice Rehnquist, concurring in *Mullaney*, had said that *Mullaney* did not conflict with *Leland*; noting also that the Supreme Court had dismissed a later appeal, which had claimed that *Mullaney* overruled *Leland*, as not presenting a substantial federal question (citing *Rivera v. Delaware*, 429 U.S. 877 (1976))).

The Court later applied this doctrine to other affirmative defenses. For example, Ohio required a defendant charged with murder to prove self-defense by a preponderance of the evidence. The Court held that this rule did not violate due process because self-defense did not negate any of the elements of murder. *Martin v. Ohio*, 480 U.S. 228, 233 (1987). This rule, the Court noted, also accorded with the common-law practice of making defendants prove affirmative defenses. Even though most states had since changed this common-law requirement, the Court declined to force all states to follow the modern trend. *Id.* at 235-36.

50. *Patterson*, 432 U.S. at 214-15 & n.15. The Court noted that such a holding might discourage progressive reforms in the criminal law by, for example, forbidding Congress to create an affirmative defense to the felony-murder rule. *Id.* at 215 n.15. If states could not put the burden of proving extreme emotional disturbance on defendants, they might not enact the defense at all. *Id.* at 207. Forbidding affirmative defenses would encourage legislators to enact simple crimes in “unqualifiedly general terms, and leave only to sentence the adjustment between offenses of lesser and greater degree.” *Id.* at 211 n.13 (quoting *Patterson v. State*, 347 N.E.2d 898, 909-10 (N.Y. 1976) (Breitel, C.J., concurring)). This statement is a prescient forecast of what is likely to happen under the elements rule. Legislatures will draft broader criminal statutes and create more discretion at sentencing, undercutting fair warning to defendants and equal treatment.

51. 477 U.S. 79 (1986).

affects punishment.⁵² Also, states have broad latitude to define crimes and procedures, and their definitions are usually dispositive.⁵³ Here, the legislature had expressly made firearm possession a sentencing fact, not an element, and courts should hesitate to thwart the legislative will.⁵⁴ There are, however, unspecified constitutional limits on the legislature's power to redefine elements as sentencing factors. States may not presume defendants guilty, nor may they relieve the prosecution of its burden of proving guilt.⁵⁵

Next, the Court upheld the sentence enhancement based on five factors. First, the enhancement did not result in a vast sentencing differential, such as that between a nominal fine and a mandatory life sentence.⁵⁶ Second, the sentencing provision did not change the maximum sentence.⁵⁷ Third, it did not create a separate offense with a separate penalty.⁵⁸ Fourth, "[t]he 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."⁵⁹ Petitioners' claim "would have at least more superficial appeal if a finding of visible possession exposed them to greater or additional punishment."⁶⁰

52. *Id.* at 84. The three dissenters disagreed, arguing that every fact that "give[s] rise both to a special stigma and to a special punishment" is an element requiring proof beyond a reasonable doubt. *Id.* at 103 (Stevens, J., dissenting); *id.* at 94 (Marshall, J., dissenting) (quoting this language from Justice Stevens's dissent). Justice Stevens tried to limit this rule to aggravating, as opposed to mitigating, facts. He claimed that the democratic process provided a better check on broad oppressive laws coupled with mitigating facts than it did on aggravating facts. *Id.* at 100 (Stevens, J., dissenting). He admitted, however, that others had criticized this distinction as formalistic. *Id.* Justice Marshall criticized and rejected Justice Stevens's "formalistic distinction between aggravating and mitigating facts," arguing that legislatures might well use mitigating facts to achieve the same results as aggravating ones. *Id.* at 94 (Marshall, J., dissenting).

53. *Id.* at 85.

54. *Id.* at 85-86.

55. *Id.* at 86-87.

56. *Id.* at 87 (quoting *Mullaney v. Wilbur*, 421 U.S. 684, 700 (1975)). The Court noted that the five-year minimum applied only to those already facing maximum sentences of between ten and twenty years. *Id.* at 87-88.

57. *Id.* at 87-88.

58. *Id.*

59. *Id.* at 88.

60. *Id.* In addition, petitioners relied on *Specht v. Patterson*, 386 U.S. 605 (1967). In that case, *Specht* was convicted of the crime of indecent liberties, carrying a ten-year maximum sentence. *Id.* at 607. The court then sentenced him under a distinct statute, the Sex Offenders Act, to an indeterminate sentence of up to life imprisonment. *Id.* at 607-08. The court imposed this sentence without any notice or hearing, based solely on a presentence psychiatric report. *Id.* at 608. *Specht* did not claim a right to notice by indictment or proof to a jury beyond a reasonable doubt. Rather, he claimed this procedure violated his due process rights to a hearing, confrontation, cross-examination, and to put on evidence. *Id.* The Supreme Court agreed, holding that the new finding of fact was not an ingredient of the original offense but rather a separate criminal proceeding. *Id.* at 608.

The *McMillan* Court found *Specht* inapposite, noting that *Specht* itself had described its psychiatric context as "radically different" from ordinary sentencing. 477 U.S. at 89 (quoting *Specht*, 386 U.S. at 608). *Specht* is also distinguishable on another ground. The procedural rights conferred in *Specht* were quite different from those required by the elements rule: indictment, jury trial, and proof beyond a reasonable doubt. Indeed, the *Specht* rights were much closer to the procedural safeguards proposed in Subsection III.B.2 of this Article.

Fifth, the enhancement was not a mere subterfuge to circumvent *Winship*. The state had not amended any existing offense nor removed any elements to create the enhancement.⁶¹ Rather, it had taken one factor traditionally used at sentencing and told courts exactly how much weight to give it.⁶² If a legislature can allow or suggest that courts consider a fact at sentencing, the Court reasoned, it can guide courts in doing so.⁶³

Thereafter, the issue of sentence enhancements lay dormant for more than a decade. It arose again only a few years ago, in what looked like a routine case of statutory construction: *Almendarez-Torres v. United States*.⁶⁴

C. Almendarez-Torres

1. *Facts and Procedural Background*

Illegally reentering the United States after deportation is a federal felony under 8 U.S.C. § 1326(a). The penalties for illegal reentry vary depending on the alien's criminal record: Under subsection (a), aliens with no criminal records face up to two years in prison.⁶⁵ Under subsection (b)(1), aliens who have been convicted of certain misdemeanors or lesser felonies face up to ten years in prison.⁶⁶ And under subsection (b)(2), aliens who have been convicted of aggravated felonies face up to twenty years in prison.⁶⁷

A federal grand jury indicted Almendarez-Torres under 8 U.S.C. § 1326 without specifying a subsection. He pleaded guilty, admitting that he had illegally reentered the country after having been deported following three aggravated felony convictions.⁶⁸ At sentencing, he claimed that subsection (b)(2) created a separate offense and that a prior aggravated felony conviction was an element of that offense. Because his indictment did not mention the earlier convictions, he reasoned, his maximum sentence was two years. The Supreme Court rejected this claim five to four, holding that subsection (b)(2) created a sentencing factor rather than an element.

61. *McMillan*, 477 U.S. at 89.

62. *Id.* at 89-90.

63. *Id.* at 92.

64. 523 U.S. 224 (1998).

65. 8 U.S.C. § 1326(a) (1994).

66. *Id.* § 1326(b)(1).

67. *Id.* § 1326(b)(2).

68. *Almendarez-Torres*, 523 U.S. at 227.

2. *The Supreme Court Majority Opinion*

Writing for the Court, Justice Breyer treated the issue as one of statutory construction. Citing *McMillan*, the Court noted that legislatures normally determine which facts are elements and which are sentencing factors.⁶⁹ The Court divined the legislature's intent from many factors. In particular, the Court stressed that recidivism is the quintessential sentencing factor. Courts routinely interpret recidivism as a sentencing factor, not an element, and no statute clearly makes recidivism an element in such cases.⁷⁰ Also, the Court noted, interpreting (b)(2) as a separate crime would harm defendants. If (b)(2) were a separate crime, the government would have to prove the aggravated felony to the jury. This evidence of an aggravated felony might well prejudice a jury, even if a stipulation kept out the felony's name or details.⁷¹ And the Court refused to construe (b)(2) as a separate crime to avoid constitutional doubt. The Court did so both because the statute was clear and because the majority had no grave constitutional doubts.⁷²

Only after construing the statute did the Court turn to the Constitution. Almendarez-Torres claimed that the Constitution requires the government to charge recidivism in an indictment and prove it to a jury beyond a reasonable doubt. The Court rejected that claim, holding that recidivism need not be an element of the offense.⁷³ In other words, the Court held that not every fact that increases a statutory maximum need be an element.⁷⁴

Applying the *McMillan* factors, the Court held that recidivism need not be an element of the offense. First, recidivism is perhaps the most traditional sentencing factor. The Court had repeatedly held that prior convictions need not be charged in an indictment.⁷⁵ Recidivism, the Court

69. *Id.* at 228 (stating that legislatures are entrusted with defining crimes) (citing *Staples v. United States*, 511 U.S. 600, 604 (1994); *McMillan v. Pennsylvania*, 477 U.S. 79, 84-91 (1986)).

70. *Id.* at 230. The Court noted that recidivism is an element where the conduct proscribed would not otherwise be unlawful, as with the crime of being a felon in possession of a firearm. *Id.* (citing 18 U.S.C. § 922(g)(1) (1994)). The Court treated the felon-in-possession statute as distinguishable, because illegally reentering is unlawful even absent a prior conviction. *Id.*

71. *Id.* at 234-35.

72. *Id.* at 238-39.

73. *Id.* at 239.

74. *Id.* at 247.

75. *Graham v. West Virginia*, 224 U.S. 616 (1912), held that after convicting a defendant, a state may file a new information seeking an enhanced penalty for recidivism. It rejected double jeopardy and due process challenges to this procedure. Recidivism, the Court reasoned, is not an element of a crime. Rather, "it is a distinct issue [going to the punishment only], and it may appropriately be the subject of separate determination." *Id.* at 625 (internal quotation marks omitted). No indictment is required, as recidivism is not part of an offense but a distinct factual issue. *Id.* at 627.

Likewise, *Oyler v. Boles*, 368 U.S. 448, 452 (1962), held that "an habitual criminal charge does not state a separate offense." *Oyler* cited *Graham* for the proposition that states can try this "distinct issue" separately. *Id.* (quoting *Graham*, 224 U.S. at 625). Due process does not require an indictment or pretrial notice that, after trial, the prosecution will accuse a defendant of being a

had reasoned, is a “distinct” issue.⁷⁶ Recidivism “does not relate to the commission of the offense, but goes to the punishment only, and therefore . . . may be subsequently decided.”⁷⁷

Second, while Almendarez-Torres faced an increased statutory maximum rather than a minimum, this difference favors defendants. Minimum sentences bind sentencing judges in a way that maximum sentences do not.⁷⁸ *McMillan*, the Court argued, had not turned on this distinction. *McMillan* had said only that an increase in the statutory maximum might add “more *superficial* appeal” to a defendant’s argument.⁷⁹ *Almendarez-Torres* refused to read more into this statement, relying instead on the tradition of recidivism.

Third, the broad range of penalties for illegal reentry was not out of the ordinary, and the Sentencing Guidelines helped to channel this discretion.⁸⁰ Fourth, the statute did not change a preexisting definition of a well-established crime, nor was there any reason to think Congress had tried to evade *Winship* by presuming guilt or restructuring elements.⁸¹

3. *The Dissent*

Justice Scalia dissented, joined by Justices Stevens, Souter, and Ginsburg. Justice Scalia read *McMillan* as hinging on whether or not a fact increases the statutory maximum sentence.⁸² *McMillan*, in his view, raised enough constitutional doubt that the Court should construe subsection (b)(2) as a separate crime.⁸³ He then construed the statutory language as open to two plausible readings.⁸⁴ He claimed that the common law and most states treat recidivism as elements to be charged in the indictment and proved to juries.⁸⁵ And he disputed the unfairness of presenting recidivism evidence to juries, claiming that limiting instructions, stipulations, and evidentiary rules can prevent jury prejudice.⁸⁶ In sum, he contended, courts

habitual offender. Due process requires only reasonable notice before sentencing and an opportunity to be heard at sentencing. *Id.*

76. *Almendarez-Torres*, 523 U.S. at 244 (quoting *Graham*, 224 U.S. at 629 (emphasis omitted)).

77. *Id.* (quoting *Graham*, 224 U.S. at 629 (emphasis omitted)). *Almendarez-Torres* further argued that courts traditionally treated recidivism as an element. The Court rejected this claim, finding that any such tradition was neither uniform nor rooted in the U.S. Constitution. *Id.* at 246-47.

78. *Id.* at 245.

79. *Id.* at 245 (quoting *McMillan v. Pennsylvania*, 477 U.S. 79, 88 (1986)) (emphasis supplied by the *Almendarez-Torres* Court).

80. *Id.* at 245-46.

81. *Id.* at 246.

82. *Id.* at 253-55 (Scalia, J., dissenting).

83. *Id.* at 260.

84. *Id.* at 265-70.

85. *Id.* at 261.

86. *Id.* at 269.

should avoid the serious constitutional question at issue. Courts should, he argued, construe facts that increase maximum sentences as elements unless Congress “unambiguously” says otherwise.⁸⁷

* * *

In short, the majority and dissent stood diametrically opposed. The majority stressed deference to legislatures, the traditional treatment of recidivism at sentencing, and the need to avoid jury prejudice. In contrast, the dissent distrusted legislatures and judges, exalted juries, read tradition differently, and construed the statute narrowly to avoid strong constitutional doubts. Before long, these brewing constitutional doubts erupted into a full-fledged battle.

D. Monge

The Court faced the same issue again just a few months later in *Monge v. California*.⁸⁸ After Monge was convicted, the judge found that a recidivism enhancement for assault applied and the judge sentenced him to eleven years.⁸⁹ The appellate courts held that the state had not proffered sufficient evidence to support the enhancement. The California Supreme Court remanded for another sentencing hearing, holding that double jeopardy did not bar another attempt to prove the enhancement.⁹⁰

The U.S. Supreme Court affirmed, five to four. While eight Justices agreed that double jeopardy does not apply to noncapital sentencing, three of the eight dissented on another ground.⁹¹ Writing for these dissenters, Justice Scalia *sua sponte* claimed that the prior assault had to be an element.⁹² Because the state court authoritatively construed its own statute, Justice Scalia could not reconstrue it to avoid constitutional doubt. Instead, he adopted the elements rule. Monge had been convicted of a crime with a seven-year maximum, but, in Justice Scalia’s view, he had been acquitted of the enhanced crime.⁹³ Thus, Justice Scalia argued, the Constitution capped his sentence at seven years, not eleven. Justice Scalia gave only one reason for the elements rule: Without such a rule, legislatures could evade

87. *Id.* at 271.

88. 524 U.S. 721 (1998).

89. *Id.* at 725.

90. *Id.* at 726-27; *People v. Monge*, 941 P.2d 1121, 1128-30 (Cal. 1997) (plurality opinion); *id.* at 1134-35 (Brown, J., concurring).

91. *Monge*, 524 U.S. at 737-41.

92. *Id.* at 728, 740-41 & n.2 (Scalia, J., dissenting). Justice Scalia’s dissent was joined by Justices Souter and Ginsburg. The fourth *Almendarez-Torres* dissenter, Justice Stevens, did not reach the issue because he thought that the Double Jeopardy Clause applied to sentencing regardless of whether the assault was an element. *Id.* at 734-37 (Stevens, J., dissenting).

93. *Id.* at 740-41 (Scalia, J., dissenting).

the rights to jury trials and proof beyond a reasonable doubt. For example, they could make “knowingly causing injury to another” a crime punishable by thirty days’ imprisonment, reserving all other gradations for sentencing.⁹⁴ To shore up the Sixth Amendment, Justice Scalia claimed that every fact increasing a maximum sentence must be an element.

The majority, through Justice O’Connor, devoted only one paragraph to rebutting this argument. It noted that the parties had not raised the argument, the lower courts had not considered it, and *Almendarez-Torres* had foreclosed it.⁹⁵ Moreover, making every fact an element at trial would be unfair to many defendants. Defendants might not want to make two contradictory arguments to a jury, denying guilt while simultaneously contesting drug quantity.⁹⁶

After *Almendarez-Torres* and *Monge*, the issue appeared to be settled. Five Justices—Chief Justice Rehnquist and Justices O’Connor, Kennedy, Thomas, and Breyer—had held that not every sentence enhancement need be an element. Four—Justices Stevens, Scalia, Souter, and Ginsburg—wanted to treat every fact that increased a statutory maximum as an element. But *Almendarez-Torres*’s reliance on recidivism left the Court room to adopt a different rule for other kinds of facts. And one year later, the Court did just that. Justice Thomas changed sides in *Jones v. United States*.⁹⁷

E. Jones

1. *Facts and Procedural Background*

Jones was a federal carjacking case. The carjacking statute, 18 U.S.C. § 2119, consisted of a main paragraph listing various elements followed by subsections setting three levels of penalties. Under subsection (1), ordinary carjackings were punishable by up to fifteen years in prison.⁹⁸ Under subsection (2), carjackings resulting in “serious bodily injury” were punishable by up to twenty-five years in prison.⁹⁹ And under subsection

94. *Id.* at 738.

95. *Id.* at 728 (majority opinion).

96. *Id.* at 729.

97. 526 U.S. 227 (1999).

98. 18 U.S.C. § 2119(1) (Supp. 1993). The version of § 2119 in effect at the time provided:

Whoever, possessing a firearm . . . takes a motor vehicle that has been transported, shipped, or received in interstate or foreign commerce from the person or presence of another by force . . . shall—

(1) be . . . imprisoned not more than 15 years . . . ,

(2) if serious bodily injury . . . results, be . . . imprisoned not more than 25 years and

(3) if death results, be . . . imprisoned for any number of years up to life.

Id.

99. *Id.* § 2119(2).

(3), carjackings resulting in death were punishable by up to life imprisonment.¹⁰⁰

A federal grand jury indicted Jones for violating § 2119 but did not specify a subsection or charge serious bodily injury.¹⁰¹ The arraiging magistrate told Jones that he faced a maximum sentence of fifteen years.¹⁰² And at trial, the jury instructions did not mention serious bodily injury.¹⁰³ The jury convicted Jones.¹⁰⁴

After trial, the presentence report noted that a victim had suffered a perforated eardrum and some permanent hearing loss. Thus, the probation officer recommended a twenty-five-year sentence.¹⁰⁵ Jones objected that serious bodily injury was an element of § 2119(2). So, Jones reasoned, it had to be charged in the indictment and proved to a jury. Because it had been neither charged nor proved at trial, he claimed, his maximum sentence was fifteen years.¹⁰⁶ The trial court disagreed, finding by a preponderance of the evidence that Jones had seriously injured the victim. Accordingly, it sentenced him to twenty-five years on the carjacking count.¹⁰⁷ The Ninth Circuit affirmed.¹⁰⁸

2. *The Supreme Court Majority Opinion*

The Supreme Court reversed, five to four. Justice Souter wrote for the majority, which consisted of the four *Almendarez-Torres* dissenters plus Justice Thomas. The Court conceded that, at first glance, the body of § 2119 appeared to contain all the elements and the subsections appeared to be mere sentencing factors.¹⁰⁹ The Court observed, however, that subsections (2) and (3) provided for much higher penalties and conditioned them on further important facts. The Court doubted that Congress would have stripped these facts of the procedural safeguards of indictment and jury trial.¹¹⁰

Second, the Court conceded that the subsections did not stand alone, but, the Court argued, the main paragraph was not self-contained either.¹¹¹

100. *Id.* § 2119(3).

101. *Jones*, 526 U.S. at 230.

102. *Id.* at 230-31.

103. *Id.* at 231.

104. *Id.*

105. *Id.*

106. *Id.*

107. *Id.*

108. *United States v. Oliver*, 60 F.3d 547 (9th Cir. 1995). The Ninth Circuit again affirmed the sentence on appeal after remand. 116 F.3d 1487 (9th Cir. 1998).

109. *Jones*, 526 U.S. at 232-33.

110. *Id.* at 233.

111. The Court noted that the first paragraph ended with “shall,” not “shall be unlawful,” “shall be punished,” “shall be guilty of,” or the like. *Id.* at 234. And while the word “shall” often separates elements from penalties, it does not always do so. *Id.*

Third, the Court focused on how the law traditionally treated harm as an element. It noted many federal and state statutes that treated harm or injury as an element, suggesting that this backdrop cast light on Congress's intent.¹¹² Fourth, the Court distinguished *Almendarez-Torres* as turning on the distinctive history of recidivism as a sentencing factor.¹¹³

After reciting these arguments, the Court conceded that the statute was open to two interpretations. It then construed serious bodily injury as an element to avoid constitutional doubts. In describing the constitutional doubts it harbored, the Court first surveyed *Mullaney* and *Patterson*. One could, the Court admitted, read *Patterson* narrowly as a ban merely on using presumptions to make elements nominal. One could also, however, read *Patterson* broadly as forbidding states to shift the burden of proving "traditional" elements of crimes.¹¹⁴ And, the Court said, *McMillan* had left open whether facts that increase the statutory maximum must be proved to a jury beyond a reasonable doubt.¹¹⁵

The Court bolstered its constitutional concerns by turning to the history of the jury in England and colonial America. It conceded that no one in that era had discussed facts that increase statutory maxima.¹¹⁶ This silence, the Court suggested, was due to the broad judicial discretion over fines and corporal punishment for misdemeanors and the fixed penalties for felonies.¹¹⁷ The Court instead cited other forms of competition between judges and juries. Juries, the Court noted, had thwarted harsh sentences by acquitting defendants or convicting them of lesser-included offenses. In response, Parliament had created new, statutory offenses and denied the right to jury trial for those offenses. And judges had tried to limit juries to finding facts in libel cases, arrogating to themselves the power to apply the law to those facts.¹¹⁸ The American colonists had insisted on the Jury Clauses to guard against just such denials and erosions of the right to jury trial.¹¹⁹

112. *Id.* at 234-35. The Court also claimed that two of the three offenses on which § 2119 was modeled made harm an element and argued that § 2119 had to be the same. *Id.* at 235-36.

113. *Id.* at 235, 248-49 & n.10. The Court also noted that prior convictions are unlike almost any other fact. A person can be convicted only after being afforded rights to indictment and proof to a jury beyond a reasonable doubt. So, there may be no need to insist on such procedures a second time when recidivism is used at sentencing. *Id.* at 249.

114. *Id.* at 241-42.

115. *Id.* at 242.

116. *Id.* at 244-45.

117. *Id.*

118. *Id.* at 245-47. The Court's opinion was unclear as to whether it was courts or Parliament that had limited juries' ability to apply libel law to the facts. A glance at the sources cited by the Court confirms that judges were arrogating this power to themselves, and that Parliament fixed this problem by enacting Fox's Libel Act. *Infra* note 205.

119. *Jones*, 526 U.S. at 247-48.

The Court conceded that today, not every fact that affects sentencing must be found by a jury.¹²⁰ In footnote six, however, the Court suggested the elements rule:

[A]ny fact (other than a prior conviction) that increases the maximum penalty for a crime must be charged in an indictment, submitted to a jury, and proved beyond a reasonable doubt. Because our prior cases suggest rather than establish this principle, our concern about the Government's reading of the statute rises only to the level of doubt, not certainty.¹²¹

The government disputed the suggested elements rule. It claimed that this rule was inconsistent with the Court's cases allowing judges to find facts at capital sentencing.¹²² In *Hildwin v. Florida*¹²³ and *Walton v. Arizona*,¹²⁴ states had required judges to find certain facts as prerequisites to imposing the death penalty. The *Jones* Court did not read these cases as permitting judges to raise maximum sentences, however. *Walton*, according to *Jones*, had characterized the judge as selecting between two alternative penalties, not as raising the sentencing ceiling.¹²⁵ Nonetheless, the *Jones* Court recognized tension between these cases and *Jones*'s suggestion that juries must find all facts that increase maximum penalties.¹²⁶

3. Concurrences

Justice Stevens joined the majority opinion but also wrote separately to embrace an even broader rule. In his view, juries must find beyond a reasonable doubt all facts that increase statutory maxima or minima or make defendants eligible for the death penalty.¹²⁷ Also concurring, Justice Scalia suggested that he would apply the elements rule to facts that alter either minimum or maximum penalties.¹²⁸

4. The Dissent

Justice Kennedy dissented, joined by Chief Justice Rehnquist and Justices O'Connor and Breyer. The dissent first objected to the majority's

120. *Id.* at 248.

121. *Id.* at 243 n.6.

122. *See id.* at 249-51.

123. 490 U.S. 638 (1989) (per curiam).

124. 497 U.S. 639 (1990).

125. *Jones*, 526 U.S. at 250-51 (distinguishing *Walton*, 497 U.S. at 648; *Hildwin*, 490 U.S. at 639-40; *Spaziano v. Florida*, 468 U.S. 447, 459 (1984)).

126. *Id.*

127. *Id.* at 252-53 (Stevens, J., concurring).

128. *Id.* at 253 (Scalia, J., concurring).

“strained” statutory interpretation.¹²⁹ Next, it disputed the majority’s readings of *McMillan* and *Almendarez-Torres*. *McMillan* had rejected the claim that any fact that affects punishment is an element.¹³⁰ And one cannot limit *Almendarez-Torres*’s logic to recidivism or to the right to indictment, the dissent added.¹³¹

The majority’s rule, the dissent predicted, would disrupt sentencing and cause much uncertainty.¹³² In particular, at capital sentencing, the majority’s rule conflicts with the settled judicial role in finding aggravating facts.¹³³ In sum, the dissent warned, the majority’s rule is both murky and ominous.

* * *

Jones was the mirror image of *Almendarez-Torres*. Four members of the *Almendarez-Torres* majority repeated their arguments in dissent in *Jones*. They wanted to defer to legislatures, stressed traditional leeway for judicial fact-finding at sentencing, and forecast that the elements rule would cause grave practical problems. Conversely, the *Jones* majority copied the *Almendarez-Torres* dissent. These Justices distrusted legislatures and judges, exalted juries, relied on traditions of jury fact-finding, and adopted a strong rule of construction to avoid constitutional doubts. Thus, the *Jones* majority took what seemed to be a clear statute, found it unclear, and used that unclarity to avoid the constitutional issue. But its reasoning, and in particular footnote six, suggested a broad constitutional rule. The Court adopted that rule one year later, in *Apprendi v. New Jersey*.¹³⁴

F. Apprendi

1. *Facts and Procedural Background*

Charles Apprendi lived in what had been an all-white neighborhood in New Jersey. In 1994, one black family moved into that neighborhood, and Apprendi fired several bullets into the family’s home. After being arrested, he admitted that he had fired the shots because the family was black and he did not want them in his neighborhood.¹³⁵ He later retracted his statement about why he had fired the shots.¹³⁶

129. *Id.* at 254-64 (Kennedy, J., dissenting).

130. *Id.* at 265 (citing *McMillan v. Pennsylvania*, 477 U.S. 79, 84 (1986)).

131. *Id.* at 270 (“[T]here is no rational basis for making recidivism an exception.” (quoting *Almendarez-Torres v. United States*, 523 U.S. 224, 258 (1998) (Scalia, J., dissenting))).

132. *Id.* at 271.

133. *Id.* at 271-72.

134. 120 S. Ct. 2348 (2000).

135. *Id.* at 2351.

136. *Id.*

A state grand jury indicted Apprendi on twenty-three counts arising out of this and other incidents. None of the counts cited the state hate-crime statute or alleged a racially biased purpose.¹³⁷ Under a plea agreement, the prosecution dismissed twenty counts and he pleaded guilty to three: Counts 3 and 18, which charged second-degree possession of a firearm with a purpose to use it unlawfully against another person or property,¹³⁸ and Count 22, which charged third-degree unlawful possession of a bomb.¹³⁹

Ordinarily, a second-degree offense is punishable by five to ten years' imprisonment.¹⁴⁰ A separate hate-crime law doubles that range for defendants who "acted with a purpose to intimidate . . . because of race."¹⁴¹ In the plea agreement, the prosecution reserved the right to seek a hate-crime enhancement on Count 18, which arose out of the shooting described above. Apprendi, in turn, reserved the right to contest this enhancement as unconstitutional.¹⁴²

The court thoroughly advised Apprendi of the statutory maxima with and without the enhancement before accepting his plea.¹⁴³ The prosecution then moved for the hate-crime enhancement, and the defense moved to invalidate the enhancement.¹⁴⁴ The judge held a series of evidentiary hearings, at which Apprendi testified, put on witnesses, and cross-examined adverse witnesses.¹⁴⁵

The court found by a preponderance of the evidence that Apprendi had acted with a racially biased purpose.¹⁴⁶ It rejected Apprendi's constitutional challenge and applied the enhancement. The court sentenced him to twelve years on Count 18 and to shorter concurrent sentences on the other two counts.¹⁴⁷ The higher state courts affirmed.¹⁴⁸

137. *Id.* at 2352; Joint Appendix at 2-12, *Apprendi*, 120 S. Ct. 2348 (No. 99-478) [hereinafter Joint Appendix].

138. Joint Appendix, *supra* note 137, at 3-4, 10; *see* N.J. STAT. ANN. § 2C:39-4(a) (West 1995); *Apprendi*, 120 S. Ct. at 2352.

139. Joint Appendix, *supra* note 137, at 12; *see* N.J. STAT. ANN. § 2C:39-3(a).

140. N.J. STAT. ANN. § 2C:43-6(a)(2).

141. N.J. STAT. ANN. §§ 2C:44-3(e), 2C:43-7(a)(3) (West Supp. 2000).

142. *Apprendi*, 120 S. Ct. at 2352.

143. Joint Appendix, *supra* note 137, at 17-24, 41-42.

144. *Id.* at 29-30, 41.

145. Petition for Writ of Certiorari at 98a-121a, 175a-299a, *Apprendi*, 120 S. Ct. 2348 (No. 99-478).

146. *Id.* at 138a-144a.

147. *Id.* at 161a.

148. *State v. Apprendi*, 731 A.2d 485, 497 (N.J. 1999); *State v. Apprendi*, 698 A.2d 1265, 1271 (N.J. Super. Ct. App. Div. 1997).

2. *The Supreme Court Majority Opinion*

The Supreme Court reversed, five to four. Justice Stevens wrote for the majority, joined by Justices Scalia, Souter, Thomas, and Ginsburg.¹⁴⁹ The Court, quoting footnote six from *Jones*, adopted the elements rule. It held that any fact (other than recidivism) that increases the statutory maximum must be proved to a jury beyond a reasonable doubt.¹⁵⁰

First, the Court considered the history of jury trials. According to the Court, at common law, the jury's verdict was tightly linked to the judgment and sentence. Once the jury returned a verdict, the judge performed a ministerial act in entering a fixed judgment. At common law, each felony carried a fixed penalty. The judge simply imposed the fixed penalty unless it was so inappropriate as to warrant pardon or commutation.¹⁵¹ There was a gap between verdict and judgment, the Court admitted, for misdemeanors. For misdemeanors, judges had broad discretion to fine, whip, or imprison defendants. Actual imprisonment, however, was rare until the late eighteenth century.¹⁵²

Even where verdict and judgment were separable, judges exercised discretion within the limits set by the law. The Court claimed that around the time of the Founding, there was no such thing as a sentencing factor that increased the maximum penalty. All such facts, it claimed, were elements to be charged in the indictment and proved to the jury. Its only sources for this assertion were quotations from late-nineteenth- and early-twentieth-century treatises.¹⁵³ The Court also noted that when statutes created aggravated

149. As an initial matter, Justice Stevens noted that the trial judge could have imposed a twelve-year sentence by making the sentences on Counts 3 and 18 consecutive. But the trial judge had not done so; instead, he had imposed a twelve-year sentence on Count 18, above the ordinary statutory maximum. The question presented was the legality of the sentence actually imposed on Count 18, not how the trial judge could have arrived at the same result by a different route. *Apprendi*, 120 S. Ct. at 2354.

