

1-1-2005

The Myth of Uniformity

Michael M. O'Hear

Marquette University Law School, michael.ohear@marquette.edu

Follow this and additional works at: <http://scholarship.law.marquette.edu/facpub>

 Part of the [Law Commons](#)

Publication Information

Published as Michael M. O'Hear, *The Myth of Uniformity*, 17 Fed. Sent'g Rep. 249 (2005). © 2005 by the Regents of the University of California on behalf of the Vera Institute of Justice. Copying and permissions notice: Authorization to copy this content beyond fair use (as specified in Sections 107 and 108 of the U. S. Copyright Law) for internal or personal use, or the internal or personal use of specific clients, is granted by the Regents of the University of California on behalf of the Vera Institute of Justice for libraries and other users, provided that they are registered with and pay the specified fee via Rightslink® on Caliber (<http://caliber.ucpress.net/>) or directly with the Copyright Clearance Center, <http://www.copyright.com>.

Repository Citation

O'Hear, Michael M., "The Myth of Uniformity" (2005). *Faculty Publications*. Paper 118.
<http://scholarship.law.marquette.edu/facpub/118>

This Article is brought to you for free and open access by the Faculty Scholarship at Marquette Law Scholarly Commons. It has been accepted for inclusion in Faculty Publications by an authorized administrator of Marquette Law Scholarly Commons. For more information, please contact megan.obrien@marquette.edu.

The Myth of Uniformity

Do the United States Sentencing Guidelines embody a coherent system of real-offense sentencing? The opinion of the “remedy majority” in *Booker* rests, in no small measure, on an assumption that they do. Writing for the remedy majority, Justice Breyer determined that Congress wanted defendants to be sentenced in a uniform fashion based on their real-world conduct, as opposed to their crimes of conviction.¹ More specifically, Congress wanted to ensure similar treatment for defendants whose real-world conduct was similar, while also ensuring different treatment for defendants whose real-world conduct was different.² In Breyer’s view, the jury-based sentencing procedures envisioned by the *Booker* dissenters would defeat Congress’s preferences because, among other things, such procedures would undermine the scheme of sentencing uniformity embodied by the Guidelines.³ Thus, he concluded, the jury-based system should be rejected in favor of a judge-based system that at least employs the Guidelines on an advisory basis. The unspoken—but critical—assumption is that adherence to the Guidelines (expected to be greater in the judge-based advisory system) actually advances uniformity goals in a robust fashion. There are good reasons, however, to doubt this assumption.

To be sure, the Guidelines do a wonderful job of identifying normally relevant offense characteristics and ensuring that they are taken into account at sentencing. Yet, this accomplishment does not amount, in and of itself, to uniformity in any meaningful sense of the term. Consider, for instance, a sentencing scheme that imposed the death penalty in every case in which a victim suffered physical injury, no matter how slight. Such a scheme would also do a wonderful job of ensuring that a normally relevant offense characteristic is taken into account at sentencing, but no one would confuse such a scheme with a coherent system of uniformity. Nor would it even necessarily do a better job of advancing uniformity goals than a scheme that imposed the same flat sentence on all offenders—a scheme that would at least have the virtue of avoiding inappropriate sentence enhancements for minor or unforeseeable injuries.⁴ The Guidelines, of course, are not nearly as crude as this hypothetical system, but their length and technical detail mask a tendency toward some of the very same types of difficulties.

This is not to say that the jury-based system of the *Booker* dissenters would do a better job of achieving uniformity than the advisory system preferred by the remedy majority. My point, rather, is that *neither* system seems especially well suited to advance uniformity objectives. In light of the Guidelines that we have—as opposed to some sort of hypothetical guidelines that actually implement a coherent vision of uniformity—the desire of Congress to achieve uniformity, standing alone, offers little help in deciding how much weight ought to be given the Guidelines or what procedures should be used to implement them. No matter how attractive uniformity might be in theory, we should not exalt uniformity to the detriment of other important objectives as we address the important questions we now confront about the future of the Guidelines.