150. *Id.* at 2355. The Court's quotation from *Jones* also encompassed notice by indictment, but the *Apprendi* holding did not go quite so far. States are not bound by the Grand Jury Clause of the Fifth Amendment. *See Hurtado v. California*, 110 U.S. 516, 538 (1884). Thus, *Apprendi* limited his argument to the rights to trial by jury and proof beyond a reasonable doubt. In the Supreme Court, he did not press any claim to notice by indictment. *Apprendi*, 120 S. Ct. at 2355 n.3. Nonetheless, the Court's reasoning and its approving quotation of *Jones* leave little doubt that it will require notice by indictment in federal cases.

151. *See* 120 S. Ct. at 2356-57 (quoting 4 WILLIAM BLACKSTONE, COMMENTARIES *369-70; John H. Langbein, *The English Criminal Trial Jury on the Eve of the French Revolution*, in *THE TRIAL JURY IN ENGLAND, FRANCE, GERMANY 1700-1900*, at 13, 36-37 (Antonio Padoa Schioppa ed., 1987)).

152. *Id.* at 2357 n.7 (citing J.H. BAKER, AN INTRODUCTION TO ENGLISH LEGAL HISTORY 584 (3d ed. 1990)).

153. *Id.* at 2357 (quoting W.N. WELSBY, ARCHBOLD'S PLEADING AND EVIDENCE IN CRIMINAL CASES 51 (15th ed. 1862) [hereinafter ARCHBOLD] (citing 2 MATTHEW HALE, PLEAS OF THE CROWN *170)); *id.* at 2357-58 (quoting ARCHBOLD, *supra*, at 188); *id.* at 2358 n.9 (quoting 2 JOEL PRENTISS BISHOP, BISHOP ON CRIMINAL LAW §§ 933-934(1), 948 (John M. Zane & Carl Zollmann eds., 9th ed. 1923) [hereinafter BISHOP ON CRIMINAL LAW]); *see also id.* at 2362 n.15 (quoting *United States v. Reese*, 92 U.S. 214, 232-33 (1875) (Clifford, J., concurring in

versions of common-law felonies, prosecutors had to charge and prove the statutory elements to trigger the aggravated penalties.¹⁵⁴ The Court concluded that judges may exercise discretion at sentencing, so long as they sentence within the statutory range.

Buried in the Court's historical argument was a second argument about notice. As one treatise put it, indictments must be specific so that defendants have enough notice to prepare their defenses.¹⁵⁵ And historically, a defendant could discern the maximum punishment by looking at the indictment and the statute.¹⁵⁶ The Court's opinion spent only two or three sentences on the notice point.

Instead of dwelling on notice, the Court made a slippery-slope argument about the importance of protecting juries. Without an elements rule, entrusting facts to judges could erode the right to a jury trial.¹⁵⁷ In this case, the Court claimed, the biased purpose is part of the mens rea, a classic jury question.¹⁵⁸ Biased purpose is part of the very commission of the offense, yet the statute took this issue away from the jury.¹⁵⁹ This erosion matters because the rights to jury trial and proof beyond a reasonable doubt are so important. Juries and the reasonable-doubt standard protect defendants who may lose their liberty and suffer stigma upon conviction.¹⁶⁰ When a fact increases the statutory maximum, "it is obvious that both the loss of liberty and the stigma attaching to the offense are heightened," the Court claimed.¹⁶¹ Because enhancements have such consequences, the Court asserted, the procedures protecting the presumption of innocence should apply equally to such enhancements.¹⁶²

To stop this slippery slope, the Court adopted a bright-line rule: Juries must find beyond a reasonable doubt any "facts [besides recidivism] that increase the prescribed range of penalties" for a crime.¹⁶³ The Court

the judgment) (citing 1 JOEL PRENTISS BISHOP, COMMENTARIES ON THE LAW OF CRIMINAL PROCEDURE § 81 (Boston, Little, Brown & Co. 2d ed. 1872) [hereinafter BISHOP ON CRIMINAL PROCEDURE])).

154. *Id.* at 2357-58 (quoting ARCHBOLD, *supra* note 153, at 51, 188 (quoting 2 HALE, *supra* note 153, at *170)).

155. *Id.* at 2356 (quoting ARCHBOLD, *supra* note 153, at 44 (stating that charges in an indictment must be precise enough "that the defendant . . . may be enabled to determine the species of offence they constitute, in order that he may prepare his defence accordingly")).

156. *Id.* at 2359 n.10. Note that the majority opinion did not do much with this argument, burying it in a single sentence in a footnote.

157. *Id.* at 2359 (quoting *Jones v. United States*, 526 U.S. 227, 247-48 (1999) (stating that the Framers feared "that the jury right could be lost not only by gross denial, but by erosion")).

158. *Id.* at 2364.

159. *Id.* at 2366 (quoting *Almendarez-Torres v. United States*, 523 U.S. 224, 230, 244 (1998), for the proposition that recidivism is not an element because it "does not relate to the commission of the offense").

160. *Id.* at 2359.

161. *Id.* (discussing *In re Winship*, 397 U.S. 358, 363 (1970)).

162. *Id.* at 2360 (discussing *Mullaney v. Wilbur*, 421 U.S. 684, 697-98 (1975)).

163. *Id.* at 2363 (quoting *Jones v. United States*, 526 U.S. 227, 252-53 (1999) (Stevens, J., concurring)). Note that the Court's own formulations of its rule vary and are slightly inconsistent.

admitted that legislatures could evade this rule by raising statutory maxima and telling judges to lower the maxima on finding mitigating facts.¹⁶⁴ But the Court stressed that a mitigating fact, unlike an aggravating one, does not carry more punishment or stigma than authorized by the jury.¹⁶⁵ It thought that “structural democratic constraints” would prevent legislatures from raising all statutory maxima.¹⁶⁶ And it hinted that if a legislature did try to evade this rule, there might be other, unspecified constitutional limits.¹⁶⁷

The Court also distinguished *Walton v. Arizona*, which allowed judges to find facts that make defendants eligible for the death penalty.¹⁶⁸ The Court asserted that juries find all elements of capital crimes. Judges, the Court claimed, merely select between the maximum sentence of death and a lesser penalty.¹⁶⁹

Finally, the Court noted in a footnote the dissent’s fears that the elements rule would invalidate the Sentencing Guidelines. In response, the Court noted that the Guidelines were not before the Court and so expressed no view on their validity. The Court did, however, quote an earlier case’s statement that the Guidelines cannot raise sentences above statutory maxima.¹⁷⁰

3. Concurrences

Justice Scalia joined the Court’s opinion and wrote a brief concurrence. He responded to Justice Breyer’s arguments about the fairness and efficiency of judicial fact-finding under the Sentencing Guidelines. Justice Breyer’s arguments, he said, sketch out a “bureaucratic realm of perfect equity” but do not bear on the meaning of the Constitution.¹⁷¹ The Framers, Justice Scalia wrote, feared judges as bureaucratic arms of the state. They saw the inefficiency of jury trials as a price worth paying to safeguard

For example, the formulation quoted in the text above would appear to forbid judicial findings of fact that increase mandatory minima. But elsewhere in the opinion the Court stated that it was not overruling *McMillan*’s narrow holding on that point. *Id.* at 2361 n.13.

164. *Id.* at 2363 n.16.

165. *Id.*

166. *Id.* The Court did not elaborate on what these constraints were or how they would operate.

167. *Id.* (citing *Patterson v. New York*, 432 U.S. 197, 210 (1977); *Mullaney v. Wilbur*, 421 U.S. 684, 698-702 (1975)).

168. 497 U.S. 639, 647-49 (1990).

169. *Apprendi*, 120 S. Ct. at 2366 (quoting *Almendarez-Torres v. United States*, 523 U.S. 224, 257 n.2 (1998) (Scalia, J., dissenting)).

170. *Id.* at 2366 n.21 (“[O]f course, petitioners’ statutory and constitutional claims would make a difference if it were possible to argue, say, that the sentences imposed exceeded the maximum that the statutes permit for a cocaine-only conspiracy. That is because a maximum sentence set by statute trumps as higher sentence set forth in the Guidelines. [United States Sentencing Guidelines Manual] § 5G1.1.” (quoting *Edwards v. United States*, 523 U.S. 511, 515 (1998))).

171. *Id.* at 2367 (Scalia, J., concurring).

freedom.¹⁷² Justice Scalia admitted that the Court's rule allowing mitigating facts to reduce statutory maxima would produce disparities. But he insisted that the elements rule was essential to give content to the right to trial by jury.¹⁷³ And he expanded upon the need for notice hinted at by the Court. The elements rule "tell[s] a prospective felon" how much prison time his crime could bring.¹⁷⁴ Under the rule, mercy may lower his punishment, but "the criminal will never get *more* punishment than he bargained for when he did the crime."¹⁷⁵

Justice Thomas wrote a lengthy concurrence, joined in part by Justice Scalia. Both Justices endorsed a rule even broader than the Court's: Every fact that by law aggravates punishment is an element, though mitigating facts are not.¹⁷⁶ This rule, Justice Thomas said, would apply to both minimum and maximum sentences, overruling *McMillan* and, he implied, *Almendarez-Torres*.¹⁷⁷

Justice Thomas then surveyed nineteenth-century state cases and a few late-nineteenth-century treatises by Bishop. These authorities, he argued, had treated all facts that increased punishment as elements.¹⁷⁸ He also relied on two of the modern commentators who had supported the elements rule.¹⁷⁹ Apart from Bishop, however, none of the historical authorities had rested on the Federal Constitution. He concluded from this history that the elements rule was part of "the original meaning of the Fifth and Sixth Amendments."¹⁸⁰

Finally, Justice Thomas set aside two other areas for future consideration. First, he viewed the death-penalty cases that allowed judges to find aggravating circumstances as *sui generis*. The Court, he noted, had forbidden automatic death eligibility upon a jury's conviction of a capital crime. He set aside whether this modern restriction on jury verdicts justified an exception to his rule. He likewise set aside whether he would apply his rule to the Sentencing Guidelines, but hinted that he would.¹⁸¹

4. *The Principal Dissent*

Justice O'Connor dissented, joined by Chief Justice Rehnquist and Justices Kennedy and Breyer. She criticized the Court's and Justice

172. *Id.*

173. *Id.*

174. *Id.*

175. *Id.*

176. *Id.* at 2368 (Thomas, J., concurring).

177. *Id.* at 2378-80.

178. *Id.* at 2369-77.

179. *Id.* at 2377 (citing Knoll & Singer, *supra* note 3).

180. *Id.* at 2378.

181. *Id.* at 2380 & n.11.

Thomas's use of history. She also noted that not once in the last 200 years had the Court applied the novel constitutional elements rule.¹⁸²

Next, Justice O'Connor criticized the Court's treatment of its precedents. The Court, she noted, read *Mullaney* as requiring proof beyond a reasonable doubt of facts that increase punishment.¹⁸³ *Patterson*, however, had rejected this very reading, distinguishing *Mullaney* as a mere ban on using presumptions to undercut elements.¹⁸⁴ The Court also relied on *McMillan*. Yet the elements rule, she noted, appears to conflict with *McMillan*'s holding, which allowed sentencing factors to create mandatory minima.¹⁸⁵ Moreover, *Almendarez-Torres* and *Monge* had squarely rejected the elements rule.¹⁸⁶ And this rule is inconsistent with *Walton*, the capital punishment case, as well.¹⁸⁷

Justice O'Connor tried to figure out what exactly the elements rule required. On one reading, a legislature could raise an overall maximum to life and then have judges find facts that set a submaximum below life.¹⁸⁸ On a second reading, the elements rule would apply to aggravating but not mitigating facts.¹⁸⁹ Legislatures could evade this rule simply by jacking up maximum sentences and allowing judges to mitigate them.¹⁹⁰ As Justice Thomas's own cases showed, whether a fact is aggravating or mitigating can be in the eye of the beholder.¹⁹¹ And there is little reason to think that "structural democratic constraints" will constrain mitigating facts any more than they do aggravating ones today.¹⁹² On either reading, the elements rule is a mere formality.

Given how formalistic these alternatives are, Justice O'Connor suspected the Court's principle was even broader. It might, she feared, reach any fact that in real terms exposes a defendant to more punishment.¹⁹³ This rule would invalidate determinate-sentencing schemes

182. *Id.* at 2381 (O'Connor, J., dissenting).

183. *Id.* at 2383-84 (citing *id.* at 2360 (majority opinion); *Mullaney v. Wilbur*, 421 U.S. 684 (1975)).

184. *Id.* at 2384 (rejecting the argument that *Mullaney* required the state to bear the burden of proof on all facts that affect "the blameworthiness of an act or the severity of punishment authorized for its commission" (quoting *Patterson v. New York*, 432 U.S. 197, 214-15 (1977)) (emphasis added by the *Apprendi* dissent)).

185. *Id.* at 2385. The Court's various formulations of its own rule are inconsistent on this point. See *supra* note 163 and accompanying text (stating that in one place, the rule appears to conflict with *McMillan*; in another place, the majority purports to preserve *McMillan*'s holding for the time being).

186. *Id.* at 2386-87.

187. *Id.* at 2387-88.

188. *Id.* at 2389-90.

189. See *id.* at 2359, 2363 & n.16; *id.* at 2368 (Thomas, J., concurring).

190. *Id.* at 2390 (O'Connor, J., dissenting).

191. *Id.*

192. *Id.* at 2391.

193. See *id.* at 2391 (quoting *id.* at 2365 (majority opinion)).

(such as the Sentencing Guidelines) and capital case law.¹⁹⁴ And it would conflict with the long history of discretionary sentencing, which often turns on facts found by judges.¹⁹⁵ If judges can have broad discretion to find facts covertly, she argued, legislatures can guide judges by prescribing overt findings of these same facts.¹⁹⁶

Finally, Justice O'Connor argued, the elements rule would undercut the very values it sought to serve. One of the main purposes of the Sixth Amendment was to check arbitrary judges. By undermining determinate sentencing, she reasoned, the elements rule will force states to give more discretion to judges, leading to more arbitrariness.¹⁹⁷

5. *Justice Breyer's Dissent*

Justice Breyer also dissented, joined by Chief Justice Rehnquist. He argued that lumping all issues together at trial would be unworkable and unfair to defendants. Defendants cannot comfortably deny guilt while also contesting sentencing facts.¹⁹⁸

Justice Breyer then reiterated Justice O'Connor's point about how judges find facts in both discretionary and determinate sentencing. If the former is permissible, he asked, why not the latter? And if Congress can delegate power to the Sentencing Commission to guide judges, why can it not guide judges itself?¹⁹⁹

Justice Breyer predicted that the elements rule would greatly disrupt sentencing and legislative guidance. This is particularly true because legislatures have set up enhancements in reliance on the Court's previous cases.²⁰⁰ And the Court's principle of jury sentencing, taken to its logical conclusion, would displace judges and sentencing commissions and disrupt all sentencing, he feared.²⁰¹

* * *

Apprendi is *Jones* taken to its logical conclusion. The majority once again exalted jury fact-finding, relied heavily on historical arguments about juries' traditional role, and refused to trust judges or legislators. The Court feared the erosion of jury trials and also hinted at the need to give fair notice to defendants of enhancements. Because *Apprendi* involved a

194. *Id.* (citing *id.* at 2364 n.11 (Thomas, J., concurring) (hinting that his rule would govern the Sentencing Guidelines)).

195. *Id.* at 2391-92.

196. *Id.*

197. *Id.* at 2393-95.

198. *Id.* at 2398 (Breyer, J., dissenting).

199. *Id.* at 2397-2400.

200. *Id.* at 2401-02.

201. *Id.* at 2402.

state statute definitively construed by a state court, the Court could not reconstrue it to avoid constitutional doubts. Thus, the Court announced a broad constitutional rule. In contrast, the dissenters once again took issue with the majority's historical account and noted that legislators and judges historically had broad discretion. The dissenters once again noted with alarm the massive practical problems that the elements rule would cause at jury trials, at sentencing, and on habeas corpus.

The dissenters' tones of alarm do not exaggerate the stakes. The elements rule will have sweeping effects, and the suggestions in the concurrences would be more sweeping still. Such a sweeping rule should rest on solid foundations in history, principle, and policy. The foundations laid by the Court and commentators, however, are too shaky to support the rule. Moreover, the rule will cause many problems at trial and on habeas corpus. Part II looks at the rule from this perspective, critiquing the Court's and commentators' reasons on their own terms.

II. EXAMINING THE ELEMENTS RULE ON ITS OWN TERMS

Commentators and the Supreme Court have given three basic reasons for the elements rule. First, there are historical arguments about the limited role of judges at sentencing. Second, commentators and judges have feared sliding down a slippery slope, along which judicial arbitrariness would steadily displace jury trials. And third, they have hinted at the need to give fair notice to defendants of the penalties they face.

None of these arguments suffices to support the sweeping elements rule. Section II.A critiques the historical evidence. It notes that the early sources do not discuss sentence enhancements and the nineteenth-century sources do not bear on the meaning of the Fifth and Sixth Amendments. Section II.B challenges the slippery-slope argument. If a slippery slope is the problem, the easy-to-evade elements rule is not the solution. Section II.C acknowledges that the notice argument, the least developed point, is the most powerful. It shows, however, that the elements rule is not tailored to the problem of notice. And Section II.D notes the practical problems the elements rule will cause at trial and on habeas corpus. Thus, even on the Court's and commentators' own terms, the elements rule is flawed.

A. *Historical Arguments*

1. *The English and Colonial Evidence*

The Court's opinions in *Jones* and *Apprendi* contain several strands of historical argument. At the outset, however, the Court in *Jones* admitted that no history squarely supported the elements rule: "[T]he scholarship of

which we are aware does not show that a question exactly like this one was ever raised and resolved in the period before the Framing.”²⁰² In other words, there were no sentence enhancements at common law, so the eighteenth-century sources did not address them. This absence of direct support undercut the remainder of the Court’s historical argument.

Instead, *Apprendi* and *Jones* stressed how limited the role of judges was at sentencing. Juries, the Court stressed, returned verdicts that led directly to fixed felony sentences.²⁰³ Juries could mitigate sentences by refusing to convict on the greatest crime charged and convicting on lesser-included offenses.²⁰⁴ Parliament tried to circumvent juries by barring jury trials for

202. *Jones v. United States*, 526 U.S. 227, 244 (1999).

203. *Apprendi*, 120 S. Ct. at 2356-57 (collecting sources); *Jones*, 526 U.S. at 244-45 (same).

204. *Jones*, 526 U.S. at 245 & n.7 (collecting sources); 4 WILLIAM BLACKSTONE, COMMENTARIES *238-39.

At early common law, the penalty for felonies was death. But if the offender qualified for benefit of clergy, the judge would order branding or imprisonment for up to one year. 18 Eliz., ch. 7, §§ II-III (Eng. 1576). Later, by the seventeenth and eighteenth centuries, judges gained discretion to transport specified classes of felons (that is, banish them to a penal colony). See 39 Eliz., ch. 4, § IV (Eng. 1597); 14 Car. 2, ch. 12, § 23 (Eng. 1662); 18 & 19 Car. 2, ch. 3, § 2 (Eng. 1666); 22 & 23 Car. 2, ch. 7, § 3 (Eng. 1670); 31 Car. 2, ch. 2, § 13 (Eng. 1679); 4 Geo., ch. 11 (Eng. 1717); 8 Geo. 3, ch. 15 (Eng. 1767). See generally D.A. THOMAS, THE PENAL EQUATION: DERIVATIONS OF THE PENALTY STRUCTURE OF ENGLISH CRIMINAL LAW 1-9 (1978).

By the mid-eighteenth century, the standard punishments for felonies were death or transportation, or whipping for larceny of goods worth less than a shilling. It appears that jurors knew which penalties attached to which felonies and calibrated their verdicts to the seriousness of the crime. So, as one study showed, juries convicted defendants of capital offenses in all cases of highway robbery and livestock theft and most cases of burglary. But for most other felonies, they usually convicted defendants of lesser noncapital offenses unless the facts were egregious (such as cases with professional or gang criminals). John H. Langbein, *Albion’s Fatal Flaws*, 98 PAST & PRESENT 96, 106 (1983) [hereinafter Langbein, *Albion’s Fatal Flaws*] (documenting criminal jury trials in the Old Bailey in London from 1754 to 1756). The pattern of fictitious verdicts correlated with culpability and the seriousness of the facts, reinforcing this conclusion. See John H. Langbein, *Shaping the Eighteenth-Century Criminal Trial: A View from the Ryder Sources*, 50 U. CHI. L. REV. 1, 40-41, 52-55 (1983) [hereinafter Langbein, *Shaping the Eighteenth-Century Criminal Trial*] (noting that recurrent verdicts of thirty-nine shillings indicate that juries were returning verdicts below the forty-shilling threshold to avoid the death penalty; guilt was rarely contested, trials were really about sentencing and mitigation, and verdicts depended on the severity of the crime); see also JOHN BEATTIE, CRIME AND THE COURTS IN ENGLAND 1660-1800, at 419-21, 424-29 (1986) (same). Moreover, the pool of jurors was small enough that most juries contained at least some jurors who had served as jurors before, and thus were more likely to have learned the penalties attaching to particular crimes. See BEATTIE, *supra*, at 385-86, 398 (observing that in the eighteenth century, on average, more than a third of all jurors had served before); John H. Langbein, *The Criminal Trial Before the Lawyers*, 45 U. CHI. L. REV. 263, 276-77 (1978).

Indeed, in some cases judges warned juries that convicting a defendant of one offense would make the death penalty mandatory, while convicting of another would allow for the lesser sanction of transportation. Juries then committed “pious perjury” by convicting of lesser offenses to spare criminals’ lives. See Langbein, *Shaping the Eighteenth-Century Criminal Trial*, *supra*, at 22, 54 (recounting one case of stealing a forty-shilling watch and more than twenty guineas in cash, in which the judge warned the jury that theft of forty shillings or more was a felony without benefit of clergy; the jury then convicted the defendant of stealing thirty-nine shillings); see also THOMAS ANDREW GREEN, VERDICT ACCORDING TO CONSCIENCE: PERSPECTIVES ON THE ENGLISH CRIMINAL TRIAL JURY, 1200-1800, at 276-78, 281-82, 294-95 (1985) (describing sentence mitigation by juries and noting that “[t]he jury was only one of several sources of

new statutory offenses, and judges arrogated to themselves the power to apply libel law to the facts. The Constitution, the Court reasoned, was designed to check these abuses.²⁰⁵

The last of these points is inapposite to the elements rule. No legislature today is trying to create crimes without jury trials, and no judge is trying to ban juries from applying law to facts. The question is the line between guilt and sentencing, an issue never addressed at the time of the Founding. Nor is judicial sentencing discretion new. At common law, judges enjoyed broad discretion to impose fines, whipping, imprisonment, and other sentences in misdemeanor cases.²⁰⁶ They also had discretion to downgrade felony sentences of transportation to branding and to trigger the pardon and commutation processes.²⁰⁷ These forms of discretion coexisted with the Jury Clauses, which guarantee jury trials for all nonpetty offenses, misdemeanors and felonies alike.²⁰⁸ In other words, colonials treated

mitigation, and even then many of its merciful verdicts were encouraged—practically commanded—by the bench”).

In other words, at common law, juries had a de facto sentencing role. *See* BEATTIE, *supra*, at 429 (“[J]urors could anticipate precisely the sentence that would follow particular decisions, and their willingness to mitigate capital charges must have been influenced to some extent by the punishments Juries thus determined not only the general issue of the accused’s guilt or innocence but also, for many of those they did convict, the sentence that would follow.”); Langbein, *supra* note 151, at 36-37; Langbein, *Shaping the Eighteenth-Century Criminal Trial*, *supra*, at 53-55.

This is not to say that judges had no role at sentencing. As noted, they played quite a role in guiding and sometimes urging juries to return lesser verdicts. They had discretion to downgrade punishments of transportation to branding. Langbein, *Albion’s Fatal Flaws*, *supra*, at 109 n.43 (citing 4 Geo., ch. 11 (Eng. 1717)). And the monarch routinely followed judges’ recommendations as to whether to grant pardons or executive clemency. *Id.* at 109-12 & n.44; *see also* BEATTIE, *supra*, at 432 (stating that, by the early eighteenth century, a judicial “reprieve was almost certain to result in pardon”); V.A.C. GATRELL, *THE HANGING TREE: EXECUTION AND THE ENGLISH PEOPLE 1770-1868*, at 201 (1994); 1 LEON RADZINOWICZ, *A HISTORY OF ENGLISH CRIMINAL LAW AND ITS ADMINISTRATION FROM 1750*, at 120-21 (1948).

205. *Jones*, 519 U.S. at 245-47 (collecting sources). The Court’s argument is particularly puzzling because, upon first reading, the Court appears to blame Parliament for the judicial arrogation of the power to apply libel law to facts. But when one goes back and reads the Court’s sources, one finds the reverse is true. Judges had arrogated this power to themselves, and it was a legislative measure, Fox’s Libel Act, that restored the jury’s power. *See generally* *Rex v. Shipley* (The Dean of St. Asaph’s Case), 99 Eng. Rep. 774, 781 (K.B. 1784); GREEN, *supra* note 204, at 318-55. The moral of the story is that legislatures could be trusted to check judicial encroachment.

206. *Apprendi*, 120 S. Ct. at 2357 n.7; *Jones*, 526 U.S. at 244 (citing BAKER, *supra* note 152, at 584; 4 WILLIAM BLACKSTONE, *COMMENTARIES* *372). The sentences of imprisonment for misdemeanors were not capped at one year, as they are in modern America. In theory, judges could order unlimited imprisonment, though in practice they appear not to have imposed sentences of longer than five years. *See* 1 JAMES FITZJAMES STEPHEN, *A HISTORY OF THE CRIMINAL LAW OF ENGLAND* 490 (London, MacMillan 1883); THOMAS, *supra* note 204, at 8-9.

207. *See supra* note 204 (discussing the judicial power to recommend pardon or clemency and noting that the monarch routinely followed judges’ recommendations).

208. At common law, all nonpetty offenses, whether felonies or misdemeanors, enjoyed the protection of jury trials. There is some disagreement about how the common law defined the category of petty offenses. The Supreme Court has implied that at common law, *mala prohibita* could be petty offenses while *mala in se* could not. *See* *District of Columbia v. Colts*, 282 U.S. 63, 73 (1930). But the historical evidence suggests that the category of petty offenses was even narrower. The elder Justice Harlan argued that the petty-offense category encompassed only those

judicial discretion to enhance misdemeanor sentences within broad ranges, and to reduce felony sentences, as consistent with the jury-trial guarantees. This broad judicial discretion undercuts the suggestion that sentencing was the sacred province of juries alone.²⁰⁹

Not only did judges have some sentencing discretion at common law, but they gained even more in the nineteenth century with the rise of the indeterminate sentence.²¹⁰ The Court did not take issue with this settled practice. On the contrary, it endorsed broad judicial sentencing discretion within a range. Judges, it reaffirmed, may find facts and use discretion at sentencing.²¹¹ The corollary is that the Sixth Amendment does not require

offenses for which Parliament had expressly authorized bench trials. *Schick v. United States*, 195 U.S. 65, 80-81 (1904) (Harlan, J., dissenting); see also Timothy Lynch, *Rethinking the Petty Offense Doctrine*, KAN. J.L. & PUB. POL'Y, Fall 1994, at 7 (arguing that Justice Harlan was correct, both as a matter of common law and because the Constitution did not carry forward any petty-offense exception); Melissa Hartigan, Comment, *Creatures of the Common Law: The Petty Offense Doctrine and 18 U.S.C. § 19*, 59 MONT. L. REV. 343, 354-56 (1995) (arguing that research shows that Justice Harlan was correct as a matter of common law). Either way, the operative distinction was petty versus serious offenses, rather than misdemeanors versus felonies. In other words, nonpetty offenses included some if not many misdemeanors (at a minimum, common-law mala in se misdemeanors). Thus, these nonpetty misdemeanors were subject both to jury-trial guarantees and to the broad judicial sentencing discretion discussed in the text.

The Court has since abandoned any attempt to define petty offenses in terms of the common law, instead redefining the category to include only those crimes punishable by six months' imprisonment or less. See, e.g., *Lewis v. United States*, 518 U.S. 322, 325-28 (1996); *Blanton v. N. Las Vegas*, 489 U.S. 538, 543 (1989).

209. Nancy King and Susan Klein go further, asserting that by 1791, "most sentences [in America] were set by judges, at their discretion, within broad statutory ranges." King & Klein, *supra* note 3 (manuscript at 42). While the sources they cite show a nascent trend away from mandatory capital sentences, however, they do not show that judicial discretion in felony sentencing was well established by 1791. They show at most that sentencing was in flux in 1791, which would undercut the argument that the Fifth and Sixth Amendments enshrine a single approach to sentencing discretion. E.g., An Act for the Punishment of Certain Crimes Against the United States, ch. 9, 1 Stat. 112, 112-19 (1790) (creating thirteen crimes with sentencing ranges from up to one year to up to seven years' imprisonment, and one, bribery of a judge, punishable by unlimited imprisonment and fines at the discretion of the judge, but also creating six capital crimes); LAWRENCE M. FRIEDMAN, A HISTORY OF AMERICAN LAW 281-83 (2d ed. 1985) (noting a move away from the death penalty beginning in the late 1780s); ADAM JAY HIRSCH, THE RISE OF THE PENITENTIARY: PRISONS AND PUNISHMENT IN EARLY AMERICA 11-12 (1992) (noting that in 1785, Massachusetts began a trend away from capital punishment toward imprisonment); MORTON J. HORWITZ, THE TRANSFORMATION OF AMERICAN LAW, 1780-1860, at 12 (1977) (noting criticism of judicial discretion in 1786).

210. See *Apprendi*, 120 S. Ct. at 2358; *Williams v. New York*, 337 U.S. 241, 246-48 (1949). In some jurisdictions, such as the federal system, indeterminate sentencing was established very early in this nation's history. E.g., KATE STITH & JOSÉ A. CABRANES, FEAR OF JUDGING: SENTENCING GUIDELINES IN THE FEDERAL COURTS 9 (1998) ("From the beginning of the Republic, federal judges were entrusted with wide sentencing discretion . . . , permitting the sentencing judge to impose any term of imprisonment and any fine up to the statutory maximum.").

211. *Apprendi*, 120 S. Ct. at 2358 ("We should be clear that nothing in this history suggests that it is impermissible for judges to exercise discretion—taking into consideration various factors relating both to offense and offender—in imposing a judgment *within the range* prescribed by statute. We have often noted that judges in this country have long exercised discretion of this nature in imposing sentence *within statutory limits* in the individual case.").

juries to find all facts at sentencing.²¹² The history cited by the Court offers no basis for reserving certain sentencing facts for juries alone.