The article develops these points as follows. Part I considers the meaning of the term “uniformity,” as the Court seems to be using it. Part II summarizes key failings of the Guidelines from a uniformity perspective. Part III discusses the implications of this uniformity analysis for the ongoing judicial debate over how much weight ought to be given the Guidelines after *Booker*. Part IV suggests a possible restructuring of the Guidelines.

I. Framing the Issue: What Is “Uniformity”?

In the *Booker* remedy opinion, Justice Breyer did not offer a rigorous definition of uniformity but, broadly speaking, seemed to have in mind the twin goals of (1) similar treatment of similarly situated offenders and (2) different treatment of differently situated offenders.⁵ Confusingly, elsewhere in the sentencing literature, these two goals are sometimes given the separate labels of “uniformity” and “proportionality,” respectively.⁶ In this article, I will follow *Booker*’s lead in using “uniformity” to embrace both objectives. Where necessary, I will disaggregate the problems of inappropriately similar treatment from inappropriately different treatment by referring to “false positives” (i.e., imposing different sentences on offenders who are actually similarly situated) and “false negatives” (i.e., imposing similar sentences on offenders who are actually differently situated).



MICHAEL M. O’HEAR

Associate Professor,
Marquette University
Law School

Federal Sentencing Reporter, Vol. 17, No. 4, pp. 249–257, ISSN 1053-9867 electronic ISSN 1533-9867

©2005 Vera Institute of Justice. All rights reserved. Please direct requests for permission to photocopy or reproduce article content through the University of California Press’s Rights and Permissions website at www.ucpress.edu/journals/rights.htm

The *Booker* remedy opinion offers hypothetical examples of these problems, which it uses to illustrate the drawbacks of a charge-offense system. First, there are the contrasting cases of “Smith” and “Jones,” both of whom violate the Hobbs Act:

Smith threatens to injure a co-worker unless the co-worker advances him a few dollars from the interstate company’s till; Jones, after similarly threatening the co-worker, causes far more harm by seeking far more money, by making certain that the co-worker’s family is aware of the threat, by arranging for deliveries of dead animals to the co-worker’s home to show he is serious, and so forth.⁷

Although the behavior of Smith and Jones is very different (in the words of the remedy majority, “the known harmful consequences of their actions are different”), a charge-offense system would result in the same sentence in both cases because the same statute has been violated.⁸ This is an example of a false negative: Jones’s sentence is not enhanced relative to Smith’s, although it should be.

Second, there are the contrasting cases of the former felons “Johnson” and “Jackson”:

[E]ach . . . engages in identical criminal behavior: threatening a bank teller with a gun, securing \$10,000, and injuring an innocent bystander while fleeing the bank. Suppose prosecutors charge Johnson with one crime (say, illegal gun possession) and Jackson with another (say, bank robbery).⁹

Under a charge-offense system, the prosecutor’s charging decision makes all the difference, producing inappropriately different sentences for Johnson and Jackson. This would be an example of a false positive.

All of this analysis, of course, assumes some sort of coherent underlying theory of sentencing. After all, every case is unique in some way. We cannot speak of uniformity in any meaningful sense without some ability to distinguish which characteristics ought to result in different sentences. Indeed, much of the appeal of uniformity rests on the assumption that uniform sentences will result in the advancement of the legitimate purposes of punishment in a consistent and principled fashion across the full range of federal criminal cases.

While the *Booker* remedy majority does not provide any systematic elaboration of its views on this point, the Smith/Jones and Johnson/Jackson hypotheticals suggest an emphasis on two types of variables in the uniformity calculus: (1) the “known harmful consequences” of the offender’s conduct and (2) the offender’s criminal history. This approach to uniformity resonates with Congress’s directive that courts “avoid unwarranted sentencing disparities among defendants with *similar records* who have been found guilty of *similar conduct*.”¹⁰ This approach also mirrors the Court’s Eighth Amendment proportionality analysis.¹¹