The Court tried to bridge this historical gap by quoting two twentieth-century Supreme Court cases. In *Williams v. New York*, the Court had endorsed the tradition of broad judicial discretion to find facts and sentence “within limits fixed by law.”²¹³ And in *United States v. Tucker*, the Court had agreed that, as a rule, sentences “within statutory limits” are not subject to collateral attack.²¹⁴ *Apprendi* tried to read these quotations as allowing discretion within statutory ranges but forbidding enhancements to those ranges. These quotations, however, are not on point. Neither case dealt with sentence enhancements, nor was either statement essential to the holding. More fundamentally, these cases contrast judges who follow the legislature’s will with renegade judges who ignore “limits fixed by law.” But renegade judges have nothing to do with *Apprendi*. Before *Apprendi*, trial judges were not rogues, jacking up sentences in violation of statutes. They were enhancing sentences within the “limits fixed by” enhancement statutes. Judges, in short, were not contravening the will of the legislature, but applying it. So *Williams* and *Tucker* are inapposite.

The majority’s final piece of historical evidence was Archbold’s 1862 treatise. Archbold wrote that when a statute creates an aggravated version of a common-law felony, the prosecution must charge and prove the statutory circumstances to trigger the statutory penalty.²¹⁵ The majority never explained why an 1862 English treatise illuminated the meaning of the Fifth and Sixth Amendments. Moreover, Archbold was discussing two separate crimes, one existing at common law and a second, aggravated version created by statute. To trigger the statutory penalties, one must of course prove the elements of the statutory offense.²¹⁶ Archbold did not

212. See *Jones v. United States*, 526 U.S. 227, 248 (1999) (“It is not, of course, that anyone today would claim that every fact with a bearing on sentencing must be found by a jury; we have resolved that general issue and have no intention of questioning its resolution.”); see also *McMillan v. Pennsylvania*, 477 U.S. 79, 93 (1986) (holding that there is no Sixth Amendment right to jury fact-finding at sentencing); *Spaziano v. Florida*, 468 U.S. 447, 459 (1984) (same).

213. 337 U.S. at 246, quoted in *Apprendi*, 120 S. Ct. at 2358; see also *id.* at 247.

214. 404 U.S. 443, 447 (1972), quoted in *Apprendi*, 120 S. Ct. at 2358; cf. *Apprendi*, 120 S. Ct. at 2358 n.9 (quoting Bishop’s 1923 treatise twice for the same proposition); *id.* at 2362 n.15 (quoting *United States v. Reese*, 92 U.S. 214, 232-33 (Clifford, J., concurring in the judgment) (paraphrasing Bishop on the same point)). For a discussion of the significance of Bishop’s writings, see *infra* notes 238-241 and accompanying text. *Reese*, one should note, had nothing to do with sentence enhancements. The issue in *Reese* was not whether to plead all the words of the statute, but whether additional details beyond the statutory words were also required. 92 U.S. at 232-33 (Clifford, J., concurring in the judgment).

215. ARCHBOLD, *supra* note 153, at 51, 188, quoted in *Apprendi*, 120 S. Ct. at 2357-58.

216. Justice O’Connor made the same point in dissent. She noted that Archbold had cited Hale’s *Pleas of the Crown*, and that the cited page said only that the prosecutor must charge and prove the elements of the statutory offense: “‘An indictment grounded upon an offense made by act of parliament must by express words bring the offense within the substantial description made in the act of parliament.’” *Apprendi*, 120 S. Ct. at 2382 (O’Connor, J., dissenting) (quoting 2 HALE, *supra* note 153, at *170).

discuss a statute creating a single crime and providing for penalty enhancements to the same crime. In sum, Archbold said only that prosecutors must charge and prove the elements; he did not define what those elements were.

The majority's account, in short, had little to do with the sentence enhancements governed by the elements rule. The Court did not reject judicial discretion and did not explain why discretion within a sentencing range is different from discretion to adjust that range. Its historical rhetoric about protecting jury trials merely set the stage for the slippery-slope argument that was to follow.

2. *Nineteenth-Century American Authorities*

Concurring, Justice Thomas looked at a more similar historical tradition. He reviewed many nineteenth-century state cases that had treated sentence enhancements as elements. These cases had required prosecutors to plead and prove to a jury all facts, including recidivism, that increased maximum sentences. Thus, Justice Thomas reasoned, the Fifth and Sixth Amendments carried forward this tradition.²¹⁷

Justice Thomas's account rested on the following argument: (1) At common law, any fact that by law aggravated punishment was an element. (2) The Fifth and Sixth Amendments enshrined this common-law tradition. Therefore, (3) the Constitution today requires that indictments charge and juries find aggravating facts beyond a reasonable doubt. Justice Thomas, however, proved neither his common-law nor his constitutional premise.

As for the common-law tradition, Justice Thomas cited cases dating mostly from the 1840s through the 1890s. Many of these cases did indeed treat aggravating facts as elements to be charged in indictments and proved to juries. All of his cases, however, were decided well after the Founding, most of them fifty to one hundred years later. To support his argument, Justice Thomas had to point to a common-law tradition *at the time of the Founding* that the Constitution enshrined.²¹⁸ He offered no evidence that the common law in the eighteenth century embodied the elements rule.

217. *Id.* at 2368-78 (Thomas, J., concurring).

218. One might reply that the nineteenth-century tradition is part of the backdrop against which the Fourteenth Amendment was adopted. On this reasoning, the Fourteenth Amendment might have incorporated and codified a practice that was well-settled in the states by that time. Neither the majority nor Justice Thomas made such an argument. Even if one were to make such an argument, however, it would not carry much weight. As the textual discussion goes on to show, nineteenth-century law was too inconsistent and in too much flux to amount to a solid, longstanding tradition of applying the elements rule. And even if there were such a tradition, there is no reason to view it as a fundamental right codified by the U.S. Constitution, as opposed to a technical common-law pleading requirement or a rule of statutory construction.

On the contrary, one of the earliest precedents adopted the opposite rule. In *State v. Smith*, an 1832 case from South Carolina, the court was unaware of any “precedent . . . of an indictment charging a former conviction for a similar offence.”²¹⁹ Recidivism was a simple matter for the court, not the jury, to decide before pronouncing sentence. “[I]t is certainly immaterial whether the first conviction is or is not recited in the record of the second, and this is in conformity with the practice here and every where else.”²²⁰

Smith shows that there was no uniform rule of charging and proving all sentence enhancements at common law. The court said that its rule prevailed “here and every where else” and that the court was unaware of any contrary practice. So *Smith*’s approach was not uncommon in the early nineteenth century.²²¹ The logical inference is that Justice Thomas’s rule did not become widespread until the late nineteenth century, and so was not firmly rooted at the time of the Founding.

Other cases likewise reflected a variety of practices. Cases from the 1830s and 1840s recognized that a prosecutor could file an information to trigger a recidivism enhancement long after trial.²²² In 1912, the Supreme Court agreed, upholding such a law in *Graham v. West Virginia*.²²³ The Court held that recidivism was not an element, but rather “a distinct issue” “go[ing] to the punishment only.”²²⁴ Thus, recidivism did not need to be charged in the indictment and could be determined separately after trial.²²⁵ Perhaps one could argue that recidivism enhancements were different from other enhancements, but Justice Thomas did not.

In short, the nineteenth-century tradition was not uniform, suggesting that the common law had no fixed rule on the subject. And Justice Thomas cited no eighteenth-century evidence linking this tradition back to the time of the Founding.

The second, constitutional premise of Justice Thomas’s syllogism was even shakier. He never explained why the Fifth and Sixth Amendments incorporated this alleged common-law tradition. None of his cases relied on

219. 42 S.C.L. (8 Rich.) 460, 460 (Ct. App. 1832).

220. *Id.*

221. The practice was not uniform. *See, e.g.*, *Commonwealth v. Smith*, 1 Mass. (1 Will.) 245 (1804) (reading a statute as requiring an allegation of the value of the goods stolen, because the punishment hinged on their value), *cited in Apprendi*, 120 S. Ct. at 2369 (Thomas, J., concurring). It is not enough for Justice Thomas to cite this lone, vague Massachusetts case in conjunction with other cases decided half a century later. To constitutionalize the elements rule, Justice Thomas had to show that it was solid, uniform, and longstanding at common law. The case discussed in the text strongly suggests that the elements rule was neither solid, nor uniform, nor longstanding.

222. *See Plumbly v. Commonwealth*, 43 Mass. (2 Met.) 413, 414 (1841) (Shaw, C.J.), described as one of two “leading antebellum cases” by Justice Thomas in *Apprendi*, 120 S. Ct. at 2372; and *Commonwealth v. Phillips*, 28 Mass. (11 Pick.) 28, 28 (1831).

223. 224 U.S. 616 (1912).

224. *Id.* at 624, 625.

225. *Id.* at 625-27.

or interpreted the U.S. Constitution. On the contrary, while a few interpreted state constitutions,²²⁶ most read as if they were about nonconstitutional rules of criminal pleading.²²⁷

Indeed, the historical authorities lumped the elements rule together with highly technical pleading rules that could not possibly be required by the Constitution. For example, Archbold listed the elements rule amidst a raft of trivial, indeed inane, pleading rules, such as: An indictment for stealing colts was not within a statute proscribing horse theft, as judges could not take notice that colts are horses.²²⁸ An indictment for stealing lambs could not be proved by showing that the defendant stole sheep.²²⁹ An indictment was defective because it charged that the defendant “did break to get out” and “did break and get out” instead of using the statutory phrase “break out.”²³⁰ And indictments could not use numerals, but rather had to write out numbers in words.²³¹ Some of Justice Thomas’s case law turned on similar technicalities.²³² These trivialities exemplify two more flaws in Justice Thomas’s argument: Just because the common law required rule X at

226. *E.g.*, *Goeller v. State*, 85 A. 954, 956 (Md. 1912); *Tuttle v. Commonwealth*, 68 Mass. (2 Gray) 505, 505 (1854); *see also Bandy v. Hehn*, 67 P. 979, 981 (Wyo. 1902) (implying but not stating that the state constitution required the recidivism allegation to be in the indictment); *cf. State v. Adams*, 13 A. 785, 786 (N.H. 1888) (rejecting the argument that the state constitution required the indictment to recite the facts of, or to refer to, a prior conviction); *State v. Freeman*, 27 Vt. 523, 526-27 (1855) (rejecting the state constitutional challenge because the offense was a minor one, without holding whether the Constitution requires an indictment for felonies).

One might argue that the inclusion of the elements rule in state constitutions shows its fundamental status, which might in turn suggest its inclusion in the Federal Constitution as well (perhaps via the Fourteenth Amendment). The number of cases resting on state constitutions was small, however, and only one (*Tuttle*) clearly did so before the adoption of the Fourteenth Amendment. One swallow does not a summer make.

227. *E.g.*, *State v. Smith*, 106 N.W. 187, 189 (Iowa 1906) (describing a statute as tracking the “general rule governing criminal procedure”); *Freeman*, 27 Vt. at 526 (holding that “according to the general rules of pleading, it is necessary to allege the former conviction, in the indictment,” but that the rule did not apply to Freeman because his offense was a minor one); *Shiflett v. Commonwealth*, 77 S.E. 606, 607 (Va. 1913) (“[B]y the rules of criminal pleading the indictment must always contain an averment of every fact essential to the punishment to be inflicted.”).

A rule of pleading could in theory have been grounded in a constitutional clause, such as a state petit- or grand-jury clause. None of the cases cited in this note, however, invoked or alluded to such federal or state constitutional provisions. Nor did any of these cases rely on the need to protect the sacred function of the grand or petit jury.

228. ARCHBOLD, *supra* note 153, at 53.

229. *Id.* at 52.

230. *Id.*

231. *Id.* at 54 (citing 2 HALE, *supra* note 153, at *170).

232. For example, one state statute enhanced the penalties for arson if the house was “lawfully occupied” at the time of the arson. *See Lacy v. State*, 15 Wis. 13, 15 (1862) (quoting WIS. REV. STAT. ch. 165, § 1(2) (1858)). An indictment charged Lacy under this subsection with “burn[ing] . . . the dwelling house of the said Manoah Griffin . . . one Manoah Griffin and the wife of the said Manoah Griffin being then in said dwelling house.” *Id.* at 14. The court refused to apply the enhancement because the indictment did not expressly allege that Manoah Griffin was lawfully in the house. While the indictment did allege that the house belonged to Manoah Griffin, the dweller was described as “one Manoah Griffin” rather than the “said Manoah Griffin” and so could not be inferred to be the same person as the owner, even though he was there at night with “the wife of the said Manoah Griffin.” *Id.* at 15. Thus, the pleading was defective.

one point does not mean that *X* was fundamental enough to be constitutionalized and frozen in place for all time. And the inclusion of the elements rule amidst this dross casts doubt on how fundamental the rule was.²³³

Some cases, in fact, expressly said that the elements rule was a mere canon of statutory construction. For instance, one of Justice Thomas's "leading antebellum cases," *Plumbly v. Commonwealth*, said nothing about the Constitution.²³⁴ *Plumbly* treated the elements rule as one of statutory construction and legislative intent.²³⁵ Federal case law agreed, treating it as "the established rule, under such statutes, *unless the statute designates a different mode of procedure*."²³⁶ And other state cases treated enhancements as elements based on statutory construction.²³⁷ This rule, in short, was one for divining the will of the legislature, not for shackling it.

In the end, only one source tied the elements rule to the Federal Constitution: Bishop's 1872 treatise. Bishop wrote: "[T]he indictment must allege whatever is in law essential to the punishment sought to be inflicted."²³⁸ He claimed that this doctrine flowed both from the common law and from federal and state constitutions.²³⁹ Bishop cited no authority

233. In construing a document or statute, one construes one provision in light of those surrounding it. This traditional canon of construction is known by its Latin name *noscitur a sociis* (a thing is known by its neighbors). BLACK'S LAW DICTIONARY 1084 (7th ed. 1999).

234. 43 Mass. (2 Met.) 413 (1841); see *Apprendi v. New Jersey*, 120 S. Ct. 2348, 2372 (2000) (Thomas, J., concurring) (describing *Plumbly* as one of two "leading antebellum cases").

235. 43 Mass. (2 Met.) at 417 (stating that the question is "what was intended by the legislature; and such intent is the governing rule in the construction of all statutes").

236. *Massey v. United States*, 281 F. 293, 297 (8th Cir. 1922) (emphasis added).

237. *E.g.*, *People v. Delany*, 49 Cal. 394, 395-96 (1874); *State v. Smith*, 106 N.W. 187, 188-89 (Iowa 1906).

238. 1 BISHOP ON CRIMINAL PROCEDURE, *supra* note 153, ch. 6, at 50; *accord id.* § 81, at 51 ("[T]he indictment must contain an allegation of every fact which is legally essential to the punishment to be inflicted . . ."); *id.* § 540, at 330 ("[T]he indictment must . . . contain an averment of every particular thing which enters into the punishment . . .").

239. *Id.* §§ 81-84, at 51-53 (common law); *id.* §§ 86-88, at 54-56 (constitutions). Bishop sharply distinguished his rule from judicial discretion to consider aggravating and mitigating circumstances:

Where the law permits the heaviest punishment, on a scale laid down, to be inflicted, and has merely committed to the judge the authority to interpose its mercy and inflict a punishment of a lighter grade, no rights of the accused are violated though in the indictment there is no mention of mitigating circumstances. The aggravating circumstances spoken of cannot swell the penalty above what the law has provided for the acts charged against the prisoner, and they are interposed merely to check the judicial discretion in the exercise of the permitted mercy. This is an entirely different thing from punishing one for what is not alleged against him.

Id. § 85, at 54. Justice Thomas quoted this passage and adopted its distinction between increasing the range and affecting the sentence within the range. *Apprendi*, 120 S. Ct. at 2378 (Thomas, J., concurring).

This argument is mistaken, at least insofar as it is a rule constraining legislatures. Whether a fact is within the range or outside the range may suggest whether the legislature meant for judges or juries to find the fact. But the distinction is useless as a rule to constrain legislatures. If a legislature wants to give this power to judges, all it need do is increase the maximum sentence and recharacterize the fact as one governing punishment within the range. See *infra* notes 259-270 and

for his constitutional musings, however. Instead, he relied on the speculative natural-law theorizing characteristic of the late nineteenth century:

Springing from natural reason, dwelling in abstract justice, [the rule] constitutes a part of our adjudged law, and is likewise confirmed by our written constitutions. Whatever amendments there may be of the laws, . . . this doctrine must abide, even in the midst of the wreck of all else pertaining to the law, until, if we could imagine such a thing, law itself is no more.²⁴⁰

This bare assertion about “our written constitutions” is no substitute for citations and proof.

Bishop, as noted, cited no earlier authorities for his constitutional claims. While several later cases cited Bishop with approval, none grounded the elements rule in the Federal Constitution.²⁴¹ Thus, Bishop’s authority extends only so far as his reasoning is persuasive. And nineteenth-century speculation about “natural reason” and “abstract justice” is no basis for discerning the original meaning of the Fifth and Sixth Amendments.

accompanying text. The rule might also have made a difference if juries were informed of what penalties their verdict would authorize. But judges do not tell juries about penalties, and indeed forbid them to consider them. *See infra* note 249. Thus, it makes no difference whether one labels the maximum twenty years plus a ten-year enhancement, or thirty years. It is simply a hoop through which legislatures must now jump.

240. 1 BISHOP ON CRIMINAL PROCEDURE, *supra* note 153, § 77, at 50. He distilled this principle from the brooding omnipresence of the common law, even though no one case stated his constitutional rule:

This principle, that the indictment must contain an allegation of every fact which is legally essential to the punishment to be inflicted, pervades the entire system of the adjudged law of criminal procedure. It is not made apparent to our understandings by a single case only, but by all the cases. Wherever we move in this department of our jurisprudence, we come in contact with it. We can no more escape from it than from the atmosphere which surrounds us.

Id. § 81, at 51. This statement, analogizing the rule to the omnipresent atmosphere, reads almost as a parody of the “brooding omnipresence” natural-law theories mocked by Justice Holmes. *S. Pac. Co. v. Jensen*, 244 U.S. 205, 222 (1917) (Holmes, J., dissenting) (“The common law is not a brooding omnipresence in the sky but the articulate voice of some sovereign or quasi-sovereign that can be identified . . .”).

241. *United States v. Reese*, 92 U.S. 214, 232-33 (1875) (Clifford, J., concurring in the judgment); *Riggs v. State*, 3 N.E. 886, 887 (Ind. 1885); *State v. Perley*, 30 A. 74, 75 (Me. 1894); *Goeller v. State*, 85 A. 954, 955-56 (Md. 1912); *Maguire v. State*, 47 Md. 485, 497 (1878); *State v. Hayward*, 83 Mo. 299, 307 (1884); *Larney v. Cleveland*, 34 Ohio St. 599, 600-01 (1878); *Hobbs v. State*, 44 Tex. 353, 354 (1875). Indeed, many of these cases cited section 81 of Bishop’s treatise, which is the introduction to the common-law discussion. *See Reese*, 92 U.S. at 232-33 (Clifford, J., concurring in the judgment); *Hayward*, 83 Mo. at 307; *Hobbs*, 44 Tex. at 354. *But cf. Goeller*, 85 A. at 955-56 (relying on various Bishop treatises in interpreting the state constitution). Note also that all of these cases came after the adoption of the Fourteenth Amendment, which precludes an originalist argument that these cases should influence its interpretation.

B. *Protecting Juries from the Slippery Slope*

The second argument for the elements rule is, as one commentator put it, the need to protect juries from “[t]he danger of the slippery slope.”²⁴² In other words, creeping judicial encroachment threatens the jury’s province as a safeguard of defendants’ rights.²⁴³

The Court agreed. Without such a rule, the Court asserted, “the jury right could be lost . . . by erosion.”²⁴⁴ Allowing judges to find these facts could lead to judicial arbitrariness and oppression.²⁴⁵ Jury trials and proof beyond a reasonable doubt protect not only the innocent from stigma and losing liberty. They also protect the guilty from added losses of liberty and stigma, the Court reasoned.²⁴⁶ When a fact enhances a maximum sentence, the Court claimed, “it is obvious that both the loss of liberty and the stigma attaching to the offense are heightened.”²⁴⁷ The Court saw the elements rule as a bulwark to prevent judicial fact-finding from eroding jury findings of these crucial facts.

Taking the last point first, it is not at all “obvious” that a sentence enhancement heightens stigma. The label attached to the enhanced crime is the same, and the label is what carries stigma. People identify a criminal as, say, a burglar, not a burglar who had a weapon; as a drug dealer, not a five-kilogram drug dealer. The verdict branded Jones a carjacker, not a carjacker-who-injured, and the injury enhancement would not have been named on Jones’s rap sheet.²⁴⁸ Even if one assumes that sentence lengths

242. Rosenberg, *supra* note 3, at 480.

243. *See id.* at 481-84.

244. *Apprendi*, 120 S. Ct. at 2359 (quoting *Jones v. United States*, 526 U.S. 227, 247-48 (1999)); *accord id.* at 2359 n.11 (praising the jury-trial guarantee of Article III, Clause 2 as needed “to guard with the most jealous circumspection against the introduction of new, and arbitrary methods of trial, which, under a variety of plausible pretenses, may in time, imperceptibly undermine this best preservative of LIBERTY” (quoting *Jones*, 526 U.S. at 248 (quoting *A [New Hampshire] Farmer*, No. 3, June 6, 1788, *reprinted in* THE COMPLETE BILL OF RIGHTS 476, 477 (N. Cogan ed. 1997)))).

245. *Id.* at 2367 (Scalia, J., concurring) (claiming that fact-finding by a judge, a lone employee of the state, is a threat to liberty).

246. *Id.* at 2360 (quoting *Mullaney v. Wilbur*, 421 U.S. 684, 697-98 (1975) (stating that criminal law “is concerned not only with guilt or innocence in the abstract, but also with the degree of criminal culpability”)). This broad reading of *Mullaney*, however, conflicts with *Patterson*. *Patterson* refused to read *Mullaney* as requiring proof beyond a reasonable doubt of all facts that increase culpability. *Patterson v. New York*, 432 U.S. 197, 214-15 & n.15 (1977); *see Jones*, 526 U.S. at 241 (likewise rejecting such a reading of *Mullaney*); *Almendarez-Torres v. United States*, 523 U.S. 224, 240-41 (1998) (same); *McMillan v. Pennsylvania*, 477 U.S. 79, 84 (1986) (same).

247. *Apprendi*, 120 S. Ct. at 2359.

248. *See* Julie R. O’Sullivan, *In Defense of the U.S. Sentencing Guidelines’ Modified Real-Offense System*, 91 NW. U. L. REV. 1342, 1376 (1997) (“The resolution of these sentencing issues does not impose upon the defendant the stigma of a criminal conviction. The defendant has no criminal judgment lodged against him, bears no notation on his rap sheet regarding such conduct, and suffers no continuing civil disabilities by virtue of it.”); *infra* text accompanying notes 465-467.

will be publicized and somehow affect stigma, there is no reason to single out sentence enhancements for special restrictions. Any stigma attached to sentence length will be the same whether caused by a sentence enhancement or an equivalent increase within the sentencing range. In other words, a five-year enhancement to a twenty-year maximum adds no more stigma than raising a twenty-year sentence to twenty-five years within a zero-to-life sentencing range. Likewise, while sentence enhancements do heighten loss of liberty, the loss is the same whether caused by an enhancement or an equivalent increase within the sentencing range. As juries need not find all facts that increase sentences, why should they have to find these particular facts?²⁴⁹

A bigger problem with this line of argument is that it rests on a speculative parade of horrors. Justice Scalia, for example, rested his entire dissent in *Monge* on just such a parade of horrors. Unless sentence enhancements are elements, he argued, a legislature could make “knowingly causing injury” punishable by up to thirty days in jail but use judicial enhancements to increase the maximum to life.²⁵⁰ Commentators have echoed this argument.²⁵¹ Realistic hypotheticals are useful for testing a rule against likely developments. Justice Scalia’s hypothetical is, however, implausible and hardly solid enough to justify a preemptive, sweeping bright-line rule. What the legislature might do, if it went insane, is not a good reason to pass on what it in fact has done.

The slippery slope is closely related to another claim of Justice Scalia’s. He insisted that the elements rule is essential to give content to the right to a jury trial. Otherwise, the right to a jury trial would be meaningless, as legislatures could redefine elements at will.²⁵² Commentators have made

249. See *Jones*, 526 U.S. at 248 (admitting that juries need not find all sentencing facts); *McMillan*, 477 U.S. at 93 (same); *Spaziano v. Florida*, 468 U.S. 447, 459 (1984) (same).

It is no answer to say that sentence enhancements are different because they increase the range authorized by the jury. See *Apprendi*, 120 S. Ct. at 2363 n.16 (suggesting that mitigating facts are different from aggravating ones because a jury verdict of guilt authorizes a certain quantum of punishment and stigma). Juries do not authorize a punishment range in any meaningful sense. Legislatures authorize punishment ranges; juries, blindfolded, write blank checks for legislatures by returning guilty verdicts. Juries are not told about penalties and indeed are forbidden to consider them. The jury is not being misled about the punishment it is authorizing. It would be different if a court told the jury its sentence could result in twenty years of imprisonment, but later imposed thirty years. But because the law keeps juries ignorant of penalties, it does not matter to them whether the nominal maximum sentence is twenty-five years minus a possible ten-year reduction, or is fifteen years plus a possible ten-year enhancement. See *infra* text accompanying notes 476-477.

250. *Monge v. California*, 524 U.S. 721, 738 (1998) (Scalia, J., dissenting).

251. See Herman, *supra* note 3, at 328-29; Priester, *supra* note 3, at 249-50, 292 n.210.

252. See *Apprendi*, 120 S. Ct. at 2367 (Scalia, J., concurring) (“What ultimately demolishes the case for the dissenters is that they are unable to say what the right to trial by jury *does* guarantee if, as they assert, it does not guarantee—what it has been assumed to guarantee throughout our history—the right to have a jury determine those facts that determine the maximum sentence the law allows. They provide no coherent alternative.”).

the same argument.²⁵³ Not every right need fetter the legislature in order to be meaningful, however. The legislature has latitude to define crimes, but procedural protections attach to whatever the legislature has chosen to define. For example, the Court has protected defendants against double jeopardy through the *Blockburger* test.²⁵⁴ *Blockburger* is very deferential to legislatures, allowing them to define elements and lesser-included offenses as they see fit.²⁵⁵ *Blockburger* simply attaches double jeopardy consequences to the elements specified by the legislature. Indeed, *Blockburger* is nothing more than a tool for divining legislative intent. If the legislature makes a contrary intent clear, courts defer to that as well.²⁵⁶ Yet the current Court continues to embrace the *Blockburger* test.²⁵⁷ The Court evidently agrees that wholesale deference to legislative definitions of crimes does not make the Double Jeopardy Clause a nullity; it still checks judges and prosecutors.²⁵⁸ The same logic applies to the Jury Clauses.

Assume for the moment, however, that the slippery slope and legislative manipulation are real problems. Even so, the elements rule is not the solution. The rule would require juries to find all facts aggravating the maximum sentence but allow judges to find mitigating facts.²⁵⁹ Commentators, surprisingly, endorse this flimsy distinction.²⁶⁰ As the

253. See Priester, *supra* note 3, at 291; Meadow, *supra* note 3, at 1604-05.

254. *Blockburger v. United States*, 284 U.S. 299, 304 (1932).

255. *Id.* (“[T]he test to be applied to determine whether there are two offenses or only one, is whether each provision requires proof of a[n additional] fact which the other does not.”).

256. See *Albernaz v. United States*, 450 U.S. 333, 340 (1981) (“The *Blockburger* test is a ‘rule of statutory construction,’ and because it serves as a means of discerning congressional purpose the rule should not be controlling where, for example, there is a clear indication of contrary legislative intent.”); *Whalen v. United States*, 445 U.S. 684, 691-92 (1980).

257. See, e.g., *Rutledge v. United States*, 517 U.S. 292, 297 (1996).

258. Many commentators agree as well. See, e.g., GEORGE C. THOMAS III, *DOUBLE JEOPARDY: THE HISTORY, THE LAW* (1998); Nancy J. King, *Portioning Punishment: Constitutional Limits on Successive and Excessive Penalties*, 144 U. PA. L. REV. 101, 114 (1995). Others prefer a more robust, legislature-constraining approach to the Double Jeopardy Clause. See, e.g., Susan R. Klein, *Double Jeopardy’s Demise*, 88 CAL. L. REV. 1001, 1002 (2000) (reviewing THOMAS, *supra*); Susan R. Klein & Katherine P. Chiarello, *Successive Prosecutions and Compound Criminal Statutes: A Functional Test*, 77 TEX. L. REV. 333, 375-82 (1998). I need not opine on which approach to the Double Jeopardy Clause is preferable. My point is more modest: that one can reasonably believe, and the *Apprendi* Court in fact believes, that *Blockburger*’s deference to legislatures leaves a meaningful role for the Clause.

259. The Court’s formulations of its own rule vary somewhat, but the various formulae seem to agree on this point. Compare *Apprendi v. New Jersey*, 120 S. Ct. 2348, 2355 (2000) (“[A]ny fact (other than prior conviction) that increases the maximum penalty for a crime must be charged in an indictment, submitted to a jury, and proven beyond a reasonable doubt.” (quoting *Jones v. United States*, 526 U.S. 227, 243 n.6 (1999))), and *id.* at 2362-63 (“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.”), with *id.* at 2363 n.16 (drawing a sharp line “between facts in aggravation of punishment and facts in mitigation”), *id.* at 2367 (Scalia, J., concurring) (suggesting that aggravating facts exceed the warning that defendants have of their maximum penalties, while mitigating ones do not), and *id.* at 2368 (Thomas, J., concurring) (defining an element as “every fact that is by law a basis for imposing or increasing punishment (in contrast with a fact that mitigates punishment)”).

260. See, e.g., Rosenberg, *supra* note 3, at 502; Note, *supra* note 3, at 1362.

Coase Theorem teaches, however, the baseline allocation of rights is immaterial to the ultimate distribution so long as the parties can contract around it without significant costs. It is child's play to redefine aggravating facts as mitigating facts. Legislatures need only raise statutory maxima and then provide for reductions if a judge finds that certain aggravating facts were absent. (Alternatively, legislatures could give judges even more power by returning to broad, indeterminate sentencing ranges.) As Justice Kennedy put it, "[n]o constitutional values are served by so formalistic an approach, while its constitutional costs in statutes struck down . . . are real."²⁶¹

Just as legislatures can turn aggravating facts into mitigating ones, so too aggravating facts are hard to distinguish from mitigating ones. For example, one can describe malice aforethought as aggravating manslaughter to murder. Or one can just as easily describe the absence of malice aforethought (heat of passion) as mitigating murder to manslaughter.²⁶² As Justice O'Connor put it, whether a fact is aggravating or mitigating "rests in the eye of the beholder."²⁶³

In response, the majority admitted that states could "hypothetically" evade its rule by raising statutory maxima.²⁶⁴ The Court dismissed this possibility in a footnote, however. It thought "structural democratic constraints" would limit maximum sentences to what is "generally proportional to the crime."²⁶⁵ Commentators likewise postulate that democratic outrage would prevent legislative evasion.²⁶⁶ They argue that the elements rule serves a truth-in-labeling function, making sentences

261. *Jones*, 526 U.S. at 267 (Kennedy, J., dissenting).