Our definition of uniformity can be refined on this basis. Holding criminal history constant, uniformity demands similar treatment for similar conduct and different treatment for different conduct. But how do we know what aspects of conduct count? The statute refers to “guilty” conduct, while the Court refers to knowledge and “harmful consequences.” Implicitly, there seems to be some contemplation of “just deserts,” i.e., a system in which sentences would be adjusted based on those aspects of an offender’s conduct that tend to aggravate or mitigate the moral blameworthiness of that conduct.¹²

This observation is not intended to equate uniformity with a strict system of just deserts. Indeed, the emphasis on criminal history, alongside conduct, is at least arguably inconsistent with just deserts.¹³ On the other hand, uniformity cannot be an infinitely malleable concept; the objective holds little intrinsic appeal if it lacks some sort of grounding in the leading recognized theories of criminal punishment, including just deserts. The Court’s apparent focus on harm and state of mind is perfectly consistent with this view. Thus, in considering how the Guidelines actually distinguish among offenders with similar records, I will assume that just deserts offers an appropriate and helpful point of reference.

II. The Guidelines’ Uniformity Problems

A. Circumstantial Evidence: History of the Guidelines

As drafted by the first Commission, the Guidelines were the product of what Justice Breyer himself has characterized as a series of “compromises,” perhaps the most important being the decision to employ past sentencing practices as a baseline.¹⁴ While past practices surely reflected a sensitivity to the “known harmful consequences” of offenders’ conduct, they may have also reflected other considerations that do not have as clear a bearing on uniformity (penalties for going to trial, amenability to rehabilitation, disadvantaged upbringing, parole expectation and so forth). Moreover, whatever uniformity could be achieved by the averaging of past practices may have been further diluted by other sorts of compromises. These include the decisions to limit the number of offense characteristics as a matter of administrative convenience¹⁵ and otherwise to deviate from past practices as to matters that were of special concern to particular commissioners, such as the sentencing of white-collar criminals.¹⁶

More importantly, while the Guidelines have retained the basic framework created by the original Commission, it is important to recognize that, with respect to many offense areas, the Guidelines have become every bit as much a product of Congress as they are of the Commission. This fact alone, on its face, should raise serious doubts about the role of uniformity in the Guidelines. Subject to two very different institutional masters, it would be rather surprising to discover that the Guidelines embody a coherent scheme of identifying and assigning weights to relevant offense characteristics—particularly given the tendency of

one of those institutions toward ad hoc gestures of condemnation of the crime *du jour*. And, indeed, Congress's habitual reliance on crude mandatory minimums—dutifully worked into the fabric of the Guidelines by the Commission, but vigorously criticized by Breyer himself¹⁷—demonstrates the tension between uniformity ideals and the reality of congressional intervention. One might also consider Congress's rejection of the Commission's attempt to soften the 100:1 crack cocaine policy,¹⁸ the substantial increases in sentences for economic crimes that were reluctantly adopted by the Commission in 2003 as a result of pressure from Congress,¹⁹ and Congress's rewriting of the Guidelines to toughen the sentencing for child pornography and sex crimes.²⁰

B. Missing Uniformity Factors

Despite their length and technical detail, the Guidelines are actually surprisingly silent on a number of important considerations bearing on the blameworthiness of an offender's conduct. Perhaps most notably, the Guidelines devote little attention to mens rea or motive.²¹ The Guidelines instead focus on the magnitude of actual or foreseeable harm. Yet, even as to harm, the Guidelines employ a scattershot, unsystematic approach.²²

Consider Smith and Jones, the two extortionists. They are supposed to get different sentences based on the different known harmful consequences of their behavior, which are apparently composed of (1) different amounts of money sought and (2) different degrees of fear instilled in the victim and his family. As to the first offense characteristic, the extortion guideline would distinguish between Smith and Jones only to the extent that the monetary loss in the two cases happened to fall into different recognized loss categories.²³ (More will be said about loss tables in Section D.) As to the second, the extortion guideline lacks any generalized offense characteristic dealing with victim fear. Instead, the guideline seems to get at fear through a host of more specific offense characteristics, such as whether there was a death threat, a firearm discharge, or a kidnapping. It is not clear whether any of these enhancements would apply to Jones or, if so, apply any differently to Jones than to Smith.