262. See King & Klein, *supra* note 3 (manuscript at 17).

263. *Apprendi*, 120 S. Ct. at 2390 (O'Connor, J., dissenting). Justice O'Connor gave another example of the "eye of the beholder" phenomenon, drawn from Justice Thomas's own case law: Wisconsin used to punish arsonists in which no person died by seven to fourteen years, or only three to ten years if "there was no person lawfully in the dwelling house." WIS. REV. STAT., ch. 165, § 1 (1858). The absence of a person appeared to be a mitigating fact, but the Wisconsin Supreme Court instead treated the presence of a person as aggravating. *Lacy v. State*, 15 Wis. 13, 15-16 (1862).

264. *Apprendi*, 120 S. Ct. at 2363 n.16.

265. *Id.*; accord *McMillan v. Pennsylvania*, 477 U.S. 79, 100 (1986) (Stevens, J., dissenting) (trusting that "the continued functioning of the democratic process" will prevent legislatures from raising all statutory maxima). The Court also stated that if a legislature went down this road, "we would be required to question whether the revision was constitutional under this Court's prior decisions." *Apprendi*, 120 S. Ct. at 2363 n.16. But if the Court was seeking to stop the slippery slope, why adopt a rule that does not do so and will require another rule to fill the void? And the Court did not suggest what other rule could possibly fill that void. *But cf.* King & Klein, *supra* note 3 (manuscript at 70-78) (proposing such a rule, consisting of a multifactor balancing test). Justice Stevens gave no explanation of how "structural democratic constraints" might operate.

266. See, e.g., Frank R. Herrmann, "Understanding" Maximum Sentence Enhancements, 46 BUFF. L. REV. 175, 213-14 (1998); Rosenberg, *supra* note 3, at 507 (dismissing the objection as "too speculative"); Note, *supra* note 3, at 1365. These articles do not explain how "structural democratic constraints" would prevent such evasion, apart from bald assertions about what voters would not tolerate. The articles listed in the following footnote expand on these arguments somewhat. See *infra* note 267.

more transparent to the public. Voters supposedly understand substantive law much better than procedures, such as burdens of proof. Likewise, they supposedly understand overall maximum sentences (substantive penalties) much better than sentence enhancements (which are obscured by the sentence-enhancement procedure). So, the argument goes, voters will put pressure on their elected representatives not to raise maximum sentences to get around the elements rule.²⁶⁷

This reliance on “structural democratic constraints” is unconvincing, for four reasons. First, it depends on a significant gulf between voters’ understandings of overall maxima and sentence enhancements. But many, if not most, voters are ignorant of both, so it is implausible to rest reform on the assumption of voter knowledge.²⁶⁸ Second, determinate sentence enhancements triggered by specific facts are clear enough to promote accountability. Accountability rests not on whether a law is substantive or procedural, but on how clear it is. If the elements rule pushes legislatures back toward broad, indeterminate sentencing ranges, voters will be even less well-informed and broad, opaque laws will produce even less accountability.

Third, the party seeking a deviation from the status quo or baseline legal rule normally bears the burden of proving the fact that supports the deviation. Under current law, the prosecution bears the burden of proving aggravators, albeit by a preponderance of the evidence. But if the elements rule forces legislatures to recharacterize aggravators as mitigators, they may put the burden of proving these mitigators on defendants. (So the elements rule would likely make burdens of proof worse for defendants.) By the commentators’ logic, voters would be oblivious to this change because they do not understand burdens of proof. So the legislature could simply point out that carjackers who do not cause injury would face only fifteen years. By this reasoning, the public would not comprehend the difference between

267. Writing in another context, Barbara Underwood has argued that “rules about proof at trial are even less accessible to popular understanding than rules about conduct in society.” Barbara D. Underwood, *The Thumb on the Scales of Justice: Burdens of Persuasion in Criminal Cases*, 86 YALE L.J. 1299, 1323 (1977). Thus, she reasoned, “it is somewhat more reasonable in the case of substantive law than in the case of rules of proof to rely on political processes to evaluate and revise the law.” *Id.* In her view, the government should bear the burden of persuasion on all facts relating to criminal culpability. *Id.* at 1340. This requirement serves a truth-in-labeling function, making legislation more transparent and more likely to provoke a democratic response. *Id.* at 1317, 1323-24. A student case note has extended this argument to sentence enhancements, claiming that a raised maximum sentence would be more transparent than a lower maximum plus aggravating facts. Recent Case, *Commonwealth v. Williams*, 733 A.2d 593 (Pa. 1999), 113 HARV. L. REV. 2140, 2145 & n.58 (2000).

268. Underwood herself notes that “popular understanding of the substantive law is notoriously deficient.” Underwood, *supra* note 267, at 1323. Her argument depends on the existence of a significant gulf between voters’ awareness of substance and procedure, but, as she admits, there is reason to question how aware voters are of substance. Empirical work is needed on this question.

aggravating a low sentence and mitigating a high one, as long as the substantive number at the end is the same.

Finally, the “structural democratic constraints” argument is in tension with the slippery-slope rationale for the elements rule: If we cannot trust legislatures with sentence enhancements, how can we trust them with maximum sentences? If “structural democratic constraints” today allow legislatures to aggravate sentences, they will equally allow legislatures to reach the same result via mitigation. Legislatures often override Supreme Court decisions with which they disagree.²⁶⁹ In a world where getting tough on crime appeals to voters, legislators have enough incentive to do so.²⁷⁰ Indeed, they can claim even more credit by raising nominal maximum sentences. At most, “structural democratic constraints” might limit sentences for those few crimes that average voters could imagine committing, such as simple assault or gun possession. As for violent, drug, serious property, and most other crimes, most voters identify with the victim, not the criminal. Raising maxima for most crimes, then, would likely be quite popular. So “structural democratic constraints” will not provide much of a check.

To be fair, the democratic process can have some effect on enhancement laws. The porous elements rule will not stop legislators bent on enhancements, however, nor is it needed to trigger democratic scrutiny. A simple rule of statutory construction, or even a clear-statement rule, could force legislators to advert to the issue before they resolve it.²⁷¹ No

269. William N. Eskridge, Jr., *Overriding Supreme Court Statutory Interpretation Decisions*, 101 YALE L.J. 331 (1991); see also King & Klein, *supra* note 3 (manuscript at 25-28 & app. A) (describing an empirical survey of seven past Supreme Court decisions upholding criminal procedures, and finding that after those rulings, substantially more states adopted the upheld procedures). Legislative action is even more likely in this case than in those studied by King and Klein, because the latter merely confirmed the status quo. In contrast, the elements rule upsets the status quo, so legislatures are more likely to seek a way to return to the previous equilibrium. *But see* King & Klein, *supra* note 3 (manuscript at 28-31) (noting that it is easy for legislatures to circumvent the elements rule as to “nested” enhancements that apply to only one or a few statutes but much harder to do the same with “add-on” enhancements that apply to the entire criminal code, as this would require rewriting the whole code).

270. The first time that Justice Stevens suggested this approach, in *McMillan*, Justice Marshall rejected the “formalistic distinction between aggravating and mitigating facts,” even though both Justices were in dissent. While Justice Marshall conceded that the “democratic process” might perhaps provide some protection, he was well aware that it might not. *McMillan v. Pennsylvania*, 477 U.S. 79, 94 (1986) (Marshall, J., dissenting).

271. These rules of construction, sometimes dubbed “quasi-constitutional law,” have many advantages over bright-line bans. See generally William N. Eskridge, Jr., *Public Values in Statutory Interpretation*, 137 U. PA. L. REV. 1007, 1028-36 (1989); William N. Eskridge, Jr. & Philip P. Frickey, *Quasi-Constitutional Law: Clear Statement Rules as Constitutional Lawmaking*, 45 VAND. L. REV. 593 (1992). They force legislatures to advert to issues that might otherwise have gotten lost in the legislative fracas. They require clearer resolutions, which are more transparent to public scrutiny. Yet they are more respectful of the judgment of a coordinate branch of government. If, upon consideration, legislators announce that a measure is necessary, the rules of construction heed that considered choice. They do not force a legislature bent on a result to contort and twist the law to reach the desired result. They are also more flexible. Courts

rigid ban on sentence enhancements is needed to do so. And the intuitions that are troubling the Court stem from two narrow categories of cases that are better handled via narrow rules. First, in *Mullaney* and *Jones*, the Court worried about letting legislatures redefine traditional common-law elements. Second, in his *Monge* hypothetical, Justice Scalia worried about enhancing a misdemeanor to a felony punishable by life imprisonment. Subsection III.B.3 suggests two narrower, fine-tuned rules that would target these extreme situations without imposing the elements rule's much broader ban.

In short, making law by parade of horrors is dubious and rests on a premature distrust of legislatures. Even if the slippery slope is a problem, the elements rule is not the solution because it is easy to evade. This manipulable rule is no bulwark against judicial arbitrariness. On the contrary, it will likely increase arbitrariness by giving prosecutors more power. Subsection III.A.2 explores this problem in detail.

C. Notice

The Court devoted the least attention to a third argument: the need for notice.²⁷² In *Apprendi*, the Court suggested that indictments must allege enhancements so that defendants have adequate notice of them. The Court praised the common law's determinate sentences because they allowed defendants to predict their sentences from the faces of the indictments.²⁷³ Later, the Court extended this point to indeterminate sentencing within a range, in which the statutes of indictment tell defendants the maximum sentences they face.²⁷⁴ Concurring, Justice Scalia devoted a paragraph to

can over time calibrate the level of presumption to the imminence of a parade of horrors, instead of preemptively foreclosing experimentation. And courts can be more sensitive to feedback they receive from the world of guilty pleas.

Here, a rule of construction could foster an open, democratic debate on the role of elements and juries in criminal justice. It would ensure truth in labeling, making the law more transparent. It would accord with the common-law rule of construction in favor of elements. It would also give the Court time to watch the real-world problems of moving facts from sentencing to the plea phase. And if the dire predictions of Sections II.D and III.A came true, Congress and courts could adapt more easily by moving facts back to sentencing. The procrustean elements rule makes it impossible to do that.

272. *Apprendi* may have downplayed this issue because *Apprendi* himself was warned repeatedly, in his plea agreement and at the plea colloquy, that the prosecution would seek the hate-crime enhancement. See Joint Appendix, *supra* note 137, at 17-24. In addition, *Apprendi* arose from a state court, so it did not present the question of a right to indictment. See *supra* note 150. The same might explain *Monge*'s omission, where state law provided for advance notice. See *Monge v. California*, 524 U.S. 721, 740 (Scalia, J., dissenting). But that is no explanation for why *Jones* paid no heed at all to the notice issue.

273. *Apprendi v. New Jersey*, 120 S. Ct. 2348, 2356 (2000) ("The defendant's ability to predict with certainty the judgment from the face of the felony indictment flowed from the invariable linkage of punishment with crime." (citing 4 WILLIAM BLACKSTONE, COMMENTARIES *369-70)).

274. *Id.* at 2359 n.10.

notice as an essential part of fairness: The law should “tell a prospective felon that if he commits his contemplated crime he is exposing himself to a jail sentence of 30 years . . . [to ensure that] the criminal will never get *more* punishment than he bargained for when he did the crime.”²⁷⁵

Commentators have likewise justified the elements rule as a way of giving notice. Priester, for example, endorses the elements rule in part because it gives notice of the maximum possible sentence.²⁷⁶ Similarly, Rosenberg notes that defendants need notice to decide whether to go to trial or plead guilty.²⁷⁷

The Court was rightly concerned with notice, because notice is a core concern of the Constitution. The Framers thought it so essential that the Bill of Rights guarantees notice to criminal defendants not once, but several times. The Fifth Amendment’s Grand Jury Clause requires notice of all “infamous” crimes via indictment or presentment.²⁷⁸ One of the key purposes of this clause is to notify defendants of the charges against them.²⁷⁹ The Sixth Amendment’s Notice Clause is even more explicit: “In all criminal prosecutions, the accused shall enjoy the right . . . to be informed of the nature and cause of the accusation.”²⁸⁰ This clause is likewise designed to notify defendants of the charges they face.²⁸¹ And the Due Process Clauses also require notice.²⁸²

275. *Id.* at 2367 (Scalia, J., concurring). Justice Scalia’s phrasing of this point is curious. He presumes, as the law generally does, that prospective felons are aware of the sentences they face. (I suspect this is in large part a legal fiction; empirical work is needed to quantify the extent to which this is true.) But if we presume that the felon is aware of the substantive prohibition, why not also presume awareness of penalty enhancements? Justice Scalia would be on firmer ground pointing out that felons who have already entered the criminal justice system need advance notice. These felons may receive notice (via indictment or arraignment) of maximum sentences but not enhancements. Indeed, Justice Scalia’s reference to the punishment the criminal “bargained for” inadvertently suggests when the criminal truly needs more notice: when entering a plea bargain or other guilty plea. *See infra* Subsections III.A.4, III.B.1.

276. *See* Priester, *supra* note 3, at 293 (apparently agreeing with Justice Scalia that “[t]he defendant’s most important interest . . . is understanding the government’s accusation in terms of the maximum possible sentence to which the defendant is exposed”).

277. *See* Rosenberg, *supra* note 3, at 487.

278. U.S. CONST. amend. V (“No person shall be held to answer for a capital, or otherwise infamous crime, unless on a presentment or indictment of a Grand Jury, except in cases arising in the land or naval forces, or in the Militia, when in actual service in time of War or public danger.”).

279. *See* Schmuck v. United States, 489 U.S. 705, 718 (1989).

280. U.S. CONST. amend. VI.

281. *See* United States v. Cruikshank, 92 U.S. 542, 557-58 (1875).

282. *See* U.S. CONST. amend. V (“[N]or [shall any person] be deprived of life, liberty, or property without due process of law.”); *id.* amend. XIV, § 1 (“[N]or shall any State deprive any person of life, liberty, or property without due process of law.”). “The core of due process is the right to notice and a meaningful opportunity to be heard.” *LaChance v. Erickson*, 522 U.S. 262, 266 (1998); *accord* *Cleveland Bd. of Educ. v. Loudermill*, 470 U.S. 532, 542 (1985); *Mullane v. Cent. Hanover Bank & Trust Co.*, 339 U.S. 306, 313 (1950). These same due process rights apply in criminal cases as well. *See Oyler v. Boles*, 368 U.S. 448, 452 (1962) (recognizing that due process requires “reasonable notice and an opportunity to be heard” on a recidivism enhancement, but holding that notice is reasonable even if not given until after trial).

Defendants need this notice to prepare their defenses.²⁸³ Charges must be precise enough “that the defendant . . . may be enabled to determine the species of offence they constitute, in order that he may prepare his defence accordingly.”²⁸⁴ Unfortunately, sometimes a defendant does not get notice of enhancements in time to “prepare his defence.” Thus, in *Jones*, the indictment did not mention the enhancement and the arraigning magistrate misled Jones, telling him the maximum sentence was fifteen years.²⁸⁵ It was not until after conviction, in the presentence report, that Jones received notice of the enhancement.²⁸⁶ In other cases, defendants likewise received notice only after pleading guilty.²⁸⁷ The Supreme Court upheld this practice, reasoning that postplea notice gave defendants enough time to contest enhancements at sentencing.²⁸⁸ This reasoning is flawed, however. Defendants need notice not only to prepare for sentencing, but also to decide whether to go to trial or plead guilty. Notice after verdict is too late to inform this decision.²⁸⁹

It is not clear that lack of notice is a widespread problem in practice, as judges, prosecutors, and defense lawyers often advise defendants of enhancements. Assume arguendo, with the Court and commentators, that it is a real problem. Even if notice is a problem, the elements rule is not the solution.²⁹⁰ First, giving talismanic significance to the charges in the

283. See *United States v. Simmons*, 96 U.S. 360, 362 (1877).

284. *Apprendi v. New Jersey*, 120 S. Ct. 2348, 2356 (2000) (quoting ARCHBOLD, *supra* note 153, at 44) (internal quotation marks omitted).

285. *Jones v. United States*, 526 U.S. 227, 230-31 (1999).

286. See *id.* at 231.

287. See, e.g., *Oyler*, 368 U.S. at 449-54; *Graham v. West Virginia*, 224 U.S. 616, 624, 629 (1912).

288. See *Oyler*, 368 U.S. at 453-54; *Graham*, 224 U.S. at 625.

289. See *infra* Subsections III.A.4, III.B.1; cf. *Bandy v. Hehn*, 67 P. 979 (Wyo. 1902) (concerning a defendant who was told only after pleading guilty to petit larceny that he would receive a recidivism enhancement, resulting in imprisonment; upon receiving notice, the defendant asked to withdraw his plea of guilty, but the court said it was too late to do so).

290. One might object that the Grand Jury Clause does not apply to the states. See *supra* note 150. Thus, the Court’s rule of notice by indictment technically would not help the many defendants prosecuted in state court. The Court has already held, however, that the Sixth Amendment requires notice of all elements of an offense, which now include facts that enhance maximum sentences. See *United States v. Cruikshank*, 92 U.S. 542, 558 (1875) (holding that the Sixth Amendment requires the allegation of “every ingredient of which the offence is composed” (quoting *United States v. Cook*, 84 U.S. (17 Wall.) 168, 174 (1872))). Thus, all states must now give notice of enhancements. They need not, however, give this notice in the charging instrument. The Sixth Amendment analysis looks beyond the face of the charging instrument, taking into account all actual notice and the likelihood of actual prejudice. Thus, as a constitutional matter, the required notice could come via a bill of particulars, discovery, or other means. Some state pleading rules would allow other forms of notice to supplement the charging instrument in considering whether an omission was a reversible error. Others would treat any omission from the charging instrument as per se reversible error. See WAYNE R. LAFAVE ET AL., *CRIMINAL PROCEDURE* 884-85, 888 (3d ed. 2000).

One should separate Sixth Amendment notice from due process notice. The former is limited to elements of the offense. It gives notice so that a defendant can procure discovery on those elements and prepare for trial or plea. The latter can extend beyond the elements of the offense to

indictment gives prosecutors power to manipulate statutory maxima through charge bargaining.²⁹¹ Second, factual allegations in the indictment, while helpful, may not be enough to give notice. In choosing whether to plead guilty or go to trial, defendants need to know how many years in prison they face. A good attorney should deduce the enhanced statutory maximum from the indictment, but the law needs safeguards in case the attorney fails to do so or makes an error.

While in some circumstances the elements rule will not give enough notice, in other circumstances it will broadcast too much. Specifically, the indictment may taint jurors by reciting prejudicial facts about the defendant, such as racial bias. In many of the cases cited by Justice Thomas, defendants objected to allegations of recidivism in the indictment. Most courts brushed off these challenges, reasoning that, because recidivism was an element, the jury was entitled to know about it.²⁹² If the Court were to extend the elements rule to recidivism, as Justice Thomas advocated, this problem would recur. A defendant has a right to stipulate to prior convictions to keep the jury from learning the details of his recidivism.²⁹³ Even so, the jury could be prejudiced just by hearing the recidivism allegation in the indictment and jury instructions.²⁹⁴ By requiring notice in the indictment, instead of via other means, the elements rule needlessly opens the door to jury prejudice.

In short, defendants do need more notice, but the elements rule is not tailored to this need. The rule gives too little notice in some cases and too much in others. Subsections III.A.4 and III.B.1 pick up these themes, exploring what rule might better serve the constitutional value of notice.

D. *Practical Problems at Trial and on Habeas Corpus*

The dissenters raised a host of practical objections to the elements rule. For one, as noted, juries may be prejudiced just by hearing of

include sentencing factors. But under current law, due process notice of sentencing factors can wait until after a trial or guilty plea, so long as it is given before sentencing. *See infra* Subsection III.B.1.

291. *See infra* Subsection III.A.2.

292. *See, e.g.*, *People v. Coleman*, 79 P. 283, 284 (Cal. 1904); *People v. Ratner*, 153 P.2d 790, 790 (Cal. App. Dep't Super. Ct. 1944) (rejecting what appears to have been a prejudice argument raised by the defendant); *State v. Scheminsky*, 174 P. 611, 611 (Idaho 1918); *People v. Sickles*, 51 N.E. 288, 289 (N.Y. 1898); *Johnson v. State*, 55 N.Y. 512 (1874).

293. *Old Chief v. United States*, 519 U.S. 172, 190-92 (1997).

294. When the prior conviction is an element, the jury needs to know the fact of the prior conviction. *See id.* at 178-79. At a minimum, it will learn this fact from the stipulation, jury instructions, and indictment. *See Almendarez-Torres v. United States*, 523 U.S. 224, 235 (1998); *Old Chief*, 519 U.S. at 191 n.10. Thus, it seems a defendant cannot keep the fact away from the jury entirely, say by agreeing to redact the entire allegation out of the indictment. Even if courts redacted recidivism allegations, this would do nothing for nonrecidivism enhancements, because *Old Chief* does not require stipulations to these facts. *Old Chief*, 519 U.S. at 190-92.

enhancements, let alone hearing evidence about them. For another, defendants face difficulties arguing alternative, inconsistent defenses to juries. The Court dismissed both of these objections, suggesting that bifurcated proceedings could shelter juries from prejudice. Bifurcation, however, is not as simple as the Court suggests. The elements rule also threatens judges' role in capital sentencing, and it is already unleashing a wave of collateral attacks. This Section looks at these four objections and finds merit in each.

First, as argued, mentioning a defendant's racial biases and the like risks prejudicing the jury.²⁹⁵ If the defendant stipulates to the enhancement, the jury will learn only the name of the enhancement from the indictment and stipulation.²⁹⁶ Even this information can be prejudicial, however.²⁹⁷ If the defendant must stipulate the right away to avoid prejudice, the elements rule is of little effect. And as a rule, prosecutors need not agree to stipulate. Except for prior convictions, they can prove facts as they see fit.²⁹⁸ If a prosecutor or defendant refuses to stipulate, the prosecution will introduce evidence that would formerly have been admissible only at sentencing. Proof of enhancements such as racial bias will often implicate a defendant's character, prior bad acts, and crimes. This evidence had been strictly limited because of its danger of prejudice.²⁹⁹ By making sentencing issues into elements of guilt, the elements rule has made inflammatory sentencing evidence relevant and admissible to prove guilt.

Second, defendants find it difficult to argue in the alternative to juries. Under the old law, a defendant could stand mute or deny guilt at trial; admit guilt while contesting enhancements at a sentencing hearing; and beg for mercy at sentencing. Now, the elements rule lumps together the first two issues. A defendant may be in the unenviable position of telling a jury: "I didn't sell cocaine. But if I did, it weighed less than five kilos, and I certainly thought it weighed less because my buddy didn't charge the price he normally charges for five kilos. But honestly, I didn't sell cocaine at all."³⁰⁰

295. See *supra* notes 292-293 and accompanying text.

296. See *Old Chief*, 519 U.S. at 190-92 (holding that the defendant has the right to insist upon a stipulation to a prior conviction); see also *Apprendi v. New Jersey*, 120 S. Ct. 2348, 2379 n.10 (2000) (Thomas, J., concurring); *Almendarez-Torres*, 523 U.S. at 269 (1998) (Scalia, J., dissenting).

297. See *Almendarez-Torres*, 523 U.S. at 235.

298. *Old Chief*, 519 U.S. at 186-92.

299. See FED. R. EVID. 404, 608, 609 (regulating evidence of prior bad acts and character); see also *Spencer v. Texas*, 385 U.S. 554, 560 (1967) (stating that evidence of prior crimes "is generally recognized to have potentiality for prejudice").

300. See *Apprendi*, 120 S. Ct. at 2398 (Breyer, J., dissenting); *Monge v. California*, 524 U.S. 721, 729 (1998).

One could imagine responding that pleading in the alternative is not desirable. One could imagine arguing that the common law has been wrong for hundreds of years, and defendants should come out and tell the truth to the jury. One could imagine such radical arguments, but these

Several Justices dismissed these problems, trusting that bifurcated trials and limiting instructions would solve them. By trying the enhancement after the other issues of guilt, they said, courts can insulate juries from prejudice.³⁰¹ Bifurcation, however, is more cumbersome than it sounds. Currently, only capital cases are bifurcated, which explains in part why each capital case costs millions of dollars.³⁰² Moreover, the word “bifurcation” assumes there will be only two parts to a trial. In a case involving multiple enhancements and multiple defendants, however, there could be dozens of separate issues. If each one can prejudice jury consideration of later issues, how can one jury try the dozens of issues seriatim? Which one goes first? Must judges now fragment multidefendant cases into a bushel of one-defendant cases? Bifurcation would also perversely encourage prosecutors to offer the most favorable plea deals to the worst offenders, whose trials would involve the most enhancements and so be most cumbersome.³⁰³

Nor are limiting instructions a panacea. True, the law often presumes that juries follow limiting instructions.³⁰⁴ But “[t]he naive assumption that prejudicial effects can be overcome by instructions to the jury all practicing lawyers know to be unmitigated fiction.”³⁰⁵ Thus, where the prejudice is

are not *Apprendi*'s arguments. *Apprendi* purported to be helping defendants and said nothing about how this rule would hurt them.

301. E.g., *Apprendi*, 120 S. Ct. at 2379 n.10 (Thomas, J., concurring); *Monge*, 524 U.S. at 739 n.1 (Scalia, J., dissenting); *Almendarez-Torres*, 523 U.S. at 269 (Scalia, J., dissenting). Note that before *Apprendi*, at least one court of appeals had held that a defendant could not bifurcate proof of a prior felony conviction from the rest of his trial, because the prior conviction was an element of the offense of being a felon in possession of a firearm. *United States v. Nguyen*, 88 F.3d 812, 818 (9th Cir. 1996). Even if this holding survives *Apprendi*'s approving references to bifurcation, it is unlikely that it will apply beyond the felon-in-possession context.

302. Cost estimates vary widely, but a conservative estimate is that the average capital case costs at least one million dollars, and more likely two or three million dollars, more than the average noncapital case. See Alex Kozinski & Sean Gallagher, *Death: The Ultimate Run-On Sentence*, 46 CASE W. RES. L. REV. 1, 14-15 & n.64 (1995) (surveying the range of estimates); see also PHILIP J. COOK & DONNA B. SLAWSON, *THE COST OF PROCESSING MURDER CASES IN NORTH CAROLINA* 1, 98 (1993); Robert L. Spangenberg & Elizabeth R. Walsh, *Capital Punishment or Life Imprisonment? Some Cost Considerations*, 23 LOY. L.A. L. REV. 45, 58 (1989); Margot Garey, Comment, *The Cost of Taking a Life: Dollars and Sense of the Death Penalty*, 18 U.C. DAVIS L. REV. 1221, 1245-62 (1985); Pamela Manson, *Matter of Life or Death: Capital Punishment Costly; Despite Public Perception, It's Cheaper To Keep Killers in Prison*, ARIZ. REPUBLIC, Aug. 23, 1993, at A1.

The separate penalty phase is a significant component of these capital costs. One of “the greatest expenses associated with capital punishment” is the “more elaborate, two-phased trial proceedings.” Ronald J. Tabak, *How Empirical Studies Can Affect Positively the Politics of the Death Penalty*, 83 CORNELL L. REV. 1431, 1439-40 (1998); see also Justin Brooks & Jeanne Huey Erickson, *The Dire Wolf Collects His Due While the Boys Sit by the Fire: Why Michigan Cannot Afford To Buy into the Death Penalty*, 13 T.M. COOLEY L. REV. 877, 897-900 (1996); Spangenberg & Walsh, *supra*, at 53-54.

303. King & Klein, *supra* note 3 (manuscript at 62-63).

304. E.g., *Blumenthal v. United States*, 332 U.S. 539, 552-53 (1947).

305. *Krulwitch v. United States*, 336 U.S. 440, 453 (1949) (Jackson, J., concurring) (citations omitted).

significant and avoidable, the Court has treated limiting instructions as inadequate.³⁰⁶

A third problem with the elements rule is how it unsettles capital sentencing. Since *Furman*³⁰⁷ and *Gregg*,³⁰⁸ the Court has required specific findings of aggravating facts to impose the death penalty. These aggravating facts are prerequisites to raising the maximum penalty to death and so should be covered by the elements rule.³⁰⁹ Over the last twenty-four years, the Court has repeatedly allowed judges to find these aggravating facts.³¹⁰ Indeed, the Court has praised judicial sentencing, because a judge's greater experience should lead to more consistent sentences.³¹¹ Relying on these decisions, some states have made judicial fact-finding part of their capital penalty-phase proceedings.³¹² The elements rule unsettles these settled procedures. Moreover, it casts serious doubt on every death

306. *E.g.*, *Gray v. Maryland*, 523 U.S. 185, 192-93 (1998); *Bruton v. United States*, 391 U.S. 123, 128-35, 129 (1968) (“[T]oo often such admonition against misuse is intrinsically ineffective in that the effect of such a nonadmissible declaration cannot be wiped from the brains of the jurors. The admonition therefore becomes a futile collocation of words and fails of its purpose as a legal protection to defendants against whom such a declaration should not tell.” (quoting *Delli Paoli v. United States*, 352 U.S. 232, 247 (1957)) (internal quotation marks omitted)); *Jackson v. Denno*, 378 U.S. 368, 388-89 (1964).

307. *Furman v. Georgia*, 408 U.S. 238 (1972).

308. *Gregg v. Georgia*, 428 U.S. 153 (1976).

309. The Court has been inconsistent in characterizing the capital cases. In *Apprendi*, the Court asserted that giving a judge the power to choose between life and death was somehow different because the legislature had already authorized the death penalty based on the jury's conviction. *Apprendi v. New Jersey*, 120 S. Ct. 2348, 2366 (2000); *see* *Almendarez-Torres v. United States*, 523 U.S. 224, 257 n.2 (1998) (Scalia, J., dissenting).

In *Jones*, the Court's wording was much more cautious. The capital precedents, it said, had “characterized the finding of aggravating facts falling within the traditional scope of capital sentencing as a choice between a greater and a lesser penalty, not as a process of raising the ceiling of the sentencing range available.” *Jones v. United States*, 526 U.S. 227, 251 (1999). But the Court saw the conflict between these cases and its elements rule: “We are frank to say that we emphasize this careful reading of *Walton's* rationale because the question implicated . . . is too significant to be decided without being squarely faced.” *Id.*

In his *Apprendi* concurrence, Justice Thomas admitted that the capital cases do involve judicial findings of aggravating facts, which would logically qualify as elements. But he left the question for another day. He noted that the Court has forbidden automatic execution upon a jury's conviction and wondered whether this ban might warrant an exception to the elements rule to compensate somehow for the judicially imposed sentencing phase. *Apprendi*, 120 S. Ct. at 2380 (Thomas, J., concurring).

In his *Jones* concurrence, Justice Stevens agreed that capital aggravating circumstances are elements. He stated explicitly that he would apply the elements rule to capital sentencing. *Jones*, 526 U.S. at 253 (Stevens, J., concurring).

310. *See, e.g.*, *Walton v. Arizona*, 497 U.S. 639, 647-49 (1990); *Hildwin v. Florida*, 490 U.S. 638, 639-40 (1989) (per curiam); *Poland v. Arizona*, 476 U.S. 147, 150 (1986); *Spaziano v. Florida*, 468 U.S. 447, 459-60 (1984); *Barclay v. Florida*, 463 U.S. 939, 944 (1983) (Rehnquist, J., plurality opinion); *id.* at 967-68 (Stevens, J., concurring); *Proffitt v. Florida*, 428 U.S. 242, 252 (1976) (opinion of Stewart, Powell, and Stevens, JJ.).

311. *Proffitt*, 428 U.S. at 252 (opinion of Stewart, Powell, and Stevens, JJ.).

312. *See supra* note 310 (enumerating cases).

sentence imposed by these states over the last twenty-four years. Prisoners can now collaterally attack every one of these death sentences.³¹³

They can likewise attack every noncapital sentence that includes a sentence enhancement.³¹⁴ The impact of these collateral attacks will be enormous. Dozens of federal statutes include sentence enhancements. These statutes include the oft-used laws forbidding drug trafficking, mail and wire

313. A motion under 28 U.S.C. § 2255 is the equivalent for federal prisoners of habeas corpus petitions for state prisoners. *See* 28 U.S.C. § 2255 (1994); *Hill v. United States*, 368 U.S. 424, 427 (1962); 2 CHARLES ALAN WRIGHT, FEDERAL PRACTICE AND PROCEDURE § 591 (2d ed. 1994). For the sake of simplicity, the following discussion will refer to all of these motions collectively as “habeas petitions” or “collateral attacks.”

Ordinarily, *Teague v. Lane* bars prisoners from raising new constitutional claims on collateral attack. 489 U.S. 288, 305-10 (1989) (O’Connor, J., plurality opinion). (*Teague*’s rule was later confirmed by a majority of the Court. *See, e.g., Penry v. Lynaugh*, 492 U.S. 302, 329-30 (1989).) There is an exception for procedures that “implicate the fundamental fairness of the trial” and “without which the likelihood of an accurate conviction is seriously diminished.” *Teague*, 489 U.S. at 312, 313 (O’Connor, J., plurality opinion). The elements rule appears to fall within this exception. The majority viewed the elements rule as a bedrock requirement of a fair trial. And under the elements rule, “an accurate conviction” now includes sentencing facts as elements. Courts have long found these sentencing facts by a preponderance of the evidence rather than beyond a reasonable doubt. The lower standard of proof means the likelihood of an accurate finding of these sentencing facts is seriously diminished. Thus, every capital conviction in which a judge found an aggravating fact is now open to attack. *See United States v. Murphy*, 109 F. Supp. 2d 1059, 1063-64 (D. Minn. 2000) (holding that the elements rule falls under this *Teague* exception and so applies retroactively to § 2255 claims). *But see Jones v. Smith*, 231 F.3d 1227, 1236-38 (9th Cir. 2000) (treating *Apprendi* as a new rule under *Teague* and refusing to apply it retroactively under the watershed exception, at least where a technical omission from the indictment was offset by actual notice to the defendant); Nancy J. King & Susan R. Klein, *Après Apprendi*, 12 FED. SENTENCING REP. 331, 333 (2000) (suggesting that the Supreme Court is unlikely to apply *Apprendi* retroactively because the elements rule goes only to the degree of culpability rather than innocence, but noting that retroactivity is a close question; also noting that *Apprendi* might apply retroactively if litigants can characterize their claims not as procedural claims, but rather as substantive ones about the definitions of crimes).

Assuming the Court treats the rule as retroactive, prisoners will be able to raise claims under the elements rule even if they have filed previous habeas petitions. The Antiterrorism and Effective Death Penalty Act of 1996 ordinarily prevents prisoners from raising claims in a second or subsequent petition. Pub. L. No. 104-132, Title I, §§ 101, 105, 106, 110 Stat. 1217, 1220 (codified at 28 U.S.C.A. §§ 2244, 2255 (West Supp. 2000)). Prisoners can, however, raise claims relying on “a new rule of constitutional law, made retroactive to cases on collateral review by the Supreme Court, that was previously unavailable.” 28 U.S.C.A. §§ 2244(b)(2)(A), 2255 (containing identical language for habeas petitions by state prisoners and § 2255 motions by federal prisoners). For such claims, the one-year statute of limitations does not begin to run until the date on which the Supreme Court initially recognizes the constitutional right. *Id.* §§ 2244(d)(1)(C), 2255(3).

314. *See supra* note 313. True, defendants waive jury trials by pleading guilty, and in their pleas often expressly waive proof beyond a reasonable doubt. *See* FED. R. CRIM. P. 11(c)(3), (4) (requiring the judge, before accepting a guilty plea, to advise the defendant of the right to trial by jury and that a guilty plea waives this right). (While the rule does not require it, as a federal prosecutor I learned that many judges advise defendants that if there were a trial, the government would have to prove their guilt beyond a reasonable doubt.) These waivers, however, may not expressly extend to sentencing facts, as judges may not have described enhancements as part of the crimes to which they were pleading guilty. *Cf.* FED. R. CRIM. P. 11(c)(1) (requiring the judge to advise the defendant of “the nature of the charge to which the plea is offered”); *Bousley v. United States*, 523 U.S. 614, 618-19 (1998) (allowing the defendant to attack collaterally a plea where the judge misdescribed the elements of the offense at the plea colloquy). The elements rule allows and invites prisoners to make these collateral attacks.

fraud, and racketeering.³¹⁵ These challenges have already begun to roll in, and the Court has begun vacating drug sentences.³¹⁶

In addition, the Court's reasoning undermines the validity of judicial findings of aggravating facts under the Sentencing Guidelines. Commentators have asserted, without much explanation, that Guidelines maxima might somehow be different from statutory maxima.³¹⁷ The Court

315. See King & Klein, *supra* note 313, at 336-39 apps. A, C (appendices listing forty-eight federal statutes that appear to have been invalidated by the elements rule, as well as five state gun-enhancement statutes and five state pretrial-release statutes that appear to have been invalidated); see also, e.g., 18 U.S.C. § 34 (1994) (raising the maximum penalty for any crime in title 18, chapter 2 to death when the crime results in death); *id.* § 43(b) (raising the maximum sentence for animal enterprise terrorism from one year to ten years or life imprisonment if there is bodily injury or death); *id.* § 111(a) (raising the maximum sentence for assaulting a federal officer from one to three years if the assault is serious); *id.* § 111(b) (raising the maximum sentence for using a weapon in assaulting a federal officer from three to ten years if the defendant used a deadly weapon or inflicted bodily injury); *id.* § 216 (increasing maximum sentences for various bribery and gratuity offenses from one to five years based on the willfulness of the violation); *id.* § 247 (increasing the maximum sentence for damage to religious property from one year to the death penalty based upon a finding of bodily injury or death); *id.* § 248(b) (raising the maximum sentence for restricting access to an abortion clinic from three to ten years or life imprisonment if the defendant caused serious bodily injury or death); *id.* § 521 (adding ten years to the maximum sentence if the defendant committed a felony offense as part of a criminal street gang); *id.* § 661 (increasing the maximum sentence for maritime theft from one to five years based on the value of the property); *id.* § 924(c) (adding five to thirty years to the maximum sentence for using a gun during a crime of violence or drug trafficking depending on the type and use of the gun) (construed in *Castillo v. United States*, 120 S. Ct. 2090 (2000), as creating separate offenses); *id.* § 924(j) (adding from six years to the death penalty for causing death in the course of violating § 924(c)); *id.* § 982 (authorizing the property forfeiture in addition to the maximum penalty for money laundering and bank secrecy crimes where a crime involves specified property or proceeds); *id.* § 1030(c)(2) (increasing the maximum sentence for computer fraud from one to five years based, *inter alia*, on the value of information); *id.* § 1033 (increasing the maximum sentence for insurance-related crimes from ten to fifteen years if a crime jeopardized the soundness of the insurer); *id.* § 1091(b) (increasing the maximum sentence for genocide from twenty years to the death penalty if death results); *id.* § 1201 (increasing the maximum sentence for kidnapping from life imprisonment to the death penalty if the victim dies); *id.* § 1341 (increasing the maximum sentences for mail and wire fraud from five to thirty years where the violation affects a financial institution); *id.* § 1958 (raising the maximum sentence for a murder-for-hire from ten years to twenty years or the death penalty if personal injury or death results); Racketeer Influenced and Corrupt Organizations Act (RICO), *id.* § 1963(a) (raising the maximum sentence from twenty years to life imprisonment where the underlying racketeering activity would be subject to life imprisonment); 21 U.S.C. § 841 (1994) (raising the maximum sentence for drug trafficking based on the quantity of drugs).

316. Three days after handing down *Apprendi*, the Court vacated and remanded a drug sentence for reconsideration in light of *Apprendi*. *Jones v. United States*, 120 S. Ct. 2739 (2000) (mem.). On the first day of the October 2000 Term, the Court vacated four drug sentences and remanded them for further consideration in light of *Apprendi*. *Burton v. United States*, 121 S. Ct. 32 (2000) (mem.); *Wims v. United States*, 121 S. Ct. 32 (2000) (mem.); *Blue v. United States*, 121 S. Ct. 31 (2000) (mem.); *Gibson v. United States*, 121 S. Ct. 31 (2000) (mem.). The Court has continued to do so since then. See, e.g., *Hughes v. United States*, 121 S. Ct. 423 (2000) (mem.); *Whitt v. United States*, 121 S. Ct. 423 (2000) (mem.).

317. See, e.g., Knoll & Singer, *supra* note 3, at 1061 n.21 (asserting that Congress's failure to adopt formally the Sentencing Guidelines criteria makes them immune from elements-rule scrutiny on purely formalistic grounds); Priester, *supra* note 3, at 292 n.210 (claiming that the Guidelines are different because Congress enacts a "generative statute" but not "nongenerative sentencing regulations").

left open the validity of the Guidelines.³¹⁸ Justice Thomas in concurrence said the rule should reach “every fact that is by law a basis for imposing or increasing punishment.”³¹⁹ He hinted that the rule would invalidate the Guidelines because they operate in effect as statutory maxima.³²⁰ Justice Thomas is correct. There is no logical difference between enhancing maxima set by Congress and maxima set by the Sentencing Commission. The elements rule, taken to its logical conclusion, could invalidate hundreds of thousands if not millions of federal and state guideline sentences.³²¹ These staggering practical consequences will likely deter the Court from extending the elements rule to its logical conclusion. Even if the Court refuses to apply its rule to the Guidelines, the current uncertainty will generate a tsunami of habeas petitions.

* * *

In sum, the elements rule rests on weak foundations and will cause problems both at trial and on habeas corpus. The only good argument is the need for notice, but the rule is not well tailored to this goal. These criticisms and problems are, however, only the tip of the iceberg. The dissent’s criticisms were correct insofar as they went, but shared the majority’s and commentators’ mistaken focus on jury trials. As Part III argues, the elements rule causes much more harm in the real world of guilty pleas. It hurts many of the defendants it purports to help by undermining the constitutional values it purports to safeguard. When one views the issue from this real-world vantage, better solutions emerge.

III. THE REAL WORLD OF GUILTY PLEAS

The majority, concurring, and dissenting opinions in *Jones* and *Apprendi* all treated jury trials as the norm. So did every brief submitted to the Court in *Apprendi*.³²² So does our popular culture.³²³ Commentators

318. *Apprendi v. New Jersey*, 120 S. Ct. 2348, 2366 n.21 (2000).

319. *Id.* at 2368 (Thomas, J., concurring).

320. *Id.* at 2380 n.11 (quoting from Justice Scalia’s dissent in *Mistretta v. United States*, 488 U.S. 361, 413 (1989), which had argued that the Guidelines were unconstitutional).

321. *Id.* at 2395 (O’Connor, J., dissenting) (noting that almost half a million cases have been sentenced under the Federal Guidelines, and that this number is likely dwarfed by the number of sentences under state guideline schemes).

322. Not one of the briefs discussed guilty pleas. See Brief of Petitioner, *Apprendi* (No. 99-478); Reply Brief of Petitioner, *Apprendi* (containing passing references to plea bargains on pages 14 and 15); Brief for Respondent, *Apprendi* (containing a passing reference to guilty pleas on page 39); Brief of Amicus Curiae the Anti-Defamation League, *Apprendi*; Brief of Amici Curiae Brudnick Center on Violence and Conflict, California Association of Human Relations Organizations, Center for Democratic Renewal, Center on Hate & Extremism, Hatewatch, National Association of Human Rights Workers, & Northwest Coalition for Human Dignity, *Apprendi*; Brief of Amici Curiae National Association of Criminal Defense Lawyers and Families Against Mandatory Minimums Foundation and Association of Federal Defenders, *Apprendi*; Brief

share this anachronistic worldview. The articles supporting the elements rule rest not on narrow doctrinal grounds, but on the broad assumption that juries still play a central role in criminal justice. So, for example, the articles exalt “the unique and venerable role of the jury” and “its traditional and revered role as a factfinder.”³²⁴ Knoll and Singer wax eloquent about how “juries exist for a reason,” not pausing to consider whether they still exist.³²⁵ And Rosenberg treats juries as if they remain important safeguards of justice in the typical case.³²⁶

These articles are not alone. Over the last ten years, law reviews have published 633 articles on criminal petit juries, but only 62 about guilty pleas or plea bargaining.³²⁷ In other words, jury-trial literature is ten times as common as guilty-plea scholarship. This disproportion suggests that in the minds of academics, trials are the norm and pleas are the exception.³²⁸ A glance at a leading criminal procedure casebook reflects a similar disproportion. Saltzburg and Capra, for example, devote 313 pages to jury trials, but only 38 to all aspects of pleas.³²⁹

Amicus Curiae of the Rutherford Institute, *Apprendi*; Brief for the United States as Amicus Curiae, *Apprendi*.

323. *E.g.*, INHERIT THE WIND (Metro-Goldwyn-Mayer/United Artists 1960); MY COUSIN VINNY (Twentieth Century Fox 1992); TWELVE ANGRY MEN (Metro-Goldwyn-Mayer/United Artists 1957); THE VERDICT (Twentieth Century Fox 1982); *Ally McBeal* (Fox television series); *L.A. Law* (NBC television series); *Law and Order* (NBC television series).

324. Note, *supra* note 3, at 1362; *see also* Murphy, *supra* note 3, at 43.

325. Knoll & Singer, *supra* note 3, at 1112.

326. Rosenberg, *supra* note 3, at 481-84.

327. On October 17, 2000, I searched the Law Reviews, Combined database on Lexis, searching through all law review articles published between January 1, 1990, and October 17, 2000. The first query retrieved all articles whose titles contained the words “jury,” “juries,” “juror,” or “jurors” and did not contain any of the following terms: “civil,” “Seventh Amendment,” “7th Amendment,” “employment,” “damages,” “tort,” “contract,” “million,” “plaintiff,” “injury,” or “grand jury.” This search yielded 633 results.

I conducted a second search on the same date, in the same database, using the same date restrictions. The second query retrieved all articles whose titles contained the words “guilt” or “guilty” within three words of “plea” or “plead” or “pleading,” and all articles whose titles contained the word “plea” within three words of “bargain” or “bargaining.” This query yielded sixty-two results. Spot checks of both sets of search results indicated that the articles were concerned with criminal juries and plea bargains respectively.

Note that the literature on guilty pleas is even less rich than these numbers suggest. Many of the articles on guilty pleas are simple case comments or annual surveys of criminal procedure. Many others merely decry the existence of guilty pleas, instead of discussing how to make them better within real-world constraints. *E.g.*, articles cited *infra* note 331.

328. I am not making the simplistic argument that the proportions of scholarship should match exactly the percentages of jury trials and guilty pleas. Plea bargaining and decisions to plead guilty happen in part in the shadow of expected trial outcomes, which explains why the average jury trial may be more important than the average guilty plea. But the disproportion set forth in the text is staggering—there is ten times as much scholarship about jury trials, even though there are twenty-four times as many guilty pleas. In other words, the literature is off by a factor of 240. Nor is it an answer to say that pleas are not intellectually interesting, important, or complex. As Part III shows, the dynamics of guilty pleas can be extremely challenging and complex.

329. STEPHEN A. SALTZBURG & DANIEL J. CAPRA, AMERICAN CRIMINAL PROCEDURE: CASES AND COMMENTARY 959-96 (6th ed. 2000) (guilty pleas); *id.* at 997-1309 (trials).

This view of criminal procedure is, however, decades out of date. Fewer than four percent of adjudicated felony defendants have jury trials, and another five percent have bench trials. Ninety-one percent plead guilty.³³⁰ Our world is no longer one of trials, but of guilty pleas. Our legal rules and scholarship must take this development into account, but they do not.³³¹ The elements rule is a classic example: In fashioning an elements rule for the four percent, the Court and commentators ignored how the rule will harm the ninety-one percent. To build a better rule, the law must translate the Constitution into the real world of guilty pleas.

Scholars must also look at how these pleas interact with our sentencing law. None of the articles or decisions delved into how the elements rule would fare under the Sentencing Guidelines or mandatory minima. This is because commentators, like the *Apprendi* Court, view “[c]riminal procedure [as], basically, a subset of constitutional law.”³³² Commentators and law

330. These figures exclude cases in which the prosecution was dropped, dismissed, or otherwise terminated before verdict. Between October 1999 and September 2000, federal courts disposed of 75,071 defendants, of whom 68,079 were disposed of by plea or trial. Of these, 63,863 (93.8%) entered pleas of guilty or nolo contendere, and 1235 (1.8%) were convicted or acquitted after bench trials. Only 2981 (4.4%) were convicted or acquitted after jury trials. *See* STATISTICS DIV., ADMIN. OFFICE OF THE U.S. COURTS, STATISTICAL TABLES FOR THE FEDERAL JUDICIARY SEPTEMBER 30, 2000, tbl.D-4 (2001) [hereinafter 2000 FEDERAL STATISTICAL TABLES]. (These data are not broken down into felonies and misdemeanors. However, the overwhelming majority of federal convictions (82.7%) are felony convictions, so the percentages for felonies are likely close to the overall percentages listed above. BUREAU OF JUSTICE STATISTICS, U.S. DEP’T OF JUSTICE, SOURCEBOOK OF CRIMINAL JUSTICE STATISTICS 1998, tbl.5.17, at 403 (1999). In any event, because the state numbers listed below dwarf the federal numbers, they swamp any small variations in the federal statistics.)

The jury-trial percentages are even lower for state cases. In 1996, the most recent year for which state data are available, the cases of roughly 997,972 state felony defendants were resolved by plea or verdict. Of these, 905,957 (90.8%) entered pleas of guilty or nolo contendere, 54,474 (5.5%) had bench trials, and only 37,541 (3.8%) had jury trials. *Id.* tbl.5.42, at 432. Note that these numbers understate the number of pleas. Sometimes, prosecutors allow felony defendants to plead guilty to misdemeanors, and these misdemeanor figures may not show up in the felony statistics. 1 WAYNE R. LAFAVE ET AL., CRIMINAL PROCEDURE § 1.3(o), at 124 n.207 (2d ed. 1999).

Because the state numbers dwarf the federal numbers, the percentages for the state and federal systems combined are almost the same as for the state data alone: 91.0% pleas, 5.2% bench trials, and only 3.8% jury trials.

331. There is an extensive literature on whether the predominance of guilty pleas is desirable or even constitutional. *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 (1983); Akhil Reed Amar, *The Bill of Rights as a Constitution*, 100 YALE L.J. 1131, 1196-99 (1991) (noting that Article III of the Constitution says “the trial of *all* Crimes . . . shall be by Jury” (emphasis added), and arguing that this mandatory language makes the right nonwaivable); Stephen J. Schulhofer, *Is Plea Bargaining Inevitable?*, 97 HARV. L. REV. 1037 (1984); Stephen J. Schulhofer, *Plea Bargaining as Disaster*, 101 YALE L.J. 1979 (1992).

Wading into this debate is far beyond the scope of this Article, nor is it necessary for present purposes. Legislatures have shown no inclination to abolish plea bargaining anytime soon. While there is a role for scholarship about what the ideal world would look like, there is also a need for scholarship about the world in which we live. In America, plea bargaining happens and will continue to happen for the foreseeable future. Whether they like it or not, scholars must make the best of this real world. This Article is a step toward doing so.

332. Stuntz, *supra* note 1, at 6.

school curricula treat sentencing as a separate course, distinct from lofty criminal procedure. This artificial separation of sentencing from criminal procedure made it easy to overlook the disastrous sentencing consequences of the elements rule.

Section III.A shows how poorly the elements rule fits into the real world of guilty pleas and sentencing. The trial right does little good when most defendants do not go to trial. On the contrary, it does much harm. It undercuts the very values it purports to safeguard, namely notice, hearing, checking arbitrariness, equal treatment, and expressing the community's conscience. Section III.B sketches out an alternative approach, one tailored to the real world of guilty pleas and sentencing. What defendants need above all is notice of the maximum sentences they face, including enhancements, before they plead guilty. This reform, along with procedural safeguards at sentencing and other rules, would fit America's changed legal landscape. These rules would also better serve the values of notice, hearing, checking arbitrariness, equal treatment, and expressing the community's conscience.

A. *Constitutional Values in a World of Guilty Pleas*

The Constitution protects a cluster of values in criminal cases. The Fifth Amendment's Grand Jury Clause and the Sixth Amendment's Notice Clause give notice to defendants, so that they can prepare their defenses.³³³ The Fifth and Fourteenth Amendment's Due Process Clauses guarantee notice and an opportunity to be heard.³³⁴ The Fifth Amendment's Due Process Clause and the Fourteenth Amendment's Equal Protection Clause guarantee equal treatment.³³⁵ And the Jury Clauses of Article III and the Fifth and Sixth Amendments serve many functions. Most importantly, they check arbitrary judges and prosecutors and imbue the law with the conscience of the community.³³⁶

The elements rule does a poor job of implementing these values in the real world of guilty pleas and sentencing. Subsection III.A.1 explains how the rule deprives many defendants of their best issues at sentencing hearings, which are their only real opportunities to be heard. Subsection III.A.2 argues that the elements rule promotes prosecutorial arbitrariness, which leads to unequal treatment. It shows how the rule's subdivision of crimes gives prosecutors more opportunities to manipulate statutory

333. See *supra* notes 278-284 and accompanying text.

334. See *supra* note 282 and accompanying text.

335. *E.g.*, *Batson v. Kentucky*, 476 U.S. 79 (1986) (holding that the defendant has equal protection rights in jury selection); *Bolling v. Sharpe*, 347 U.S. 497, 499 (1954) (holding that the Fifth Amendment's Due Process Clause contains an equal protection component).

336. See *Amar, supra* note 331, at 1183-85, 1187-89.

maxima during plea bargaining to favor some defendants. Subsection III.A.3 explains how the elements rule reduces the role of the conscience of the community. By shackling legislative efforts to give guidance, the rule encourages unfettered delegations of power to unaccountable sentencing commissions. Finally, Subsection III.A.4 turns to notice, noting defects in the notice provided by the elements rule.

1. *Opportunity To Be Heard*

Due process guarantees defendants opportunities to be heard. The ninety-one percent of defendants who plead guilty will never be heard at trial. Their only hearings are at sentencing, where they can dispute various enhancements and other sentencing facts.³³⁷ Or rather, they were able to dispute these enhancements before the elements rule. Now that the Court has made enhancements into elements of crimes, they must surrender these issues and allocute to them when they plead guilty.

The idea that rights can hurt defendants and deprive them of hearings is counterintuitive. But the elements rule does more than confer new rights; in effect, it *mandates* that enhancements be tried at trial if at all. In doing so, it takes away the more valuable right to try enhancements at sentencing after pleading guilty.³³⁸ The elements rule also gives prosecutors two chances to prove enhancements against defendants who go to trial. These are the features that really hurt defendants. Note that what matters is not hurting defendants per se, but depriving them of their rights to be heard under the Due Process Clause. The elements rule hurts defendants by making it too costly for them to exercise their rights to be heard.

The elements rule does confer a countervailing benefit, namely a right to have enhancements proved beyond a reasonable doubt at trial. But prosecutors can circumvent this right by trying again to prove

337. See FED. R. CRIM. P. 32(c)(1). True, defendants who plead guilty are also heard at plea hearings when they allocute to the offenses they have committed. See FED. R. CRIM. P. 11(c). These hearings, however, are largely pro forma. The parties do not put on witnesses or introduce other extrinsic evidence, apart from the defendant's own statement (and perhaps representations by counsel), and the defendant has no opportunity to contest any matters. The only opportunity for a contested or evidentiary hearing is at sentencing.

In many jurisdictions, a defendant has a right to an adversarial preliminary hearing. In the federal system and many state systems, however, prosecutors can and do bypass these hearings by obtaining indictments before the dates for preliminary hearings. *E.g.*, FED. R. CRIM. P. 5(c) (stating that a hearing before a magistrate judge is not required where an indictment was previously obtained). And even when preliminary hearings do take place, the standard of proof is too low to afford much of a safeguard. Prosecutors need show only probable cause for the magistrate judge to believe that the defendant committed the charged crime. LAFAVE ET AL., *supra* note 290, at 714.

338. This is because a defendant who pleads guilty must allocute to every element of the offense. Because enhancements are now elements, defendants must allocute to them as well and are then estopped from relitigating the issues at sentencing.

enhancements at sentencing under a lower standard of proof, so few defendants will benefit.³³⁹ And few defendants can afford to go to trial to exercise the reasonable-doubt right, while many more could have had sentencing hearings but have now lost that right. Because defendants' hearing rights benefit little from the reasonable-doubt standard, while suffering much from the deprivation of sentencing hearings, defendants on the whole will be worse off.

a. The Status Quo Ante. To explain, it will help first to describe the federal and similar state systems, before discussing what the elements rule has changed. A federal defendant in essence has a right to two trials: First, there is a jury trial, at which a jury must find the elements of the offense beyond a reasonable doubt. Second, there is a sentencing hearing, at which a judge must find sentencing facts by a preponderance of the evidence. It is helpful to think of these as "the first trial" and "the second trial," respectively.

The benefits of waiving the first trial by pleading guilty are massive. First, timely guilty pleas earn defendants reductions of two or three levels under the Sentencing Guidelines.³⁴⁰ Second, most judges reward defendants who plead guilty by sentencing them at the bottom of the applicable sentencing ranges. These two benefits together give defendants roughly a thirty-five percent (or more) reduction for pleading guilty.³⁴¹ Third, in drug cases, prosecutors do not file recidivism enhancements against defendants who plead guilty. These enhancements would trigger mandatory minima of ten years, twenty years, or life imprisonment. Most prosecutors reserve these enhancements for defendants who force them to go to a first trial.³⁴² Fourth, defendants who plead guilty avoid possible

339. The prosecution can again try to prove these issues at sentencing in order to enhance the defendant's Sentencing Guidelines calculations, subject only to the statutory minimum and maximum set by the jury's verdict at trial. *See infra* notes 348-352 and accompanying text.

340. U.S. SENTENCING GUIDELINES MANUAL § 3E1.1 (1998) (giving a two-point reduction for acceptance of responsibility, plus a third point for a timely guilty plea where the adjusted offense level is sixteen or higher); Michael M. O'Hear, *Remorse, Cooperation, and "Acceptance of Responsibility": The Structure, Implementation, and Reform of Section 3E1.1 of the Federal Sentencing Guidelines*, 91 NW. U. L. REV. 1507, 1534 (1997) (noting that in many but not all federal districts, judges automatically award acceptance-of-responsibility discounts to all defendants who plead guilty).

341. *See* O'Sullivan, *supra* note 248, at 1415 & n.274 (collecting sources for the thirty-five percent figure). My own inspection of the Sentencing Guidelines table shows that, for long sentences, the discount is somewhat more—sometimes up to forty to forty-five percent. *See* U.S. SENTENCING GUIDELINES MANUAL ch. 5, pt. A (1998) (sentencing tables). Note that, if judges saw the harsh consequences of the elements rule, they might sympathize with defendants' predicaments and be a little more lenient on defendants who go to trial. But under the prevailing view, which holds that the elements rule is favorable to defendants, the judges would be just as harsh if not harsher. By bringing the adverse consequences of the elements rule to light, this Article may change that dynamic. (This is loosely analogous to a general principle of quantum physics: The act of measuring a phenomenon changes the phenomenon being measured.)

342. *See* 21 U.S.C. § 851(a)(1) (1994) (giving prosecutors complete power over whether or not to file a recidivism enhancement, which is known as a prior felony information). As a federal

obstruction-of-justice enhancements for perjurious trial testimony.³⁴³ Fifth, prosecutors may implicitly reward defendants for pleading guilty by choosing to forgo other enhancements as well.³⁴⁴ A defendant who insists on a first trial on any factual issue (including enhancements) and loses then forfeits these massive benefits.³⁴⁵

prosecutor in the U.S. Attorney's Office for the Southern District of New York, I learned and implemented the unwritten policy that, except as described in the next paragraph, we were never to file prior felony informations against defendants who timely pleaded guilty but were always to file these informations several weeks before trial. As I understood it, this policy applied even when a defendant pleaded guilty to certain counts of an indictment (such as base offenses), so long as the defendant was insisting on a trial on one or more counts (such as enhanced offenses) that were subject to prior felony informations. Some other prosecutors' offices may have left the decision about whether to file prior felony informations up to the discretion of individual prosecutors. I am told that at least one office, the U.S. Attorney's Office for the District of Connecticut, requires prosecutors to file prior felony informations in all cases except where they receive specific permission not to do so. (I am indebted to Andrew Levine for the information in the preceding sentence.) This last policy removes the coercive pressure to plead guilty in return for massive sentence increases for all recidivists. But where office policy gives prosecutors discretion over prior felony informations, they are likely to use them as prosecutors do in the Southern District of New York. In other words, prosecutors have strong incentives to lighten their dockets by forgoing these enhancements in return for guilty pleas (see the last paragraph of this footnote).

Prosecutors also file prior felony informations against defendants cooperating in the prosecution of other individuals. These informations give the prosecution added leverage and thereby ensure continued truthfulness and cooperation. If the cooperator remains cooperative, however, the prosecution then files a motion pursuant to 18 U.S.C. § 3553(e) (1994). This motion gives the court the power to impose a sentence below the statutory minimum as a reward for cooperation and so counteracts the prior felony information.

One could ask, if prosecutors have such powers under current law, why do they not use them to extract as much prison time as possible out of defendants? Why is the gulf between trial and plea numbers so vast? Why do prosecutors not turn some of this gulf into increased prison time for pleading defendants? The answer is that the prior felony information is a sledgehammer, not a scalpel. The amount of time involved is so large and packaged in such discrete chunks (ten years, twenty years, or life) that it cannot be parceled out. Using it forces defendants to trial. Thus, the prior felony information is a powerful deterrent to trial but cannot be used more subtly to influence the terms of the plea. By restraining themselves and hewing to the trial/plea line, prosecutors make the trial deterrent all the more credible. If the prosecutors exercised the power more wantonly, they would force more defendants to trial, something few prosecutors want to do. Besides, prosecutors' offices must control the use of this sledgehammer via clear guidelines. The trial/plea distinction is a much easier and more administrable line to draw than various nuances about when prior felony informations can or cannot be used as bargaining chips.

343. See U.S. SENTENCING GUIDELINES MANUAL § 3C1.1 & application note 4(b) (1998) (mandating a two-point enhancement for obstruction of justice, including perjury).

344. See, e.g., *id.* § 2D1.2(a)(1) (adding two points to the base offense level for drug crimes committed near protected locations or involving pregnant or underage people). The prosecution is unlikely to ask for this enhancement if the defendant pleads guilty, and the Probation Department and court may fail to raise the issue if the prosecution has not. (Why do prosecutors not use school-zone enhancements in guilty pleas? Perhaps out of a sense that it is piling Pelion upon Ossa; in other words, that it is excessive to force such an enhancement upon a defendant who already agrees to a massive amount of time and spares the government a trial. Prosecutors do not always seek the absolute heaviest sentences possible; often, they act as if their main goal were to avoid trial.) Note that the school-zone enhancement reaches all offenses within 1000 feet of a school or college. 21 U.S.C. § 860(a). In most urban areas, this area encompasses almost the entire city.