Or consider Johnson and Jackson, the bank robbers. It appears that the Guidelines would mandate similar sentences based on the identical harm threatened and caused by their conduct. But should they really be treated the same? Assume that Jackson's gun was loaded but Johnson's was not, indicating that Jackson's conduct was more dangerous and reflected an aggravated mens rea. Assume further that Jackson was negligent in injuring the bystander while fleeing but that Johnson undertook all reasonable precautions to prevent any harm from occurring during flight. Neither of these, nor numerous other potential distinguishing characteristics, would necessarily be considered under the Guidelines,²⁴ leading to a false negative in the sentencing of the two robbers.

To be sure, "missing" sentencing factors may, in principle, be taken into account through the departure mechanism, but departures are never required under the Guidelines, and, in practice, upward departures, at least, are quite rare.²⁵ Likewise, missing factors may be taken into account in the selection of a sentence within a range, but, again, this would be a purely discretionary matter, and, in any event, Guidelines ranges are generally quite narrow. To the extent that a sentencing judge does take into account the full range of factors bearing on just punishment, this will, in many cases, be more a result of the limited unguided discretion allowed by the Guidelines and less a result of any mandatory uniformity embedded in the Guidelines.²⁶

C. Dubious Factors

The Guidelines not only exclude some factors that ought generally to be part of the uniformity analysis but also include a variety of factors that may undermine uniformity. For instance, the acceptance of responsibility provision gives rise to substantial sentence differences based chiefly on the mode of conviction.²⁷ Thus, Jackson and Johnson—despite identical conduct and identical records—would almost certainly receive quite different sentences under the Guidelines if one pled guilty and the other went to trial. The acceptance provision may or may not be good policy, but it seems at odds with uniformity objectives.²⁸ The substantial assistance and fast-track provisions raise similar concerns.²⁹

Apart from such general provisions, the Guidelines also include a host of dubious offense-specific aggravating factors. In the robbery guideline, for instance, why are carjackers singled out for a special enhancement? Why should the robbers of financial institutions receive longer sentences than other sorts of robbers? It is hard to see why targeting a car or a bank merits extra punishment per se in the same sort of way as does inflicting bodily injury during the course of a robbery.

D. The Problem of Weighting

Even if the Guidelines were to include all of the relevant uniformity factors (and exclude all of the irrelevant one), the Guidelines would still not function effectively as a uniformity regime if they did not ascribe an appropriate weight to each of the factors. Consider Smith and Jones again. If Jones's conduct is only a little more blameworthy than Smith's, but the Guidelines mandate a dramatically longer sentence for Jones, then the cause of uniformity may actually be undermined, not advanced, by the Guidelines. Or consider the possibility of a third hypothetical extortion defendant, Baker, who presents a different mix of aggravating circumstances that place him between Smith and Jones in blameworthiness. If Baker's offense characteristics are given too much weight, or Jones's too little, then Baker and Jones might be treated the same—a false negative.

And the Guidelines have indeed been accused of giving an unjustifiably heavy weight to numerous offense and

offender characteristics. For instance, two of the most prominent targets of criticism in this regard have been the provisions dealing with crack and illegal reentry by felons.³⁰

More generally, weighting problems are suggested by the standard Guidelines approach to aggravating circumstances: grouping together broad ranges of conduct with “cliffs” at the margins, i.e., tipping points at which very small differences in blameworthiness produce sometimes quite different offense levels. Offenders at the bottom of the range for a given offense characteristic will have that characteristic overweighted relative to other offenders with the same characteristic, while offenders at the top will have the characteristic underweighted. Thus, a robber causing a \$50,000 loss is treated the same as a robber causing an \$11,000 loss but differently than a robber causing a \$50,001 loss. Robbers who possess or brandish a firearm receive the same offense level—although these categories of possessing and brandishing encompass a broad range of conduct—but the moment the firearm is discharged a quite different offense level becomes applicable—even though one can readily imagine circumstances in which the discharger creates little more danger or fear than the brandisher. (Imagine, for instance, a wild shot in the air at some distance from a safely concealed victim versus a gun brandished menacingly just inches from the victim’s face.)