345. See U.S. SENTENCING GUIDELINES MANUAL § 3E1.1 application notes 2, 3 (1998) (stating that a guilty plea is "significant evidence" of acceptance of responsibility, but that the

A defendant can ask for a bench rather than a jury trial. In the federal and many state systems, however, the prosecution must consent to this request.³⁴⁶ Whether it is a jury or a bench trial, the first trial costs the defendant the benefits of pleading guilty, but the second, sentencing trial does not. The defendant can waive the right to a first trial, pleading guilty in exchange for various benefits, while still enjoying a second trial at sentencing. Before the elements rule, enhancements were issues for the second trial, not the first one. So a defendant was able to get all the rewards for waiving the first trial while still contesting enhancements at the second trial. The judge finds facts at the second trial, and the judge (unlike a jury) knows what sentencing consequences flow from particular findings. If the

reduction does not apply to defendants who deny elements of factual guilt at trial; drawing no distinction between bench and jury trials); *see also* United States v. Portillo-Valenzuela, 20 F.3d 393, 395 (10th Cir. 1994) (holding that “in all but rare cases going to trial will preclude reduction for acceptance of responsibility”). The requirement of pleading to the entire crime has an odd effect. After the elements rule, a defendant may be willing to allocute to exactly as much as she was willing to admit before the elements rule. But now that the elements rule has expanded crime *X* to include element *Y*, a defendant who will allocute to *X* but not *X + Y* can no longer plead guilty and so cannot get the benefits of a guilty plea.

Defendants may try to circumvent this problem by entering partial guilty pleas. In other words, they might plead guilty to base offenses while going to trial on enhanced offenses. Though the Guidelines do not speak to this precise scenario, this maneuver might work in part. If defendants were acquitted of the enhancements, they would likely get reductions for acceptance of responsibility on the base offenses to which they pleaded guilty. Prosecutors might perhaps also file prior felony informations only as to the enhanced offenses, so that acquittals on the enhanced offenses would avert the recidivism enhancements. Nonetheless, judges would likely continue to sentence closer to the tops of the applicable ranges to penalize defendants for burdening them with cumbersome jury trials. *See supra* note 341. Moreover, partial guilty pleas have their own problems: Juries may be skeptical of defendants’ admissions that they sold drugs, but not quite as much as they were charged with. And if defendants lose on enhancement issues, they lose all the benefits of pleading guilty and may also face obstruction-of-justice enhancements if they falsely denied the enhancement issues during their plea allocutions.

346. *See* George Fisher, *Plea Bargaining’s Triumph*, 109 YALE L.J. 857, 1072 (2000); Adam H. Kurland, *Providing a Federal Criminal Defendant with a Unilateral Right to a Bench Trial: A Renewed Call To Amend Federal Rule of Criminal Procedure 23(a)*, 26 U.C. DAVIS L. REV. 309, 321-23 & nn.39, 40, 42, 43 & 45 (1993) (listing eight states that allow bench trials upon a defendant’s unilateral request; thirteen states that allow bench trials with the consent of the judge; twenty-three states plus the District of Columbia that require prosecutorial and court approval; five states that require prosecutorial but not court approval; and one state that forbids waivers of felony jury trials). In those states where prosecutors cannot veto bench trials, judges have substantial power to mitigate the harmful effects of the elements rule. Bench trials are speedier than jury trials and so are less likely to antagonize the judge. Moreover, in states without sentencing guidelines, judges have more discretion in sentencing. In states with neither guidelines nor prosecutorial vetoes on bench trials, defendants can ask for bench trials and agree to stipulate to all elements except enhancements (such as drug quantity). Bench trials on enhancements could be very brief affairs, and judges might reward defendants for sparing them jury trials, reducing sentences almost as much as if the defendants had pleaded guilty. Of course, much of this leniency depends on the judge. If the judge views the defendant as causing trouble by exercising the new elements-rule right and going to trial, the judge likely will not be sympathetic. If, on the other hand, the judge sees the defendant as simply trying to get around the harmful effects of the elements rule, the judge may be more sympathetic. Thus, if this Article succeeds in changing the prevailing view that the elements rule is favorable to defendants, it may incline judges to be more sympathetic and so to soften some of the direr outcomes predicted in this Subsection. *See supra* note 341.

judge felt the prosecution was being too harsh, the judge was able to reduce the penalty by finding enhancement facts in favor of the defendant. (It was particularly easy to do so on amorphous enhancements, such as whether a defendant was motivated by racial bias or foresaw a coconspirator's quantity of drugs.) This power operated as a check on prosecutors and encouraged them to strike more generous deals.³⁴⁷ Prosecutors in effect had to share plea-bargaining power with judges and accommodate their sense of justice. And it was easy for defendants to secure hearings.

b. The Changes Wrought by the Elements Rule. The elements rule changes this scheme in three ways. It creates rights to (1) indictment and (2) proof beyond a reasonable doubt at a (3) jury trial.

(1) The indictment right does not affect the defendant's sentence or hearing rights, unless the prosecutor engages in charge bargaining as described in Subsection III.A.2. I disregard this right for the time being.

(2) The reasonable-doubt standard helps defendants' rights to hearings, but not nearly as much as one might think. Even if a jury convicts a defendant of the base offense and acquits of the enhanced crime, the jury's acquittal is not dispositive. The conviction of the base offense and acquittal of the enhanced offense at the first trial set the statutory minimum and maximum penalties. Within this broad statutory range, the judge considers far more information at the second, sentencing trial in determining the sentence under the Sentencing Guidelines. Thus, the prosecution gets a second bite at the apple.³⁴⁸ At sentencing, the prosecution can ask the judge to sentence the defendant based on wrongdoing that was not included in but related to the offense of conviction, in addition to the offense of conviction itself. The prosecution can use uncharged conduct, conduct over which there was no jurisdiction, and even conduct of which the defendant was acquitted as part of the broader evidence of *relevant conduct*.³⁴⁹ The

347. See Fisher, *supra* note 346, at 1071.

348. Note that there is a ratchet effect at work here, so that this phenomenon operates in only one direction. For example, a defendant who pleads guilty may now have to admit that she had at least five kilograms of cocaine, estopping her from later claiming a lower drug quantity. Likewise, a jury's conviction at the first trial announces that there was proof beyond a reasonable doubt of at least five kilograms of cocaine, estopping later relitigation of that fact. If there is proof beyond a reasonable doubt, a fortiori there is the proof by a preponderance of the evidence needed for the second trial. But the lower standard of proof at the second trial introduces an asymmetry. So while a conviction at the first trial estops relitigation at the second one, an acquittal does not. A jury's acquittal of an enhanced offense and conviction of a base one shows only that there was no proof beyond a reasonable doubt. The prosecution can relitigate the very same issue at the second trial because of the lower standard of proof. See *United States v. Watts*, 519 U.S. 148, 155-57 (1997) (per curiam). Thus a defendant cannot seek to lower the amount of relevant conduct at sentencing, while the prosecution can seek to raise it.

349. Under the Sentencing Guidelines, judges sentence based on all relevant conduct. Relevant conduct is much broader than the offense of conviction. It includes conduct on which the jury acquitted but which the judge finds was proved by a preponderance of the evidence. U.S. SENTENCING GUIDELINES MANUAL §§ 1B1.3, 1B1.4 (1998); see 18 U.S.C. § 3661 (1994) (providing that courts at sentencing must be able to consider without limitation all information

prosecution can use any and all of these bases to enhance the sentence, subject only to the statutory maximum for the offense of conviction.³⁵⁰ At this second trial, the prosecution need prove relevant conduct by only a preponderance of the evidence, not beyond a reasonable doubt.³⁵¹

So the higher standard of proof at the first trial helps only when the defendant's Sentencing Guidelines range is bumping up against the statutory maximum. In such a case, the statutory maximum set at the first trial caps the use of relevant conduct at the second trial. When the defendant is not up against the maximum, the prosecution can try again to prove enhancements at the second trial by a preponderance, undercutting the reasonable-doubt standard at the first trial.³⁵² Thus, in areas where statutory maxima are already very high, such as narcotics, the elements rule does defendants almost no good. And it also means the legislature can get around the burden of proof simply by raising the maxima. Prosecutors can likewise get around the rule by effectively raising the statutory maxima. They can do this by charging multiple counts, then seeking consecutive sentences that aggregate to what the enhanced sentence would have been.³⁵³

Nonetheless, if the Court had done nothing but raise the standard of proof for enhancements, defendants would have been clear winners (setting aside possible legislative circumvention). They would have had all the advantages of the old system plus some benefits from a higher standard of

about defendants' "background, character, and conduct"); *Watts*, 519 U.S. at 152-57 (holding that courts may consider acquitted conduct as part of the relevant conduct at sentencing); *Witte v. United States*, 515 U.S. 389, 397-98, 402-04 (1995) (rejecting a double jeopardy challenge to a prosecution for cocaine trafficking based on conduct that had previously been used as relevant conduct to enhance a marijuana-trafficking sentence within the statutory range). And it goes beyond conduct that was or could have been charged, to include all other conduct in furtherance of a joint criminal activity that was reasonably foreseeable. U.S. SENTENCING GUIDELINES MANUAL § 1B1.3(a)(1)(B) (1998). Thus, for example, the dollar value of loss attributable to a fraud ring includes not only fraud committed in the jurisdiction, but also fraud committed elsewhere even if there is no federal jurisdiction over those frauds.

350. This limit means little where statutory maxima are already high. For example, because the maxima for drug offenses are so high, they do not help most drug defendants. *See* 21 U.S.C. § 841(b)(1)(B) (1994) (forty-year maximum for between half a kilogram and five kilograms of cocaine); *id.* § 841(b)(1)(C) (twenty-year maximum for less than half a kilogram of cocaine).

351. *See Watts*, 519 U.S. at 152-57; *Witte*, 515 U.S. at 397-404.

352. For an example that illustrates and may clarify the statement in the text, see *infra* Subsection III.A.1.c (recounting the example of Al: Even if the jury acquits Al of having forty-two kilograms of cocaine beyond a reasonable doubt, and convicts him of having only two kilograms, the prosecution gets a second chance to prove the forty-two kilograms to the sentencing judge by a mere preponderance of the evidence; thus, the proof of relevant conduct by a preponderance at sentencing undercuts the reasonable-doubt standard at trial).

353. *See Apprendi v. New Jersey*, 120 S. Ct. 2348, 2400 (2000) (Breyer, J., dissenting) (noting that judges and prosecutors can use uncharged crimes to enhance sentences by, for example, charging five counts of embezzlement, then imposing consecutive maximum sentences to punish an uncharged murder).

proof at the first trial.³⁵⁴ The elements rule would have strengthened their rights to hearings, instead of weakening them.

(3) The third right, the right to a jury trial, is the one that undermines defendants' rights to hearings in the world of guilty pleas. In giving the right to contest enhancements at the first trial, the Court took away the right to contest enhancements at sentencing after pleading guilty.³⁵⁵ The real harm to defendants comes from depriving them of this hearing. The incentives to plead guilty are so powerful that the first trial right is merely theoretical for vast classes of defendants. Going to first trials on any issues, including enhancements, would jeopardize these incentives while conferring at best speculative benefits at sentencing. The acquitted conduct, if proved by a preponderance of the evidence, would still harm them at sentencing (unless they were up against statutory maxima). And defendants cannot rely on judges' knowledge of sentencing consequences to temper the outcomes of first trials. Because defendants have no unilateral right to waive juries and have bench trials, prosecutors can prevent judges from mitigating sentence enhancements. (Defendants want bench trials when they think judges would check overreaching prosecutors, but this is precisely when prosecutors do not want to let defendants have bench trials.)³⁵⁶ Judges thus have less power to check harsh prosecutors. Defendants are left with trial rights they cannot afford to exercise. In exchange for these trial rights, they have been robbed of colorable sentencing issues. Before the elements rule, they could have contested these issues at sentencing while pleading guilty. They must now surrender hearings on these issues with their guilty pleas, because they must plead guilty to every element of each offense.

While the elements rule costs defendants an option, it gives another option to prosecutors. For enhancements subject to the elements rule,

354. The same would be true if, after the elements rule, defendants had the rights (1) to waive jury trial and opt for a bench trial unilaterally, without the prosecution's consent, and (2) to do so only as to certain elements (such as enhancements) while pleading guilty to the rest, and (3) to do so without giving up the benefits of pleading guilty to the whole indictment, and (4) to avoid any risk that the same conduct tried before the judge would then be proved as relevant conduct at sentencing. These changes would effectively neutralize the overwhelming pressure to plead guilty to all of the elements charged in the indictment. I am indebted to George Fisher for this point.

355. Again, this is because a defendant cannot plead guilty without allocating to every element of the offense. Now that enhancements are elements, defendants must allocate to them when they plead guilty and are then estopped from relitigating them at sentencing.

356. In other words, if going to a jury trial is a zero-sum game, whenever it is advantageous for a defendant to do so it will be disadvantageous for the prosecution and the prosecution will veto the jury trial. (In fact, because jury trials are more cumbersome than bench trials, going to a jury trial is a negative-sum game, which only exacerbates the prosecution's incentive to veto jury trials.) Another way of looking at the problem is that prosecutors may want to hedge their bets by trying to prove enhancements to two different fact-finders. If the jury dislikes an enhancement, the prosecution can still try to prove it to the judge. Bench trials involve putting all the prosecutor's eggs in the judge's basket, so that bench trials are more attractive to the defendant and less attractive to the prosecutor.

prosecutors now have two bites at the apple. They can try to prove the enhancement once at the first trial (if there is one), and again as relevant conduct at sentencing. At sentencing, the prosecution need prove the enhancement by only a preponderance of evidence. This loophole swallows up the benefits of proof beyond a reasonable doubt at the first trial. It also puts added pressure on defendants not to go to trial at all. By giving prosecutors another option, while taking away defendants' rights to try enhancements at sentencing after pleading guilty, the rule undermines defendants' rights to be heard.

One might point out that what undermines hearings is not the elements rule per se, but how it interacts with other rules. These other rules include relevant conduct, prosecutors' power to veto bench trials, and the massive benefits of pleading guilty to every single element of the offense. There is some truth to this objection. In theory, one could solve these problems either by eliminating the elements rule or by eliminating these other rules (though doing the latter would be politically impractical).³⁵⁷ The Court, however, has shown no inclination to change these other rules. It has, for example, repeatedly affirmed the use of relevant conduct to enhance sentences, even where that includes acquitted conduct.³⁵⁸ New rules must fit the legal landscape of the old, but the elements rule does not. When combined with the massive benefits for pleading guilty, the elements rule deprives defendants of sentencing hearings. It does so even more when one adds relevant conduct and vetoes on bench trials to the mix.

One might also object that my analysis works only for defendants who plead guilty without plea agreements. Most enhancement issues are instead wrapped up in package deals in plea bargains and so would not be contested at sentencing.³⁵⁹ While most defendants do enter plea bargains, the terms of those plea bargains are shaped by what the outcomes would be without plea bargains. It is often said that prosecutors and defendants plea-bargain in the shadow of expected outcomes at trial. More precisely, they plea-bargain in the shadow of expected outcomes should the plea-bargaining process break down. These alternative outcomes include not only the outcomes of jury

357. It is probably politically impractical to eliminate the massive benefits of pleading guilty, as this would result in many more trials and so cost much more money. So long as the massive benefits of pleading guilty remain in place, the elements rule will cause at least some of the perverse results described in the text.

358. See, e.g., *United States v. Watts*, 519 U.S. 148, 152-57 (1997) (per curiam); *Witte v. United States*, 515 U.S. 389, 397-98, 402-04 (1995).

359. I have been unable to find any statistics on the percentage of defendants who plead guilty with and without plea agreements. When I was a federal prosecutor, I was not aware of any reporting of these statistics. In my experience, a substantial majority of defendants pleaded guilty with plea agreements, but a significant minority (perhaps one-quarter) pleaded guilty without agreements. Moreover, the terms of the plea agreements were shaped by the knowledge of what terms a defendant could get by pleading guilty without an agreement. Thus, the plea-without-agreement number served as a starting point for further plea bargaining.

trials, but also the outcomes of guilty pleas without plea agreements. Because plea bargains are struck in the shadow of these outcomes, plea terms will reflect the costs of these rights. To the extent that the reasonable-doubt standard strengthens the hands of those few defendants who can credibly threaten to go to trial, prosecutors will offer them more generous terms. Conversely, to the extent that the jury-trial right undermines sentencing hearings after guilty pleas without agreements, prosecutors will offer less favorable terms. For many defendants, the harm of losing a hearing option will greatly outweigh the benefit.³⁶⁰

Prosecutors need not burden themselves by going to trial more often in order to reap this windfall. Now that the elements rule has strengthened their bargaining position, prosecutors can drive harder plea bargains. They can calibrate the added harshness of those bargains to produce the same number of guilty pleas as before, secure in the knowledge that most defendants will plead to avoid worse outcomes at trial.

c. An Illustration: The Example of Al. Take, for example, a typical federal drug trafficking defendant with one prior drug felony conviction. You can call him Al. Suppose that Al was caught with two kilograms of cocaine on his person, but alleged accomplices or coconspirators Betty and Chris had another forty kilograms.³⁶¹ Suppose also that the prosecution

360. I am setting aside for now the intriguing likelihood that the barriers to going to trial are so enormous, and the threat of going to trial so incredible, that in some categories of cases plea-bargaining outcomes may become decoupled from expected trial outcomes. There is, however, good reason to question the conventional wisdom about bargaining in the shadow of expected trial outcomes. For example, in securities class actions, agency costs and other barriers to trial cause these actions to settle for roughly one-quarter of the damages demanded, regardless of their merits. In other words, because the parties know that they will never go to trial, they settle cases independently of expected trial outcomes. See Janet Cooper Alexander, *Do the Merits Matter? A Study of Settlements in Securities Class Actions*, 43 STAN. L. REV. 497 (1991). I am indebted to Jon Molot for sharpening my thinking on this point.

Here is another way of looking at the same phenomenon: The conventional view assumes that the only possible outcomes are a jury verdict or a plea agreement. If the expected value of a jury verdict, discounted by the possibility of acquittal, is nine years, then, to induce a plea, any plea bargain offer must track this number and offer nine years or less. (I am setting aside for present purposes the ways in which risk aversion or risk preference may adjust this number up or down.) But in the real world, there are *three* numbers in play: the outcome after trial, the outcome under a plea agreement, and the outcome after a guilty plea without an agreement. If the expected value of a jury verdict, discounted by the possibility of acquittal, is twenty years, and the expected value of a guilty plea without a plea agreement is ten years, then plea bargaining happens not in the shadow of the twenty-year figure, but rather in the shadow of the ten-year figure. Moving the expected value of a trial up to twenty-two years or down to eighteen years by fiddling with a trial right (such as the reasonable-doubt provision of the elements rule) does not affect plea bargaining. This is true because everyone knows the defendant will never go to trial; at worst, he will simply plead guilty without an agreement and get at most ten years. The only subject of negotiation is whether the defendant will enter an early guilty plea (sparing both sides a great deal of work) and surrender a possible sentencing hearing in exchange for some additional benefit in a plea agreement.

361. Agents commonly catch drug couriers red-handed with modest quantities, while also finding circumstantial evidence linking them to other people with much larger quantities. Another common scenario involves a telephone conversation intercepted by a wiretap, in which the

refused to cut a favorable plea bargain with Al and wished to hold him jointly responsible for the entire forty-two kilograms. Under the old law, Al pleaded guilty to drug trafficking without a plea agreement³⁶² and without allocuting to a specific quantity of cocaine. He then disputed the drug quantity at sentencing, claiming the others were not his accomplices or coconspirators or that he could not reasonably have foreseen the additional forty kilograms. By pleading guilty, Al gained at least three benefits: First, he got a three-point sentencing reduction under the Guidelines for acceptance of responsibility.³⁶³ Second, most judges would have given Al a sentence at the bottom of the range because he spared the court a trial. Third, the prosecution would not have sought to enhance Al's sentence based on his recidivism. Prosecutors reserve such enhancements for defendants who force them to go to trial.³⁶⁴ If the prosecution did not prove that Al was responsible for the forty-two kilograms, his sentence would likely have been five years and three months.³⁶⁵ If it proved his responsibility, his sentence would likely have been ten years and one month.³⁶⁶

defendant is caught talking with another person about drugs. In this scenario, the precise quantity of drugs might be unclear. In addition, the defendant's statements might or might not be interpreted as aiding or abetting or as part of a conspiracy.

362. Alternatively, the prosecution and Al might have entered into a plea bargain that left open the issue of drug quantity, agreeing to disagree at the sentencing hearing.

363. U.S. SENTENCING GUIDELINES MANUAL § 3E1.1 (1998) (subtracting two points for any guilty plea and a third point for a timely plea where the adjusted offense level is sixteen or higher).

364. *Supra* note 342.

365. Assume that Al had been sentenced to at least sixty days on his prior drug conviction but had no other qualifying convictions. This would put him in Criminal History Category II. *See* U.S. SENTENCING GUIDELINES MANUAL § 4A1.1(a) (1998) (specifying that the criminal history described above results in two or three criminal history points); *id.* at ch. 5 pt. A (sentencing table) (showing that two or three criminal history points places a defendant in Criminal History Category II). A drug quantity of two kilograms of cocaine results in a base offense level of twenty-eight, which, for a defendant in Criminal History Category II, means a base sentencing range of 87 to 108 months. *See id.* § 2D1.1(c)(6) (setting a base offense level of twenty-eight for two to three and one-half kilograms of cocaine); *id.* at ch. 5 pt. A (sentencing table) (showing that the intersection of the base offense level row and the criminal history category column results in a sentencing range of 87 to 108 months). Al's guilty plea would ordinarily entitle him to a three-level reduction for acceptance of responsibility, lowering his adjusted offense level from twenty-eight to twenty-five. Thus, his sentencing range would have been sixty-three to seventy-eight months. *See id.* § 3E1.1 (adjusting the offense level downwards by three levels for acceptance of responsibility); *id.* at ch. 5 pt. A (sentencing table) (showing that the intersection of the adjusted offense level row and the criminal history category column results in a sentencing range of sixty-three to seventy-eight months). The mandatory minimum would have been five years, which is sixty months. 21 U.S.C. § 841(b)(1)(B) (1994). Absent any unusual aggravating facts, the judge likely would have sentenced Al to five years and three months, the bottom of the range, as an implicit reward for pleading guilty.

366. This analysis again assumes Al had been sentenced to at least sixty days on his prior drug conviction and had no other convictions. So once again, his Criminal History Category would be II. *See* U.S. SENTENCING GUIDELINES MANUAL § 4A1.1(a)(b) (1998) (specifying that the criminal history described above results in two or three criminal history points); *id.* at ch. 5 pt. A (sentencing table) (showing that two or three criminal history points puts a defendant in

Now let us see how the same scenario plays out under the elements rule. Trafficking in half a kilogram to five kilograms of cocaine carries a statutory minimum sentence of five years and a maximum of forty years.³⁶⁷ Trafficking in five or more kilograms of cocaine carries a statutory minimum sentence of ten years and a maximum of life.³⁶⁸ In other words, a drug quantity of five kilograms of cocaine or more increases Al's maximum penalty from forty years to life. Thus, the five-kilogram quantity is now an element of the offense. Al has an unenviable set of choices. If he pleads guilty, he must allocute to at least five kilograms, which triggers a ten-year mandatory minimum.³⁶⁹ This minimum is almost five years above what he

Criminal History Category II). The base offense level for forty-two kilograms of cocaine is thirty-four, which standing alone would mean a sentencing range of 168 to 210 months. *See id.* § 2D1.1(c)(3) (setting a base offense level of thirty-four for fifteen to fifty kilograms of cocaine); *id.* at ch. 5 pt. A (sentencing table) (showing that the intersection of the base offense level row and the criminal history category column results in a sentencing range of 168 to 210 months). The three-level reduction for acceptance of responsibility would reduce the adjusted offense level from thirty-four to thirty-one. *See id.* § 3E1.1 (adjusting the offense level downwards by three levels for acceptance of responsibility). On this analysis, his sentencing range would have been 121 to 151 months. *See id.* at ch. 5 pt. A (sentencing table) (showing that the intersection of the adjusted offense level row and the criminal history category column results in a sentencing range of 121 to 151 months). The mandatory minimum would have been ten years. 21 U.S.C. § 841(b)(1)(A). The judge likely would have sentenced Al to the bottom of the range, ten years and one month, absent any unusual aggravating facts, as an implicit reward for sparing the court a trial. The analysis also assumes that Al would not have perjured himself at the sentencing hearing, which would have opened him to an obstruction-of-justice enhancement.

367. 21 U.S.C. § 841(b)(1)(B).

368. *Id.* § 841(b)(1)(A).

369. *Id.* Note that Al might not be able to allocute honestly to only five kilograms. He might have to allocute to more to get over the five-kilogram threshold. For example, if the forty kilograms was a single shipment, Al might have to admit to knowing about the whole shipment unless he can somehow claim that he thought his accomplices or coconspirators were going to have a smaller quantity than they did.

This result is not a peculiarity of mandatory minima, though minima do exacerbate the phenomenon. The crux of the problem is that the elements rule requires defendants to allocute to more than they previously had to if they want to plead guilty. For example, imagine that the drug laws contained no mandatory minima. The elements rule would nonetheless require Al to allocute to at least five kilograms, because five kilograms is now an element of the enhanced offense. In the scenario in the preceding paragraph, he might have to allocute to the entire forty-two kilograms. If Al's Criminal History Category was II and he could somehow allocute to only five kilograms, his base offense level would be thirty-two, which would mean a range of 135 to 168 months after a trial. *See U.S. SENTENCING GUIDELINES MANUAL* § 2D1.1(c)(4) (1998) (setting a base offense level of thirty-two for trafficking in between five and fifteen kilograms of cocaine); *id.* at ch. 5 pt. A (sentencing table) (providing that a Criminal History Category of II and an offense level of thirty-two results in a sentencing range of 135 to 168 months). If one subtracts three points for acceptance of responsibility, Al's adjusted offense level after pleading guilty to five kilograms would be twenty-nine, resulting in a sentencing range of 97 to 121 months. *See id.* § 3E1.1 (providing for a three-level reduction for acceptance of responsibility); *id.* at ch. 5 pt. A (sentencing table) (providing that a Criminal History Category of II and an adjusted offense level of twenty-nine results in a sentencing range of 97 to 121 months). So Al's minimum sentence for a five-kilogram plea would be eight years and one month.

If Al had to allocute to the entire forty-two kilograms, his base offense level would be thirty-four, which would mean a sentencing range of 168 to 210 months after trial. *See id.* § 2D1.1(c)(3) (setting a base offense level of thirty-four for trafficking in between fifteen and fifty kilograms of cocaine); *id.* at ch. 5 pt. A (sentencing table) (providing that a Criminal History Category of II and

would have gotten if he had won the sentencing hearing under the old law. It is only one month below what he would have gotten if he had lost the hearing. The elements rule has cost Al the opportunity to contest almost five years of his sentence.

Al cannot get around this ten-year minimum by going to trial. If Al goes to trial, the prosecution will file a prior felony information, reciting Al's prior conviction. Whether to file such an information is entirely up to prosecutors,³⁷⁰ and they routinely do so against defendants who go to trial. The prior felony information raises Al's mandatory minimum to ten years if he is convicted of having less than five kilograms.³⁷¹ If he had five kilograms or more, it raises his mandatory minimum to twenty years.³⁷² In short, going to trial doubles the penalties Al faces, making trial suicidal for Al.³⁷³

The situation is not much better for Al if he does not have a prior drug felony on his rap sheet. If he takes the witness stand, he risks an obstruction-of-justice enhancement for perjury.³⁷⁴ Assume that he does not testify. Even so, he loses credit for acceptance of responsibility because he chose to go to trial.³⁷⁵ He also antagonizes the judge, who will probably

an offense level of thirty-four results in a range of 168 to 210 months). If Al pleaded guilty and earned a three-level reduction for acceptance of responsibility, his adjusted offense level would be thirty-one, resulting in a sentencing range of 121 to 151 months. *See id.* § 3E1.1 (providing for a three-level reduction for acceptance of responsibility); *id.* at ch. 5 pt. A (sentencing table) (providing that a Criminal History Category of II and an adjusted offense level of thirty-one results in a sentencing range of 121 to 151 months). Thus, Al's likely sentence would be ten years and one month. Either way, the elements rule raises his best-case outcome far above the five years and three months he could have gotten before the elements rule.

370. *See* 21 U.S.C. § 851(a)(1).

371. *Id.* § 841(b)(1)(B).

372. *Id.* § 841(b)(1)(A).

373. Al might try to circumvent this problem by pleading guilty to the base offense while going to a first trial on the enhanced offense. If Al is lucky, the prosecution might choose not to file a prior felony information on the base offense, but only on the enhanced offense. (This might depend on whether the prosecution found the press of enhancement trials burdensome.) Thus, if Al won the first trial, he would not face the stiff recidivist penalties. His best-case scenario would be almost as good as it would have been before the elements rule: somewhere between five years and three months and six years and six months, depending on whether the judge sentenced him at the top of the range as a penalty for going to trial. *See* U.S. SENTENCING GUIDELINES MANUAL § 2D1.1(c)(6) (1998) (setting a base offense level of twenty-eight for two kilograms); *id.* § 3E1.1 (providing for a three-level reduction for acceptance of responsibility); *id.* at ch. 5 pt. A (sentencing table) (providing that a Criminal History Category of II and a base offense level of twenty-eight would result in a sentencing range of 87 to 108 months; also providing that if the offense level were adjusted down three levels to twenty-five for acceptance of responsibility, the sentencing range would be sixty-three to seventy-eight months). But if he lost, he would still face a minimum of twenty years' imprisonment (because of the statutory mandatory minimum). By making Al's worst-case outcome much worse, while not improving his best-case outcome at all (and perhaps worsening it somewhat), the elements rule hurts Al and deters him from seeking a hearing.

374. *Id.* § 3C1.1 & application note 4(b) (1998) (adding two points for obstruction of justice, including perjury).

375. *See id.* § 3E1.1 cmt. nn.2-3 (stating that the reduction ordinarily applies to defendants who plead guilty but does not apply to defendants who deny elements of factual guilt at trial).

sentence him closer to the top of the sentencing range. The prosecution may even seek a school-zone sentence enhancement, which it probably would have forgone otherwise.³⁷⁶

If the prosecution proves only that Al is responsible for two kilograms, his sentence will likely be near nine years (without the school enhancement) or eleven years and three months (with it). If the prosecution proves the entire forty-two kilograms, Al's sentence will likely be near seventeen and a half years (without the school enhancement) or twenty-one years and ten months (with it).³⁷⁷ If the prosecution proves only two kilograms beyond a reasonable doubt, but forty-two kilograms by a preponderance of the evidence, the jury will convict of the lesser amount. Nonetheless, the judge will hold the whole forty-two kilograms against Al at sentencing. The judge will give him the same sentence as if the jury had convicted him of the entire amount, namely, almost seventeen and a half years or twenty-one years and ten months. Once again, going to trial almost doubles the penalties Al faces.

Al might try to get around this problem by pleading guilty to the two kilograms and going to trial only to dispute the additional forty kilograms.³⁷⁸ While the Guidelines are not clear on this point, Al might be able to get credit for acceptance of responsibility on the base offense, though not if he is convicted of the enhanced offense. If the prosecution does not prove the additional kilograms by a preponderance of the evidence, Al's sentence will likely be worse than it would have been before the elements rule: between five years and three months and eight years and one month, depending on the school-zone enhancement and whether the judge penalizes Al for going to trial.³⁷⁹ If the prosecution proves the entire

376. *Supra* note 344.

377. This analysis assumes that Al has two or three criminal history points, which is the equivalent of one serious prior conviction or two or three minor ones and puts him in Criminal History Category II. On this assumption, a forty-two kilogram conviction at trial would carry a base offense level of thirty-four and a sentencing range of 168 to 210 months without the protected-location enhancement, or an adjusted offense level of thirty-six and a sentencing range of 210 to 262 months with the enhancement. On the same assumption, a two-kilogram conviction at trial would carry a base offense level of twenty-eight and a sentencing range of 87 to 108 months without the protected location enhancement, or an adjusted offense level of thirty and a sentencing range of 108 to 135 months with the enhancement. *See* U.S. SENTENCING GUIDELINES MANUAL § 2D1.1(c)(3) (1998) (setting a base offense level of thirty-four for between fifteen and fifty kilograms); *id.* § 2D1.1(c)(6) (setting a base offense level of twenty-eight for between two and three and one-half kilograms); *id.* § 2D1.2(a)(1) (prescribing a two-level enhancement for drug trafficking in a protected location); *id.* § 4A1.1 (setting forth Criminal History Category rules); *id.* at ch. 5 pt. A (sentencing table).

378. For present purposes, assume that Al manages to allocute to the two kilograms without admitting or denying the forty kilograms, thereby avoiding an obstruction-of-justice enhancement.

379. Al's sentencing range would be sixty-three to seventy-eight months without the school-zone enhancement, or seventy-eight to ninety-seven with it. As suggested earlier, the judge would likely sentence Al toward the top of the range to penalize his going to trial, unless the judge saw how the elements rule harms defendants. *See supra* notes 341, 346. This analysis once again assumes that Al has two or three criminal history points, which is the equivalent of one serious

forty-two kilograms by at least a preponderance of the evidence, then the analysis in the preceding paragraph applies. Thus, Al will face almost seventeen and a half years or twenty-one years and ten months in prison.

In other words, all of Al's options are worse than they were before the elements rule. A straight guilty plea now gets him a minimum of ten years, which is almost double what he could have gotten before. If he pleads guilty, he no longer gets a meaningful hearing at sentencing. Al can get around this ten-year sentence only by rolling the dice and going to trial. His best-case outcome at trial is as bad as or worse than the five years and three months he could have gotten by pleading guilty before the elements rule. And his worst-case outcome is much worse, roughly double the worst-case outcome (ten years and one month) he faced before the elements rule. If Al wants to avoid the disastrous worst-case outcome, he must plead guilty. The elements rule has cost him his right to contest drug quantity at sentencing after pleading guilty. The prosecution knows that Al's options without a plea bargain are now much worse than they were before the elements rule. Thus, if it bargains at all, it can drive a harder bargain in the shadow of these worse outcomes, secure in the knowledge that Al must jump at the chance to plead.

d. The Broader Categories of Defendants Harmed and Helped by the Elements Rule. Al is just one example of the broader categories of defendants whose hearing rights are harmed by the elements rule. The rule harms many defendants by forcing them to choose between the benefits of pleading guilty and a hearing on enhancements. They used to be able to enjoy both, but must now jeopardize the former if they want the latter. Many will choose the bird in the hand (the certainty of massive plea benefits) over the bird in the bush (a possible win on enhancements). The result is fewer hearings and less judicial oversight of prosecutorial charging decisions, so prosecutors can be more aggressive in charging enhancements.

It is true that some defendants will benefit from the elements rule. But, for two reasons, the number of beneficiaries is likely to be small. First, a defendant cannot benefit unless she is up against a statutory maximum, for otherwise the relevant-conduct rule undercuts any benefits from the reasonable-doubt standard. Second, a defendant cannot benefit from this new trial right unless it is beneficial or at least plausible for her to go to

prior conviction or two or three minor ones and puts Al in Criminal History Category II. See U.S. SENTENCING GUIDELINES MANUAL §§ 2D1.1(c)(3), (6), 2D1.2(a)(1), 4A1.1 (1998); *id.* at ch. 5 pt. A (sentencing table) (prescribing a base offense level of twenty-eight and a sentencing range of 87 to 108 months for a two-kilogram conviction after trial, minus three levels for acceptance of responsibility, which would result in an adjusted offense level of twenty-five and a sentencing range of sixty-three to seventy-eight months after a guilty plea to two kilograms; a two-level enhancement for committing the offense in a protected location would adjust the offense level up to twenty-seven and the sentencing range upwards to seventy-eight to ninety-seven months).

trial. If the elements rule makes the threat of trial plausible, she may extract a better plea bargain. If not, the elements rule is at best useless and at worst harms sentencing hearings.³⁸⁰ The largest category of beneficiaries comprises those who would have gone to trial or were on the cusp of going to trial even without the elements rule.³⁸¹ Currently, only nine percent of all defendants go to bench or jury trials. Even if the elements rule made it plausible for another nine percent to go to trial or threaten to do so, the overwhelming majority (eighty-two percent) would not benefit. Indeed, many members of the eighteen-percent minority would not benefit either: Many of them do not face sentence enhancements. And many others are not up against statutory maxima and so derive no benefits from the reasonable-doubt standard at trial, because it is undercut by the preponderance standard at sentencing. Moreover, those most likely to benefit are the least sympathetic. Hardened recidivists are more likely to benefit from the reasonable-doubt right than are first-time offenders, because long criminal histories inflate sentencing-guidelines calculations up to statutory maxima.³⁸² Prosecutors will also offer deeper discounts to the worst offenders facing the most complex enhancements, in order to avoid cumbersome bifurcated trials.³⁸³ Finally, any benefits accruing to the minority who go to trial are far from certain. If the elements rule proves to be cumbersome, legislators, judges, and prosecutors might increase pressures to plead guilty, thus reducing the number of trials.³⁸⁴

380. See *supra* note 360 (suggesting that, where the threat of going to trial is not plausible, plea-bargaining outcomes may not track trial outcomes).

381. Two classes of defendants are likely to benefit from the elements rule. First, for some defendants the size of enhancements will be so massive as to dwarf the benefits from pleading guilty, *and* the proof of the enhancements will be more than a preponderance of the evidence but less than proof beyond a reasonable doubt, *and* statutory maxima would prevent the prosecution from using much of the acquitted conduct as relevant conduct at sentencing. This conjunction of three independent circumstances seems likely to occur in only a very small percentage of cases. Even these defendants, however, will gain only a partial benefit. If the jury convicts on the base offense but acquits on the enhancement, the judge will still use the acquitted conduct to increase the defendant's sentence up to the base offense's statutory maximum. If Congress, in response to the elements rule, raises statutory maxima, even this limited benefit will evaporate. Prosecutors can often do the same thing by charging multiple counts and using consecutive sentences to raise the overall statutory maximum.

Second, the rule will also benefit those defendants against whom the evidence of guilt is so weak that they would have gone to trial or would have been very close to doing so even before the elements rule, *and* whose likely sentences if enhanced would exceed the statutory maxima for the unenhanced crimes. The benefit on sentence enhancements may tip the balance of the undecided in favor of going to trial. But given that only nine percent of all defendants currently go to trial, the number of defendants in this class is likely to be much smaller than the numbers harmed. See *supra* note 330. And this greater number of trials may be cancelled out by the use of more pressure to make defendants plead guilty. See *infra* note 384.

382. This is not true, however, where recidivism enhancements raise statutory maxima significantly, as the recidivism enhancements for repeat drug dealers do. See *supra* note 342 and accompanying text (discussing prior felony informations).

383. See *supra* text accompanying note 303.

384. In the real world, where prosecutors hold huge leverage to induce guilty pleas and ninety-one percent in fact plead guilty, one might speculate that the elements rule will actually

Without empirical work, one cannot quantify how many defendants will lose sentencing issues and hearings under the elements rule. But there is reason to think it is a very large number. Drug defendants make up the largest class of federal criminals, amounting to more than a third of federal defendants and more than half of all federal inmates.³⁸⁵ Most repeat drug dealers, and many first-time ones, will face Al's predicament under federal law or similar state statutes. And while most states do not use the federal system's broad approach to relevant conduct,³⁸⁶ the sizable incentives to plead guilty are common to most systems. Half of all states either have sentencing guidelines in place or are considering enacting them.³⁸⁷ Even in indeterminate-sentencing states, guilty pleas earn similar quid pro quos, albeit less overt ones.³⁸⁸ Defendants must jeopardize these valuable benefits if they wish to exercise their theoretical trial rights. For many, the only option will be to plead guilty. The law now forces them to surrender sentencing hearings and issues with their guilty pleas. The nominal right to be heard at trial in practice undercuts the due process right to be heard at sentencing.

decrease the number of jury trials. The rule might do so by adding one or more additional guilt phases between the main trial on guilt and sentencing. *See supra* notes 301-302 and accompanying text (discussing bifurcation and how it could lead to dozens of separate hearings in complex multidefendant cases). This development would make trials much more cumbersome. The more cumbersome the trial, the less likely the legislature, judge, and prosecutor will want to go through with it, and the more pressure they will bring to bear on the defendant to plead guilty. One can only speculate that this will happen, but there are interesting historical parallels. As John Langbein has noted, the high evidentiary requirement of confessions in the Middle Ages led to torture to overcome the hurdle and extract the needed confessions. Just so, the increase in procedural rights and cumbersomeness of trials led to the pressures to plead guilty that have made trials an endangered species. *See* John H. Langbein, *Torture and Plea Bargaining*, 46 U. CHI. L. REV. 3 (1978). The latter could easily happen again, only more so.

385. Of the 75,071 criminal defendants whose cases were disposed of in federal district court during the twelve months ending September 30, 2000, 27,220 (36.2%) were charged with drug offenses. *See* 2000 FEDERAL STATISTICAL TABLES, *supra* note 330, tbl.D-4. The actual number is probably higher, as this figure excludes defendants whose major offense was not a drug offense but who had one or more drug charges as well. Moreover, drug sentences are longer than those for many other types of crimes. Thus, the percentage of federal inmates doing time for drug offenses is even higher—well over half of all federal inmates. *See* King & Klein, *supra* note 313, at 340 n.7 (noting that in 1998, 58.9% of federal inmates were incarcerated for drug-related offenses and that nearly a quarter of all federal and state inmates are serving sentences or awaiting trial for drug offenses).

386. *See* MICHAEL TONRY, SENTENCING MATTERS 68 (1996).

387. Seventeen states currently use guidelines, and another eight are considering enacting them. *See* Richard S. Frase, *Is Guided Discretion Sufficient? Overview of State Sentencing Guidelines*, 44 ST. LOUIS U. L.J. 425, 427, 446 tbl.1 (2000).

388. *See* U.S. SENTENCING COMM'N, SUPPLEMENTARY REPORT ON THE INITIAL SENTENCING GUIDELINES AND POLICY STATEMENTS 48 (1987) (revealing, in an empirical survey of pre-Guidelines sentencing practice, that guilty pleas earned reductions of between thirty and forty percent); Stephen Breyer, *The Federal Sentencing Guidelines and the Key Compromises upon Which They Rest*, 17 HOFSTRA L. REV. 1, 28 (1988) (same).

2. *Checking Arbitrariness*

The Due Process and Equal Protection Clauses check arbitrariness and ensure equal, predictable treatment. Yet in practice, the elements rule will lead to more arbitrariness, not less. In the real world of guilty pleas, it gives prosecutors more power to manipulate plea bargaining and makes it harder for judges to check prosecutors. This arbitrariness will benefit some lucky individual defendants, at the expense of systemic equal treatment and at the risk of possible discrimination.

One might ask why we should worry about lighter sentences for some defendants, particularly if one thinks sentences are too harsh. If some defendants are better off, no defendants are worse off, and the prosecution wants to deal, who is to complain? This would be an appropriate way to view a contractual exchange between private parties, as a consensual deal makes both sides better off. But a plea bargain is not an ordinary contract. For example, prosecutors have incentives to ease their workloads by striking more lenient deals with aggressive defense lawyers. This prosecutorial self-interest in favoring strong defense lawyers, however, is at odds with society's interest in equal treatment. Prosecutors may also consciously or unconsciously favor defense lawyers with whom they are friendly. Society has a strong interest in giving prosecutors less room to favor clients of strong or well-connected defense lawyers over other defendants.³⁸⁹ Likewise, society has a strong interest in making it harder for biased prosecutors to favor white defendants over black ones, for example. Even if individual prosecutors treat like cases alike, different prosecutors and prosecutors' offices will behave differently, creating disparities. The legislature has an interest in equalizing sentences, so that a defendant's punishment does not depend on which prosecutor or which office gets the case.

Prosecutors have power to select what offenses to charge. Prosecutors can use this power arbitrarily to favor certain defendants over others who have committed the exact same crime. This maneuver is known as charge bargaining, in which prosecutors allow some defendants to plead guilty to less serious charges and drop the remaining charges. Charge bargaining

389. Admittedly, stronger defense lawyers might have a better chance of getting their clients acquitted, which lowers the expected value of their clients' sentences on going to trial and so would be expected to have a ripple effect on the plea bargains offered to their clients. One can debate whether the efficiency of offering more generous plea terms to dispose of these cases should trump the interest in equal sentences for equally culpable offenders. But even if one sets this effect aside, prosecutors still have an incentive to strike more lenient deals with strong defense lawyers. Doing so allows prosecutors to avoid difficult and acrimonious negotiations and pretrial preparation, even if they would eventually have won at trial. To this extent, plea-bargaining outcomes can fail to reflect likely outcomes at trial. *Cf.* Alexander, *supra* note 360 (noting an analogous disconnect between the merits of securities class actions and settlement figures because of agency costs and other factors).

threatens to undermine equal treatment, at least when it involves forgoing readily provable charges. The Guidelines check this power by adopting a modified real-offense system.³⁹⁰ Take an example based on *Jones*: The prosecutor and defense lawyer recite in a plea bargain that the defendant carjacked a car, without mentioning injury. The Probation Department, however, conducts an independent investigation and may well find that the victim suffered serious bodily injury or death. Under the relevant conduct guideline, the court makes independent factual findings.³⁹¹ If it finds that in fact the victim suffered serious bodily injury or death, it sentences the defendant accordingly. The only limits on the court's power to pierce charge bargains are the statutory minima and maxima for the offenses charged.³⁹² Because the statutory maximum for carjacking was life imprisonment under the old law, however, the parties could not use the maximum to tie the judge's hands.³⁹³ The knowledge that the court had the power to pierce charge bargains limited parties' willingness to conceal or downplay facts. In other words, judges checked prosecutors.

Even under the old law, prosecutors and defense lawyers could charge bargain to manipulate some statutory minima and maxima.³⁹⁴ The elements rule has poured more fuel on this fire by fragmenting crimes, creating many more maxima.³⁹⁵ In the example above, the prosecutor now charges the defendant with simple carjacking. Because the elements rule has fragmented the unified carjacking statute into three separate offenses, the defendant's statutory maximum is now fifteen years. It does not matter how much injury or death the defendant caused. The prosecutor has now tied the judge's hands by choosing to cite the weakest statutory subsection. If the prosecutor refuses to strike a deal with another, identical defendant, that

390. See U.S. SENTENCING GUIDELINES MANUAL ch. 1, pt. A, at 5-6 (1998); cf. William L. Gardner & David S. Rifkind, *A Basic Guide to Plea Bargaining Under the Federal Sentencing Guidelines*, CRIM. JUST., Summer 1992, at 14, 16-17 (discussing the tradeoffs of pre- and post-indictment bargaining under the Guidelines); O'Sullivan, *supra* note 248, at 1348.

391. See U.S. SENTENCING GUIDELINES MANUAL § 1B1.3 (1998).

392. See U.S. SENTENCING GUIDELINES MANUAL § 5G1.1(a)-(b) (1998); *Edwards v. United States*, 523 U.S. 511, 514-15 (1998). Note that the parties are still able to bargain over charges when the bargains reflect the lack of proof of those charges. If the prosecution does not have enough proof of a charge and drops it as part of a plea bargain, the court will not count the charge as relevant conduct because of the weakness of proof. The elements rule does not change this.

393. 18 U.S.C. § 2119(3) (1994).

394. See O'Sullivan, *supra* note 248, at 1355; Daniel J. Sears, *Federal Guidelines Sentencing—The Good, the Bad, and the Ugly*, COLO. LAW., Feb. 1997, at 3, 4; Daniel J. Sears, *Practice Under the Federal Sentencing Guidelines: Bargaining for Freedom*, 22 COLO. LAW. 485, 486-87 (1993).

Note that charge bargaining takes place not only after indictment, but even before prosecutors decide what charges to file. See *United States v. Moody*, 206 F.3d 609, 616-18 (6th Cir.) (Wiseman, J., concurring) (recognizing that precharge bargaining is becoming increasingly common, but refusing to treat failure to bargain precharge as a ground for a claim of ineffective assistance of counsel), *cert. denied*, 121 S. Ct. 301 (2000).

395. See Ross, *supra* note 3, at 200 (predicting such a consequence of an elements rule before *Apprendi* was handed down).

defendant receives much harsher punishment. What matters is no longer the defendant's real offense, but the statutory maximum of the statute cited in the indictment. It is up to the prosecutor to decide which statute to cite.³⁹⁶ These statutory maxima bind the court, no matter how arbitrary or disparate the results. Thus, judges now have less power to check prosecutors.

The point of this Article is not to defend all of real-offense sentencing, let alone the particular version adopted by the Guidelines. There is already an extensive literature on this point,³⁹⁷ and wading into that marsh is beyond the scope of this Article. My point is much more modest: Congress, in creating unified crimes such as the carjacking statute, made a reasonable effort to create a coherent penalty structure. By checking charge bargaining, these measures limited arbitrariness and disparities. Arbitrariness, disparity, and potential discrimination are antithetical to equal protection and due process. By fragmenting these statutes, the elements rule disserves these values.

3. *Conscience of the Community*

The elements rule also disserves another value: expressing "the conscience of the community."³⁹⁸ As Section II.B argued, legislatures will likely evade the elements rule by redefining aggravating factors as mitigating ones. To the extent that they do so, the elements rule will be of no effect. To the extent that they do not, they will likely delegate even more power to unelected sentencing commissions, which will stifle the conscience of the community.

The Jury Clauses protected the jury as a populist check on arbitrary judges. The Framers wanted juries to ensure a democratic, representative element in the judiciary. The jury was "a political institution,"³⁹⁹ "*the democratic branch of the judiciary power.*"⁴⁰⁰ As Akhil Amar put it, "the

396. See Frank O. Bowman, III, *Completing the Sentencing Revolution: Reconsidering Sentencing Procedure in the Guidelines Era*, 12 FED. SENTENCING REP. 187, 189 (2000) (pointing out that requiring enhancements to be charged in the indictment would give "the government . . . unreviewable discretion to expand or contract a defendant's sentencing exposure by altering the contents of the charging instrument. . . Prosecutors would be at liberty to manipulate sentencing outcomes . . . either unilaterally or in order to effectuate a plea agreement").

397. See, e.g., STITH & CABRANES, *supra* note 210, at ch. 4 (criticizing the Guidelines' real-offense system, noting that the procrustean Guidelines pose their own dangers of arbitrariness by ignoring relevant factors, and advocating greater judicial power to counterbalance prosecutorial power); O'Sullivan, *supra* note 248, at 1343 & n.2 (defending modified real-offense sentencing, but collecting citations of critical commentary to the contrary).

398. *Witherspoon v. Illinois*, 391 U.S. 510, 519 & n.15 (1968).

399. 1 ALEXIS DE TOCQUEVILLE, *DEMOCRACY IN AMERICA* 283 (Phillips Bradley ed., Alfred A. Knopf 1945) (1835).

400. *Essays by a Farmer* (pt. 4), MD. GAZETTE, Mar. 21, 1788, reprinted in 5 THE COMPLETE ANTI-FEDERALIST 36, 38 (Herbert J. Storing with Murray Dry eds., 1981). The Anti-

overall jury system was fundamentally populist and majoritarian.”⁴⁰¹ Many Framers and colonists, Amar noted, analogized juries to mini-legislatures.⁴⁰²

In adopting the elements rule, the Court anachronistically assumed that the populist battle is still one of juries versus judges. But today, grand juries are rubber stamps.⁴⁰³ Hardly any petit juries are left, and even the few remaining juries have almost no role at sentencing. The remaining populist battleground is between elected legislatures and unelected sentencing-commission bureaucrats. It appears unlikely that a majority of the Court will extend the elements rule to invalidate sentencing guidelines, though Justices Scalia and Thomas would do so.⁴⁰⁴ Unless and until the Court goes that far, the elements rule will limit legislatively created enhancements but not identical ones created by unelected sentencing commissions.

Thus, under the elements rule, an unelected Sentencing Commission will somehow exercise power that Congress cannot. If Congress wanted to retain sentencing power, it would have to submit to the cumbersome procedures entailed by the elements rule, such as bifurcation.⁴⁰⁵ Instead,

Federalists fought for and won the Bill of Rights, including its jury-trial guarantees; thus, their views on the need for juries proved to be influential.

401. Amar, *supra* note 331, at 1185.

402. *See id.* at 1187-89.

403. As Chief Judge Sol Wachtler observed, any prosecutor can get a grand jury to indict a ham sandwich. Jim Felman, *An Interview with Sol Wachtler*, FED. LAW., May 1999, at 40, 45. In other words, grand juries do almost nothing to screen and regulate prosecutorial charging. This is not to deny that they can still serve other functions, such as investigation. *See generally* Renée B. Lettow, Note, *Reviving Federal Grand Jury Presentments*, 103 YALE L.J. 1333 (1994).

404. *See Apprendi v. New Jersey*, 120 S. Ct. 2348, 2380 n.11 (2000) (Thomas, J., concurring) (quoting Justice Scalia’s dissent in *Mistretta v. United States*, 488 U.S. 361, 413 (1989), which had argued that the Guidelines were unconstitutional); *supra* text accompanying notes 317-320. Even if the Court were to apply the elements rule to invalidate the Guidelines, this would simply substitute judges and probation and parole officers for sentencing commissioners. Legislatures would likely return to broad, open-ended delegations of power to judges. Judges could then find facts covertly within broad sentencing ranges, but legislatures could not require sentencing based on overt findings of the same facts. Judges would still find facts in their own heads but would not announce their findings. They would simply utter inscrutable sentences within unbounded ranges. Probation and parole officers would play similar roles long after sentencing, with even less scrutiny and fewer safeguards. Unaccountable judges, probation officers, and parole officers would exercise much more power in thousands of low-visibility decisions. These decisions would likely be unreviewable on appeal. While the elements rule purports to use populist juries to check arbitrary judging, arbitrariness would undercut transparency and democracy. (One could, in response, debate the countervailing benefits of allowing judges and probation and parole officers to mete out more individualized justice and the possibility of legislative guidance via factors and criteria rather than rigid enhancement rules. Even if these benefits were to materialize, they would come at the expense of democratic guidance.)

405. *See supra* text accompanying notes 301-302 (noting the large practical difficulties that the elements rule would cause if it led to bifurcation of trials). These logistical difficulties would attach to any solution in which Congress simply submitted to the elements rule. For example, if Congress were to set penalty *A* for simple assault, penalty *B* for assault with a deadly weapon, and penalty *C* for assault with a biased purpose, it would in essence be submitting to the elements rule’s cumbersome procedures and fragmentation of crimes. This would bring about bifurcation, more charge bargaining, and all of the other practical problems noted in Sections II.D and III.A.

Congress will broadly delegate power to the Sentencing Commission, but will not be able to instruct the Commission on setting maxima and enhancements. (Such instructions would run afoul of the elements rule.) Congress may perhaps signal vague displeasure with certain Guidelines, or delete the occasional enhancement. At most, it might exhort the Commission to consider certain factors without assigning any weight to them. But Congress dare not tell the Commission clearly to enact enhancements, lest its guidance amount to an enhancement law in violation of the elements rule. The rule in effect tells Congress to say nothing or at most to speak obliquely, in a way not transparent to voters. The political process will no longer check and critique sentence enhancements. The result may be more arbitrariness and less accountability.

In other areas of law, the Constitution disfavors unchecked delegations of power lest they obscure accountability. This is the nondelegation doctrine.⁴⁰⁶ “Congress [violates this rule where it] has declared no policy, has established no standard, has laid down no rule” to guide the exercise of power.⁴⁰⁷ If delegated power is “not canalized within banks that keep it from overflowing” but is rather “unconfined and vagrant,” it violates this rule.⁴⁰⁸ Delegation is proper so long as Congress lays down “an intelligible principle to which the [delegatee] is directed to conform.”⁴⁰⁹ This doctrine extends beyond administrative law to criminal law as well. As Dan Kahan has argued, judges use the rule of lenity as a nondelegation doctrine for criminal laws.⁴¹⁰

What is so odd about the elements rule is that it is a reverse nondelegation doctrine—a must-delegate doctrine, if you will. The elements rule nominally shores up jury fact-finding at the expense of judges, but there are almost no juries left and they have no sentencing power. Sentencing power is now in the hands of the unelected Sentencing Commission, except to the extent that Congress limits and guides it. But because the elements rule fetters statutory sentence enhancements, Congress will delegate even more power to the Commission. And the way Congress delegates must change. Statutory sentence enhancements

For these reasons, Congress is much more likely to circumvent the elements rule, either by redrafting aggravators as mitigators, *see supra* Section II.B, or by delegating even more power to the Sentencing Commission to achieve the same result.

406. *See, e.g.*, *A.L.A. Schechter Poultry Corp. v. United States*, 295 U.S. 495, 541-42 (1935); *Pan. Ref. Co. v. Ryan*, 293 U.S. 388, 430 (1935). While the Court has not struck down any statutes on nondelegation grounds since the New Deal, the doctrine retains life as an aid to statutory construction and as the basis for the occasional invalidation of an agency’s regulations. *See, e.g.*, *Indus. Union Dep’t v. Am. Petroleum Inst.*, 448 U.S. 607 (1980) (plurality opinion) (construing a statute narrowly to avoid the nondelegation doctrine); *Nat’l Cable Television Ass’n v. United States*, 415 U.S. 336 (1974) (same).

407. *Pan. Ref.*, 293 U.S. at 430.

408. *Schechter Poultry*, 295 U.S. at 551 (Cardozo, J., concurring).

409. *J.W. Hampton, Jr. & Co. v. United States*, 276 U.S. 394, 409 (1928).

410. Dan M. Kahan, *Lenity and Federal Common Law Crimes*, 1994 SUP. CT. REV. 345.

“canalized [the Commission’s discretion] within banks” and set out “intelligible principle[s].”⁴¹¹ These statutes are now unconstitutional. Today, delegation to the Sentencing Commission is permissible only if the community’s conscience does *not* limit it. In short, the elements rule will further dilute the influence of the conscience of the community by further insulating sentencing commissions from democratic control.

4. *Notice*

Due process protects not only an opportunity to be heard, but notice as well. So do the Grand Jury and Notice Clauses. It is not clear how large a problem notice is in practice; many judges, prosecutors, and defense lawyers already advise defendants of enhancements. Assume arguendo, with the Court and commentators, that notice is a significant problem. The elements rule does provide some notice by indictment, but the rule is not well tailored to its goal.

First, requiring notice in the indictment gives prosecutors power to manipulate statutory maxima by bargaining over the charges in the indictment. This may lead to arbitrariness and disparity.⁴¹² Second, indictments may prejudice those few defendants who go to trial, by broadcasting enhancement allegations to juries.⁴¹³

Third, the elements rule requires federal indictments to allege the facts underlying enhancements. This is helpful to defendants, insofar as it goes. In a world of guilty pleas, however, the law should provide more information to insure knowing and voluntary pleas.⁴¹⁴ Before they plead guilty, defendants need to know the maximum penalties they face, including enhancements.⁴¹⁵ Good attorneys can, and do, calculate the enhanced penalties from an indictment, but the law needs safeguards in case the attorney does not do so or makes an error. A better rule would ensure that defendants know the enhanced statutory maxima to which they are pleading. The elements rule does not.⁴¹⁶

* * *

The Court is rightly concerned about protecting the values of notice, hearing, and the conscience of the community. (It should have added equal treatment and checking arbitrariness to its list as well.) But it views its job

411. *See supra* text accompanying notes 408-409.

412. *See supra* Subsection III.A.2.

413. *See supra* text accompanying notes 292-293.

414. *See infra* Subsection III.B.1.

415. FED. R. CRIM. P. 11(c)(1) does require notice of maximum penalties in federal cases, but the rules do not govern state cases and do not mention sentence enhancements.

416. *See infra* Subsection III.B.1 (proposing a rule that would do so).

solely as safeguarding jury trials that no longer exist. Thus, its wooden attempt to implement the Fifth and Sixth Amendments founders on the shoals of plea bargaining and modern sentencing. For as long as we live in a world of guilty pleas, we must think harder about how to implement these constitutional values in the real world. The next Section proposes rules designed for this changed landscape, alternatives to the elements rule that would work better in the real world.

B. *Solutions Tailored to the World of Guilty Pleas*

1. *Notice of Maximum Sentences*

The best insight in *Apprendi* is the one the Court developed least. A defendant needs advance notice of sentence enhancements “in order that he may prepare his defence accordingly.”⁴¹⁷ Traditionally, the law gave this notice by indictment. (Once again, grant, for the sake of argument, the Court’s and commentators’ assumption that notice is a problem.)

In a world of guilty pleas, “his defence” is no longer evidence, questioning, or argument at trial. There is no trial. The analogue is the defendant’s decision to enter a guilty plea. The law must guarantee defendants the information they need most before they plead guilty. The Court, however, has focused on side issues, distractions. There is a great show of reciting many trial rights at guilty plea allocutions. Defendants are told that they have the right to a jury, to confront witnesses, and not to incriminate themselves.⁴¹⁸ These trial rights, however, are theoretical and largely irrelevant to defendants who plead guilty.⁴¹⁹ What defendants really need to know is the price of the plea bargain, the maximum sentences to which they are agreeing, including enhancements. Without this information, the plea is not fully intelligent and voluntary.

All guilty pleas must be intelligent and voluntary.⁴²⁰ Due process requires “real notice of the true nature of the charge” before a plea of guilty.⁴²¹ Courts must question each defendant to ensure “a full understanding of what the plea connotes and of its consequence.”⁴²² A plea is intelligent if the defendant understands “the nature of the charge against

417. *Apprendi v. New Jersey*, 120 S. Ct. 2348, 2356 (2000) (quoting ARCHBOLD, *supra* note 153, at 44).

418. See FED. R. CRIM. P. 11(c)(3); *Boykin v. Alabama*, 395 U.S. 238, 243 (1969).

419. For example, the pressures to plead guilty force defendants to surrender their privileges against self-incrimination. See Langbein, *supra* note 384, at 12-14, 21.

420. *Brady v. United States*, 397 U.S. 742, 747-48 & nn.4-6 (1970); *Boykin*, 395 U.S. at 241, 242 (using the words “knowing” and “intelligent” interchangeably).

421. *Smith v. O’Grady*, 312 U.S. 329, 334 (1941) (decrying deception by prosecutors concerning the punishment range).