To be sure, there may be some balancing out in the aggregate: some defendants will benefit unfairly by being at the top of one range but suffer an offsetting detriment by being at the bottom of another range. Additionally, the cliff problems are theoretically softened by the Guidelines’ use of overlapping sentencing ranges, so that different offense levels do not necessarily produce different sentences. Given the relative infrequency of top-of-the-range sentences in practice, though, this safeguard may be more theoretical than real.³¹ In any event, the uniformity concerns remain as to defendants who happen to be at the top or bottom of the range with respect to multiple offense characteristics. This concern is exacerbated by the ability of police and prosecutors to manipulate some offense characteristics (e.g., drug trafficking quantities in sting operations) so as to ensure that defendants “fall off the cliff.”

A related weighting concern is the double-counting of offense characteristics. Many guidelines contain a variety of overlapping characteristics that seem to be getting at the same underlying concerns. When several such characteristics are triggered in the same case, offense levels may quickly mount out of proportion to the real blameworthiness of a defendant’s conduct. Thus, for instance, in the robbery guideline, the weapons and abduction characteristics both seem to be getting at fear and danger caused by the defendant’s conduct, and they will often be found together—presumably, few abductions occur without the defendant at least brandishing a weapon of some kind. When both characteristics are present—for instance, when a gun-toting robber briefly abducts a victim (without causing any bodily injury)—the total enhancement may

amount to ten levels. By comparison, this is considerably more than the six-level enhancement for causing a *life-threatening* bodily injury. Similarly, I have shown elsewhere how a minor spill of nontoxic chemicals may trigger several overlapping characteristics in the environmental sentencing guideline such that the spiller may receive a longer sentence than other environmental criminals who put lives in imminent danger.³²

Yet, there is a more fundamental weighting problem embodied in the Guidelines’ version of uniformity: its insistence on national uniformity, without regard to varying community norms across the country. After all, there is no objective, universal yardstick to tell us how many months in prison a death threat is worth, or the discharge of a firearm, or a serious bodily injury. The moral significance of such offense characteristics cannot be meaningfully assessed without reference to community norms, and some communities will surely weight some offense characteristics quite differently than will other communities. For instance, as Judge Broderick once observed in these pages, gun possession means something quite different in the hunting cultures of western states than it does in violence-weary inner cities in the East.³³ A national-average weighting scheme is almost certainly getting the weights quite wrong in a great many districts. Indeed, Congress actually contemplated this possibility and authorized the Commission to incorporate regional variation into the Guidelines.³⁴ While the Commission did not take advantage of this authority, regional variation should not be considered inconsistent with Congress’s vision of uniformity.

E. Complexity

The Guidelines’ notorious technical complexity undermines their uniformity objectives in at least two respects. First, as Professors Ruback and Wroblewski have demonstrated, the sheer number of sentencing factors recognized by the Guidelines diminishes the reliability of sentencing decisions, that is, the tendency of the process to produce the same result if repeated.³⁵ Sentencing factors tend to be defined in the Guidelines such that reasonable people may differ as to how they should be applied in particular cases. Think, for instance, of the distinctions made in the robbery guideline between “bodily injury” and “serious bodily injury.” The presence of many such factors can have a dramatic cumulative effect on overall reliability.³⁶ Empirical studies of Guidelines application support this concern, showing, for instance, that different probation officers given the same set of facts come up with quite different offense level calculations.³⁷

Second, complexity diminishes the motivation of judges.³⁸ Research from organizational psychology demonstrates that “professionals do not like to be told how to do their jobs, particularly by nonprofessionals,” often causing detailed rules and procedures to be ignored by those charged with implementing them.³⁹ As Ruback and Wroblewski argue, this dynamic of resentment has played an important role in federal sentencing: “Judges dislike

their supervisors (i.e., Congress and the Sentencing Commission) specifying procedures for them as if they were hourly workers, not professionals.”⁴⁰ There should be little wonder that studies found circumvention of the Guidelines—i.e., the manipulation of fact-finding and charges of conviction to avoid imposition of a sentence required by the Guidelines—to be common,⁴¹ while judges were said to be in a state of “rebellion” against the drug guidelines.⁴²

In light of these reliability and motivation concerns, there is good reason to doubt whether the system as a whole has done—or is capable of doing—a good job *in practice* of ensuring similar sentences for the similarly situated and different sentences for the differently situated. These practical concerns would be just as strong even if the Guidelines somehow contained all of the right offense characteristics and weighted them all appropriately. Indeed, there may be something of a Catch-22 here: the better the Guidelines succeed in specifically identifying and weighting all factors relevant to uniformity, the more complex the Guidelines will grow—and the less likely that the Guidelines will result in consistent outcomes in practice.

F. Prosecutorial Control

Prosecutors retain vast plea-bargaining and charging discretion under the Guidelines. To be sure, the drafters of the Guidelines sought to constrain the significance of that discretion through a real-offense sentencing system, which theoretically diminishes the importance of the charges of conviction. However, the drafters apparently did not fully appreciate the extent of prosecutors’ ability to undermine the real-offense system through such mechanisms as fact bargaining and substantial assistance departures.⁴³ On the whole, judges have shown little willingness (or perhaps capacity) to insist on real-offense sentences in the face of a competing prosecutorial agenda.⁴⁴ Thus, in practice, the effectiveness of the Guidelines uniformity scheme is limited by the motivation of prosecutors to implement that scheme faithfully. However, prosecutors must balance fidelity to the Guidelines against a host of competing interests, such as speedy case processing. Studies of Guidelines circumvention tend to support suspicions that prosecutors have limited allegiance to the uniformity principle.⁴⁵

G. Summary

The Guidelines do not embody a robust system of sentencing uniformity, but, to borrow Breyer’s own characterization, a set of “compromises.”⁴⁶ Uniformity has been merely one of numerous competing considerations of public policy and political expedience.

The Commission’s compromises may or may not have been good ones. I do not mean to suggest here that the Commission *ought* to have embraced strict uniformity as the dominant consideration in the Guidelines sentencing process, which would have entailed, among other things,

grappling with the extraordinarily difficult challenge of prosecutorial discretion. Indeed, in light of the Catch-22 of complexity, uniformity in any truly rigorous form is probably unattainable.

In any event, the point is that questions about the desirability of judicial adherence to the Guidelines *as they exist* cannot usefully be equated with questions about the desirability of uniformity in sentencing as an abstract proposition.

Consider the remedy dispute in *Booker* itself. Assume, for the sake of argument, that Breyer is correct in one of his central premises: under a voluntary Guidelines system, sentences will more frequently be enhanced on the basis of applicable Guidelines factors than they would be under the dissenters’ jury-based system. Even if this claim were true, the result would not necessarily be greater system-wide uniformity. The majority’s remedy will likely produce fewer false negatives (i.e., treating differently situated defendants the same), but it may produce more false positives, due, for instance, to the Guidelines’ reliance on more or less arbitrary cliffs and other dubious sentencing factors.⁴⁷ It is not clear, *a priori*, which effect (false negatives under the dissenters’ remedy or false positives under the majority’s) would be more pronounced. (Nor is it clear, in a criminal justice system that otherwise seeks to err in favor of the defendant, that false negatives should be regarded as an equally serious problem as false positives.) In the end, it seems to me that a preference for uniformity does not really answer the remedy question; the question must instead be answered by reference to other considerations, such as the likely transaction costs of the jury system or the contribution of jury decision making to the perceived legitimacy of sentencing outcomes.