422. *Boykin*, 395 U.S. at 244; *accord Brady*, 397 U.S. at 748.

him,” is of sound mind, and has competent counsel.⁴²³ Counsel is important to permit “an intelligent assessment of the relative advantages of pleading guilty.”⁴²⁴ A plea is voluntary if the defendant is not bribed, threatened, or misled and is “fully aware of the direct consequences.”⁴²⁵

The “direct consequences” should include the maximum penalty to which the defendant is agreeing, including any enhancements. Defendants cannot make “an intelligent assessment of the relative advantages of pleading guilty” without notice of enhancements. Most courts treat the base maximum sentence as a direct consequence and require notice to defendants.⁴²⁶ The Federal Rules of Criminal Procedure enshrine this requirement for federal cases.⁴²⁷

The Court has not, however, extended this logic to sentence enhancements. It has held that due process does not require preplea notice of a recidivism enhancement. It has treated enhancements as entirely separate from the underlying offenses. So, the Court has reasoned, postplea, presentencing notice suffices to enable defendants to contest enhancements at sentencing.⁴²⁸

These cases miss the point. Defendants need notice not only to contest enhancements at sentencing, but to decide whether to plead guilty and on what terms. Nonetheless, these cases remain the law.⁴²⁹ It is time to overrule these cases and to require preplea notice of all statutory maxima, including enhancements.

423. *Brady*, 397 U.S. at 756; accord *Bousley v. United States*, 523 U.S. 614, 618 (1998).

424. *Brady*, 397 U.S. at 748 n.6.

425. *Id.* at 755 (quoting *Shelton v. United States*, 246 F.2d 571, 572 n.2 (5th Cir. 1957) (en banc) (internal quotation marks omitted), *rev'd on other grounds*, 356 U.S. 26 (1958)).

426. See, e.g., *United States v. Salmon*, 944 F.2d 1106, 1130 (3d Cir. 1991); *Micheaux v. Collins*, 911 F.2d 1083, 1089 (5th Cir. 1990). *But cf.* *Kelleher v. Henderson*, 531 F.2d 78, 81-82 (2d Cir. 1976) (holding that Rule 11(c)(1) is not a constitutional requirement and refusing to invalidate a plea for failure to advise the defendant of the maximum sentence, because the judge had warned the defendant that he faced a “stiff sentence”); *People v. Piffer*, 198 N.W.2d 907 (Mich. Ct. App. 1972); *Gibson v. State*, 641 So. 2d 1163, 1166 (Miss. 1994) (treating the failure to advise as a harmless error, and assuming without an evidentiary hearing that it did not affect the decision to plead).

427. FED. R. CRIM. P. 11(c)(1).

428. *Oyler v. Boles*, 368 U.S. 448, 452-53 (1962); *Graham v. West Virginia*, 224 U.S. 616, 624-25 (1912).

429. The Court reaffirmed *Graham* and *Oyler* as recently as two years ago in *Almendarez-Torres v. United States*, 523 U.S. 224, 243-44 (1998). *Apprendi* did not touch the holdings of these cases, as the Court carved out recidivism as an exception to its rule. Even if one views *Apprendi* as having undercut the bases of these cases, lower courts must continue to follow them. “[I]f a precedent of [the Supreme] Court has direct application in a case, yet appears to rest on reasons rejected in some other line of decisions, the Court of Appeals should follow the case which directly controls, leaving to [the Supreme] Court the prerogative of overruling its own decisions.” *Agostini v. Felton*, 521 U.S. 203, 237 (1997) (quoting *Rodriguez de Quijas v. Shearson/Am. Express, Inc.*, 490 U.S. 477, 484 (1989)). Accordingly, lower courts continue to rely on *Oyler* to uphold enhancements in the absence of advance notice. E.g., *United States v. Moore*, 208 F.3d 411, 414 (2d Cir. 2000); *United States v. Kennedy*, 952 F. Supp. 5, 7 (D.D.C. 1997).

This notice, however, should not have to be in the indictment. Allowing notice via other means would prevent indictments from prejudicing juries in those few cases that go to trial.⁴³⁰ More importantly, allowing courts to give notice would diminish prosecutors' ability to manipulate statutory maxima. Under the elements rule, notice depends entirely on prosecutors. A prosecutor could, for example, agree with defense counsel not to cite the serious-bodily-injury enhancement to the carjacking statute in the indictment. Doing so would cap the sentence at fifteen years, the statutory maximum, even if the defendant had killed the victim.⁴³¹ The elements rule makes this manipulation possible by giving talismanic significance to the indictment.

A better rule would require courts, not prosecutors, to give notice of the overall statutory maxima, including enhancements. The best time to give notice is at plea allocutions.⁴³² (Ideally, the court would give notice at arraignment as well, to facilitate plea bargaining. This notice should be encouraged but not required, as arraignment may be too early to foresee which enhancements may apply.) At both arraignment and plea hearing, the judge addresses the defendant in open court.⁴³³ This targeted colloquy is more likely to make an impression on the defendant than a statutory citation buried in an indictment. Judges can confirm on the record that a defendant understood a plea colloquy, forestalling any claim that a lawyer misadvised the defendant. If an enhancement surprises a defendant, perhaps because counsel overlooked it, the court can adjourn the plea until the defendant has thought it over.⁴³⁴ Perhaps most defendants already receive this notice from

430. See *supra* notes 292-298 and accompanying text. For example, Congress has required pretrial notice of recidivism in drug cases via prior felony informations. 21 U.S.C. § 851(a) (1994). Because the information is separate from the indictment, it does not go to the jury and so cannot prejudice the jury.

431. See Bowman, *supra* note 396, at 188-89 (noting that putting the notice requirement in the hands of prosecutors may invite manipulation); *supra* Subsection III.A.2.

432. As for those few defendants who go to trial, prosecutors could serve them with separate informations before trial, laying out the enhancements. Putting the enhancement in a separate document, not the indictment, would avoid prejudicing the jury. Prosecutorial manipulation of enhancements is not a concern in the few cases going to trial, as the parties are in an adversarial, rather than a collusive, stance. Cf. 21 U.S.C. § 851(a) (laying out a similar scheme for prior felony informations for drug convictions).

433. See FED. R. CRIM. P. 10, 11(c).

434. In ordinary cases, the judge could simply recite the various permutations. For example, under the carjacking statute, the judge could say: "The maximum penalty is ordinarily fifteen years, or twenty-five years if I find a victim suffered bodily injury, or life if I find a victim died." The judges could use a routine checklist of enhancements, especially for the most commonly charged crimes. In complex cases, where many conceivable enhancements are scattered across far-flung statutes, the court might seek the assistance of the probation officer who is to prepare the presentence report. The officer could begin reviewing arrest reports and police notes before the plea and advise the court whether there are any special enhancements that might apply on the facts. (The officers must already do the same work to prepare for sentencing, see FED. R. CRIM. P. 32(b); it would simply happen a bit earlier in the process.) The court and probation officer need not make a conclusive determination about enhancements this early in the process; they need only recite all the possibilities on the facts. Knowing of this independent investigation, prosecutors

their lawyers, but the law should ensure that all defendants receive clear notice from the court.⁴³⁵ (It would be desirable to go further and advise defendants before they plead guilty of their Sentencing Guidelines ranges as well. This, however, would require doing presentence investigations before plea hearings, which is likely not feasible given time constraints.)

2. *Procedural Protections at Sentencing*

As noted, because most defendants never go to trial, their only real opportunity to be heard is at sentencing. To ensure an adequate hearing, the law should translate some of the procedural safeguards of trials into sentencing hearings. While many safeguards will not apply or be workable at sentencing, rights of compulsory process, cross-examination, and discovery should apply.

It is neither possible nor desirable to make sentencing resemble trial in every respect. For example, the rules of evidence do not apply at sentencing.⁴³⁶ Some commentators have proposed extending the rules of evidence to sentencing,⁴³⁷ but this proposal is unwarranted. At trial, the rules use judges as gatekeepers to keep lay jurors from improperly weighing evidence.⁴³⁸ Judges who find facts at sentencing would find it hard simultaneously to be gatekeepers who keep themselves from learning of evidence.⁴³⁹ True, some might be able to perform the requisite mental gymnastics, and rules of evidence would limit the grounds a judge could articulate in a ruling, but these effects would be modest at best. Moreover,

would ordinarily disclose to the court all possible enhancements. Indeed, the court could question counsel at the plea about whether they are aware of any other enhancements that might conceivably apply.

435. It is possible that a court may overlook an obscure enhancement, even after questioning counsel, and learn of it later from the presentence report. This is particularly likely to be true if new facts develop: if, for example, the probation officer finds a new witness or learns that an ailing victim has died. Where such facts would increase the sentence above the statutory maximum stated at the plea, the best solution may be a presumption in favor of allowing defendants to withdraw their pleas. Steven D. Clymer, *Assessing Proposals for Mandatory Procedural Protections for Sentencings Under the Guidelines*, 12 FED. SENTENCING REP. 212, 215 (2000). (The presumption might be rebutted if, for example, defense counsel had mentioned the enhancement to the defendant.) There is already an analogue in the Federal Rules, allowing withdrawal of a plea when the judge rejects a plea agreement. FED. R. CRIM. P. 11(e)(4).

436. See 18 U.S.C. § 3661 (1994) (stating that hearsay is admissible at sentencing); FED. R. EVID. 1101(d)(3).

437. See, e.g., Sara Sun Beale, *Procedural Issues Raised by Guidelines Sentencing: The Constitutional Significance of the "Elements of the Sentence,"* 35 WM. & MARY L. REV. 147, 157-58 (1993); José A. Cabranes, *The U.S. Sentencing Guidelines: Where Do We Go from Here?*, 12 FED. SENTENCING REP. 208, 209-11 (2000); Mark David Harris, *Raising the Quality of Evidence at Sentencing*, 5 FED. SENTENCING REP. 102 (1992).

438. E.g., FED. R. EVID. 103(c); see, e.g., *Daubert v. Merrell Dow Pharms.*, 509 U.S. 579, 597 (1993) (noting a "gatekeeping role for the judge"); Christopher B. Mueller, *Meta-Evidence: Do We Need It?*, 25 LOY. L.A. L. REV. 819, 822-23 (1992) (discussing the hearsay doctrine).

439. Mirjan Damaška, *Free Proof and Its Detractors*, 43 AM. J. COMP. L. 343, 350-52 (1995).

experienced judges are better able to discount hearsay and the like than are lay jurors.⁴⁴⁰ Historically, judges have relied on hearsay at sentencing to gain the fullest possible picture of the defendant's character and background.⁴⁴¹ Constricting the evidence at sentencing would, to an extent, make judges exercise their discretion in the dark.⁴⁴²

While the sources of information should not be constricted, defendants should be able to challenge them at sentencing. Defendants already have the right to counsel at sentencing,⁴⁴³ as well as discovery of evidence that "would tend to . . . reduce the penalty."⁴⁴⁴ But they have no constitutional rights to compulsory process, confrontation, or cross-examination at sentencing.⁴⁴⁵ As argued, sentencing judges need leeway to consider hearsay that would be inadmissible at trial. At the very least, however, defendants should be able to use subpoenas to bring hearsay declarants into court and cross-examine them at sentencing. To exercise this right, defendants need to know what evidence is being used against them. Thus, discovery rules should extend to sentencing and to the facts underlying the presentence report.⁴⁴⁶ This discovery should not be too burdensome. Federal prosecutors already disclose much of this information before pleas and, as a matter of policy, provide more discovery before sentencing.⁴⁴⁷

These rules strike a reasonable balance. Where there is no genuine dispute or where the judge will disregard the disputed issue, the use of hearsay would save much time and expense. On those crucial contested issues, however, these rules would provide adequate safeguards. They would translate the values of the Confrontation, Compulsory Process, and Due Process Clauses into the world of guilty pleas.⁴⁴⁸

440. Clymer, *supra* note 435, at 213.

441. *E.g.*, *Williams v. New York*, 337 U.S. 241, 246-47, 250-51 (1949).

442. *Id.*

443. *Mempa v. Rhay*, 389 U.S. 128 (1967).

444. *Brady v. Maryland*, 373 U.S. 83, 88 (1963).

445. *See Williams*, 337 U.S. at 245; *United States v. Petty*, 982 F.2d 1365, 1367-69 (9th Cir. 1993) (collecting cases); *United States v. Silverman*, 976 F.2d 1502, 1511 (6th Cir. 1992) (en banc); *see also United States v. Prescott*, 920 F.2d 139, 144 (2d Cir. 1990) (holding that the defendant has no absolute right to call witnesses at sentencing). *See generally* David A. Hoffman, Note, *The Federal Sentencing Guidelines and Confrontation Rights*, 42 DUKE L.J. 382 (1992) (bemoaning the absence of confrontation rights); Note, *An Argument for Confrontation Under the Federal Sentencing Guidelines*, 105 HARV. L. REV. 1880, 1883 (1992) (same).

446. *Bowman*, *supra* note 396, at 192; Richard Smith-Monahan, *Unfinished Business: The Changes Necessary To Make Guidelines Sentencing Fair*, 12 FED. SENTENCING REP. 219, 219-20 (2000).

447. *See* FED. R. CRIM. P. 16; U.S. DEP'T OF JUSTICE, UNITED STATES ATTORNEY'S MANUAL 9-27.750 (1997), *reprinted in* 12 FED. SENTENCING REP. 233, 238 (2000).

448. I am not worshipping or reifying these procedures as ends in themselves, but rather valuing them as means to achieve fair outcomes. These rights will do so by their very existence, even when they are not in fact exercised. Many defendants will not in fact exercise these hearing rights, instead striking more generous plea bargains with prosecutors. But these plea bargains will take place in the shadow of these hearing rights and so be influenced by them. Thus, the mere

Professor Sara Sun Beale and others have proposed going even further, raising the idea of proof beyond a reasonable doubt at sentencing.⁴⁴⁹ This change, however, would go too far. The procedures proposed above would improve fact-finding at sentencing by ensuring adversarial testing. There is little need to require proof beyond a reasonable doubt as well. That standard is appropriate to protect the innocent from being branded criminals. As Pollock and Maitland put it, “felony” is “as bad a word as you can give to man or thing.”⁴⁵⁰ Because the innocent have a powerful interest in avoiding wrongful conviction, an interest far more important than punishing the guilty, the scales of justice tilt strongly toward them.⁴⁵¹ Once the state has proved a crime beyond a reasonable doubt, however, the defendant is already stigmatized.⁴⁵² The enhancement adds no additional stigma.⁴⁵³ And while the guilty do have an interest in not being overpunished, the state has a countervailing interest in not underpunishing them. By definition, the guilty need some punishment; the only question is how much. There is no compelling reason to tilt the scales so strongly in favor of criminals, systematically underpunishing them. Moreover, judges can exercise discretion and find facts covertly at sentencing without any burden of proof at all.⁴⁵⁴ Requiring proof beyond a reasonable doubt would only invite wholesale legislative circumvention via increased judicial discretion, giving defendants less protection. Requiring clear and convincing evidence might be a reasonable compromise. This verbal formula, however, might not differ much in practice from a preponderance of the evidence, and this standard might still trigger legislative circumvention.⁴⁵⁵

Judged by these standards, Apprendi’s own plea was abundantly fair. Apprendi himself had notice, lots of it. His plea agreement expressly warned him that the prosecution would seek the hate-crime enhancement.⁴⁵⁶ His lawyer reviewed the charges, maxima, and plea

possibility of judicial hearings will check prosecutorial excesses, even if the actual number of hearings is small.

449. *E.g.*, Beale, *supra* note 437, at 160; Cabranes, *supra* note 437, at 210.

450. 2 SIR FREDERICK POLLOCK & FREDERIC WILLIAM MAITLAND, *THE HISTORY OF ENGLISH LAW* 465 (Boston, Little, Brown & Co. 2d ed. 1899), *quoted in* *Staples v. United States*, 511 U.S. 600, 618 (1994).

451. *See In re Winship*, 397 U.S. 358, 363-64 (1970).

452. *See, e.g.*, *McMillan v. Pennsylvania*, 477 U.S. 79, 92 n.8 (1986) (noting that criminal sentencing occurs only after a defendant has been found guilty beyond a reasonable doubt).

453. *See supra* note 248 and accompanying text.

454. *See, e.g.*, *McMillan*, 477 U.S. at 91 (“Sentencing courts have traditionally heard evidence and found facts without any prescribed burden of proof at all.”). *Apprendi* preserved this discretion. *Apprendi v. New Jersey*, 120 S. Ct. 2348, 2358 (2000) (“We should be clear that nothing in this history suggests that it is impermissible for judges to exercise discretion—taking into consideration various factors relating both to offense and offender—in imposing a judgment *within the range* prescribed by statute.”).

455. *See Bowman*, *supra* note 396.

456. *Apprendi*, 120 S. Ct. at 2352.

agreement with him ahead of time.⁴⁵⁷ His plea colloquy went over, in great detail, his maximum sentences with and without the enhancement. The judge and his lawyer warned him that his combined maximum sentences were twenty years without the enhancement and thirty years with it. Apprendi said he understood these maxima.⁴⁵⁸ He had an adversarial evidentiary hearing on the biased-purpose issue.⁴⁵⁹ He testified at that hearing. His lawyer put on eight other favorable witnesses and confronted and cross-examined the state's witnesses.⁴⁶⁰ His twelve-year sentence, while greater than the ten-year unenhanced maximum on Count 18, was well within the notice he received. Apprendi's notice was abundant, clear, and timely. He had a full and fair opportunity to be heard. Due Process and the Sixth Amendment require nothing more.

3. *Historic Elements and Facts That Create Felonies*

A less procrustean, less sweeping rule might achieve the Court's core goals without wreaking havoc in the world of guilty pleas. The elements rule is procrustean in treating all sentence enhancements (except recidivism enhancements) alike. The intuition that drove the cases, though, focused on a narrower subset. Specifically, the real concern is redefining traditional elements of traditional crimes. Recall *Mullaney*, where Maine tried to brand Wilbur a murderer while presuming malice aforethought against him.⁴⁶¹ *Mullaney* stressed the long common-law tradition of making this element crucial to culpability.⁴⁶² Recall *McMillan's* and *Almendarez-Torres's* statements about how legislatures were not evading *Winship* by redefining traditional elements.⁴⁶³ Recall *Jones's* suggestion that states cannot omit traditional elements from crimes.⁴⁶⁴ All of these cases trade on the same intuition: There is something special about monkeying with bedrock elements of well-known common-law crimes.

In this light, one of *Apprendi's* more dubious assertions makes sense. The Court asserted that "it is obvious" that sentence enhancements increase stigma.⁴⁶⁵ As a general matter, it is not at all "obvious." When legislatures create new crimes, the stigma depends on the legislature's labeling and

457. Joint Appendix, *supra* note 137, at 17-19.

458. *Id.* at 17-24.

459. Petition for Writ of Certiorari at 98a-121a, 175a-299a, *Apprendi* (No. 99-478).

460. *Id.*

461. *Mullaney v. Wilbur*, 421 U.S. 684, 685-87 (1975).

462. *Id.* at 696.

463. *Almendarez-Torres v. United States*, 523 U.S. 224, 243-44 (1998); *McMillan v. Pennsylvania*, 477 U.S. 79, 89 (1986).

464. *Jones v. United States*, 526 U.S. 227, 241-42 (1999) (suggesting a possible interpretation of *Patterson*).

465. *Apprendi v. New Jersey*, 120 S. Ct. 2348, 2359 (2000).

definition of the crime.⁴⁶⁶ People and rap sheets identify criminals by the offense of conviction, not the enhancement.⁴⁶⁷ But in *Mullaney*, where a common-law element changed the common-law label from manslaughter to murder, the stigma was obvious. Common-law crimes have established stigma commensurate with their elements.⁴⁶⁸ Many people know that murder is in cold blood and manslaughter is in the heat of passion. Maine's sin in *Mullaney* was deception: It branded Wilbur a murderer while proving only those elements that, in people's minds, made him a manslaughterer.

The elements rule makes sense for the classic elements of classic common-law crimes, particularly mens rea requirements. Manipulating common-law elements deceives criminals about the stigma they will earn and deceives the public about the severity of crimes. To brand someone a murderer, a manslaughterer, a rapist, a burglar, or an arsonist, the legislature must use truth in labeling. In this sense, and in this sense only, is the majority right about stigma.

One would not want to freeze all common-law elements in place. The requirement that a burglary occur at night, for example, is not central to stigma or culpability. Thus, one faces difficult line-drawing issues in deciding which elements are crucial to stigma. As a first approximation, one could limit this rule to mens rea requirements, though presumably certain elements of the actus reus would be covered as well. (For instance, it would be deceptive to brand as a murderer someone whose victim did not die.)⁴⁶⁹

The advantages of limiting the elements rule to historic elements of common-law crimes are clear. Legislatures would have a choice. They could either tap into common-law stigma, in which case truth in labeling would require the core common-law elements. On the other hand, they could create brand new crimes, such as voluntary homicide, or illegal reentry, or carjacking. With these new crimes, legislatures should be able to

466. See Dan M. Kahan, *Gentle Nudges vs. Hard Shoves: Solving the Sticky Norms Problem*, 67 U. CHI. L. REV. 607, 624 (2000) (noting that reluctant decisionmakers would be more likely to enforce the new statutory crimes of "indecent contact" or "indecent assault" than rape in some circumstances because "its very name signals that it warrants less condemnation than does rape").

467. See *supra* note 248 and accompanying text.

468. Cf. *Patterson v. New York*, 432 U.S. 197, 226-27 (1977) (Powell, J., dissenting) (noting that certain elements of common-law crimes made a significant difference to both stigma and punishment and proposing that the Constitution require the prosecution to bear the burden of proving such facts).

469. This principle also suggests its own limit. Because the reasoning hinges on truth in labeling, the legislature could take other steps to change these elements while maintaining truth in labeling. For example, the label of the crime could mention the omission of the element, or an advertising campaign could educate the public of the change. While the law has traditionally presumed that citizens know the law, the approach in the text turns this fictitious presumption on its head. It presumes that citizens know well-settled common-law crimes, but that they do not know the intricate details of modern statutes unless the legislature takes affirmative steps to make these details known.

define them however they please.⁴⁷⁰ Either way, the stigma attaching to the defendant would still be up to the conscience of the community: the jury (in those few jury trials) and the legislature. This narrower rule would stop possible erosion without overly disrupting the world of guilty pleas.

The same logic applies to any enhancement that turns a misdemeanor into a felony. As Pollock and Maitland noted, the label “felon” carries a venerable common-law pedigree and stigma.⁴⁷¹ The truth-in-labeling principle means that to brand someone a felon, one must convict the criminal of a felony. In addition, a fact that enhances penalties to felony levels triggers certain constitutional procedural protections.⁴⁷² Thus, any fact that turns a misdemeanor into a felony should be an element of the offense.⁴⁷³

Juries were meant to check the imposition of stigma, as reflected in the Grand Jury Clause’s limitation to “infamous” crimes.⁴⁷⁴ Jury determination of stigma remains a viable role in the few jury trials, whereas today jury determination of punishment is an unmitigated fiction. Judges tell juries nothing about punishment, admonish them not to consider it, and give them no role in it.⁴⁷⁵ There is no sense in which a jury verdict authorizes any particular level of punishment. The verdict simply writes a blank check for legislators and judges to fill in.⁴⁷⁶ It makes no difference to the jury whether the maximum is fifteen years with a possible ten-year enhancement, or twenty-five years with a possible ten-year reduction. Jury

470. One could conjure up fanciful violations of the truth-in-labeling principle, such as a “homicide” crime that did not make killing an element. One could also imagine legislative efforts to brand someone by broadcasting the fact of the enhancement—if, for example, New Jersey distributed flyers branding Apprendi a violent racist, even though his offense of conviction did not include biased purpose as an element. Perhaps the truth-in-labeling rule would need to extend to these situations as well, requiring that they be elements of the offense. (Intuitively it might seem that the very imposition of a bias-crime enhancement would stigmatize Apprendi, but it cannot do that unless the state broadcasts this fact.)

471. See *supra* note 450 and accompanying text.

472. See, e.g., *Lewis v. United States*, 518 U.S. 322, 328 (1996) (holding that any statute authorizing an incarceration for more than six months triggers the constitutional right to a jury trial); King & Klein, *supra* note 3 (manuscript at 8) (noting that rights to grand and petit juries are at stake).

473. Even before *Apprendi*, several cases used this reasoning to treat as elements facts that elevate misdemeanor penalties to felony levels. See, e.g., *United States v. Stone*, 139 F.3d 822, 837-38 (11th Cir. 1998); *United States v. Sharp*, 12 F.3d 605, 606 (6th Cir. 1993); see also King & Klein, *supra* note 3 (manuscript at 8 n.21) (collecting sources).

474. U.S. CONST. amend. V; see also William F. Jung, Note, *Recognizing a Corporation’s Rights Under the Indictment Clause*, 1983 U. ILL. L. REV. 477, 488, 499-509 (proposing that the Grand Jury Clause should apply to corporations because it was meant to check imposition of stigma and infamy, and arguing that grand jury secrecy was meant to protect innocent defendants from such infamy); cf. Akhil Reed Amar, *Foreword: Sixth Amendment First Principles*, 84 GEO. L.J. 641, 661 (1996) (suggesting that the Grand Jury Clause’s check on “infamous” crimes is analogous to laws restricting defamation).

475. See, e.g., *Shannon v. United States*, 512 U.S. 573, 587 (1994) (reiterating “the rule against informing jurors of the consequences of their verdicts” and rejecting a proposed exception).

476. *Supra* note 249.

control of punishment has long been a lost cause; we must trust the constitutional principle of proportionality to prevent excessive punishment.⁴⁷⁷ In contrast, jurors know the stigma they are imposing, because they know the names of the crimes charged. The proposed narrowing of the rule would protect defendants against more stigma than their juries or their pleas authorized.

* * *

The above account suggests better ways to implement the Constitution's values in the world of guilty pleas. The rules of notice and historic elements would ensure fair warning and make the law more truthful and transparent to democratic scrutiny. And the rules of notice and procedural protections would ensure a fair opportunity to be heard. These rules would do so without creating more statutory maxima for prosecutors to manipulate arbitrarily in charge bargaining. The populism of legislatures and the hearings at plea colloquies and sentencing can implement these values in our changed legal landscape.

IV. CONCLUSION

The opinions in *Jones* and *Apprendi* and the academic articles supporting them are not isolated errors or flukes. They are symptomatic of criminal procedure's broader, anachronistic obsession with jury trials at the expense of pleas, and its divorce of trials from sentencing. It is revealing that none of the articles, opinions, or briefs even considered, much less heeded, the guilty plea and sentencing considerations set forth above. The myth of the jury trial is deeply embedded in our culture, our psyche, and our law school curricula. But today, the reality is but a faint echo of a bygone era. American policy makers have shown no signs of uprooting the status quo. Now that guilty pleas are a fixture of the legal landscape, scholarship must adapt to take them into account.

This is not to say that scholars should abandon study of jury trials. Jury trials remain important, particularly because plea bargaining happens in part in the shadow of expected trial outcomes. It is time, however, to think hard about tailoring other rules of criminal procedure to the world of guilty pleas. This case study has illuminated some of the principles that need to guide this inquiry. One principle is that criminal procedure must protect defendants by ensuring a fair balance of power between judges and prosecutors. The Court and commentators failed to see that the elements

477. See *Harmelin v. Michigan*, 501 U.S. 957, 997-98 (1991) (Kennedy, J., concurring in part and concurring in the judgment, joined by O'Connor and Souter, JJ.); *Solem v. Helm*, 463 U.S. 277 (1983).

rule gives prosecutors much more power in plea bargaining and makes it harder for judges to check them. A similar concern for the dynamics of plea bargaining should inform other procedural rules.

For example, the standards for ineffective assistance of counsel are based on the model of the trial. *Strickland v. Washington*⁴⁷⁸ asks whether there is a reasonable probability that an error affected the outcome. This test works well for trials. Courts can mentally add a new piece of evidence or argument to the scales and ask whether it would have changed a hypothetical jury's verdict. The Court has extended this same rule to guilty pleas, asking whether the defendant would have insisted on going to trial.⁴⁷⁹ This rule is more problematic for pleas, because it depends on defendants' self-serving post hoc statements about what they would have done. This rule also ignores errors affecting how favorable the plea bargain was, even if the defendant would not have gone to trial. Perhaps a different rule for pleas might better reflect the dynamics of plea bargaining.

Another example is the prosecutor's duty under *Brady* to disclose exculpatory and impeachment evidence.⁴⁸⁰ The Court treats evidence as material if there is a reasonable probability that it would have affected the outcome of the trial.⁴⁸¹ Once again, this test works much less well for pleas than for trials. How does one make a counterfactual prediction when the person being predicted cannot be trusted to tell the truth? Does one try to predict whether the defendant would have gotten a better plea bargain? Or does one instead try to predict whether the defendant would have gone to trial, or whether the defendant would have been convicted at trial? In other words, should it matter if the withheld evidence would have affected the dynamics of plea bargaining? And how early must prosecutors turn over this evidence? Is *Brady* as important for defendants who admit guilt, and is it a right waivable in the plea bargain? Can a deep discount for an early plea bargain be viewed as compensation for forgoing the possibility of *Brady* material? The literature is beginning to address the waiver questions, but there is much more to ask.

In addition, criminal procedure needs to fill the vacuum left by the disappearance of jury trials. For instance, jury trials served as morality plays, educating the defendant and community and serving an expressive, cathartic function. Now that jury trials are rare, the law needs other ways to educate, express, and allow catharsis in guilty-plea cases. Perhaps judges should play a more overtly moralistic role at plea hearings and sentencings. Perhaps victims could tell their stories at these proceedings, venting their

478. 466 U.S. 668 (1984).

479. See *Hill v. Lockhart*, 474 U.S. 52 (1985).

480. *Brady v. Maryland*, 373 U.S. 83, 87 (1963); see *United States v. Bagley*, 473 U.S. 667 (1985); *United States v. Agurs*, 427 U.S. 97 (1976).

481. *Bagley*, 473 U.S. at 682.

sorrows and pains and maybe even expressing forgiveness. Perhaps defendants could be made to learn about the impact of their crimes and could be encouraged to express remorse and try to make amends. Perhaps defendants and victims could even meet as part of the healing process. All of these procedures could help to fill the void left by the disappearance of jury trials.

Scholars also need to pay more attention to sentencing. Sentencing remains an academic backwater, divorced from criminal law and procedure. Yet it greatly influences the choices prosecutors and defendants make in filing charges, plea bargaining, and going to trial. Sentencing considerations should inform the choice of procedures further upstream. Academic proposals and new judicial decisions ignore the sentencing implications of rules at their peril. It is astonishing that more than a decade after the enactment of the Sentencing Guidelines, there are still no textbooks devoted to them. It is time to integrate sentencing into law-school curricula and academic criminal scholarship. For example, this Article exposes how complete prosecutorial control over recidivism enhancements (such as prior drug felony informations) puts enormous pressure on defendants to plead guilty.⁴⁸² Yet the criminal procedure literature has ignored this problem. The time is ripe for more literature discussing the structure of mandatory minima and enhancements. If we are to retain these enhancements at all, we must propose better ways to implement them, or else justify why such massive pressure to plead guilty is desirable.

This Article's case study, while not a comprehensive answer, is a step toward asking the right questions. Scholarship need not denigrate or ignore jury trials, particularly because plea bargaining happens in part in the shadow of expected trial outcomes. But scholars must also adapt to reflect the real world of guilty pleas and sentencing.

482. *Supra* note 342 and accompanying text.