III. Implementing *Booker*

Viewing the Guidelines as an ad hoc set of compromises, rather than a coherent uniformity scheme, may help to illuminate some of the important questions that are being raised as lower courts try to implement *Booker*. In particular, this part will address the debate between Judges Adelman and Cassell as to the weight that should be given the Guidelines in the new “advisory” regime.

In *United States v. Ranum*, Judge Adelman held that “under *Booker*, courts must treat the guidelines as just one of a number of sentencing factors.”⁴⁸ In particular, Adelman held that sentencing judges must now weigh all of the sentencing purposes and considerations identified in 18 U.S.C. § 3553(a), which are sometimes in conflict with the Guidelines.⁴⁹ Judges, he suggested, must resolve those conflicts on a case-by-case basis, sometimes rejecting the Guidelines, “so long as the ultimate sentence is reasonable and carefully supported by reasons tied to the § 3553(a) factors.”⁵⁰

Judge Cassell rejected this approach in *United States v. Wilson*, holding that variances from the Guidelines should be “rare,” i.e., “only in unusual cases for clearly identified and persuasive reasons.”⁵¹ In particular, Cassell objected

to Adelman's reliance on offender characteristics at sentencing.⁵²

While the Cassell-Adelman debate touches on a number of points, the most important for present purposes is Cassell's claim that "[h]eavy reliance on the Guidelines is the only way to implement the congressional directive [contained in § 3553(a)] for courts to 'avoid unwarranted sentencing disparities among defendants with similar records who have been found guilty of similar conduct.'"⁵³ Cassell argued

If each judge follows his or her own views of 'just punishment' and 'adequate deterrence,' the result will be a system in which prison terms will depend on what the judge ate for breakfast on the day of sentencing and other irrelevant factors. . . .

. . . The result will almost inevitably be that defendants sentenced in the Eastern District of Wisconsin will serve different sentences for the same offense than similarly-situated defendants sentenced in the District of Utah.⁵⁴

This reasoning suffers from the same flaw as that of the remedy majority in *Booker*, that is, equating Guidelines adherence with uniformity. Yet, as we have seen, the Guidelines distinguish among defendants based on a host of considerations that have little bearing on the relative seriousness of their records or blameworthiness of their conduct. Moreover, even if adherence to the Guidelines does diminish the importance of what the judge had for breakfast on the day of sentencing, it may actually *enhance* the importance of what the *prosecutor* had for breakfast on the day of plea bargaining. (Indeed, in a Guidelines regime that embodies not a scientific system of uniform punishment but a series of ad hoc compromises, one wonders about the digestive influences that have operated over the years on key Commissioners and members of Congress.)

Nor is there any reason to believe that the Cassell test will do anything to ensure that sentences in Wisconsin are the same as sentences in Utah. Sentences varied considerably from district to district even before *Booker*,⁵⁵ and it is hard to see how those patterns will change under *any* version of the *Booker* regime. Such variation is a predictable result of the Guidelines' complexity and (in the absence of rigorous, binding, nationwide standards for charging and plea bargaining) the exercise of prosecutorial discretion. Indeed, it is not even clear that such variation is a bad thing; variation may reflect legitimate differences in community values and crime-control needs.⁵⁶

In sum, while there may be a number of compelling reasons for sentencing judges to give "great weight" to the Guidelines, a desire to "avoid unwarranted sentencing disparities among defendants with similar records who have been found guilty of similar conduct" should not be counted among them. Quite the contrary, variances from the Guidelines may be used appropriately in some cases to

soften a number of the disparities that are woven into the fabric of the Guidelines themselves, such as the crack-cocaine disparity, the implicit trial penalty of § 3E1.1, and the endemic phenomena of cliffs and double-counting.

IV. Legislative Responses: Restructuring the Guidelines After *Booker*

If the pre-*Booker* Guidelines regime were viewed as a functional, coherent system of sentencing uniformity, then there might be a strong case for a legislative response to *Booker* that would effectively restore the old regime. On the other hand, viewing the old regime as a set of ad hoc compromises diminishes the immediate appeal of restoration. Thus, I share the view of many organizations and commentators that the Commission and Congress should take a wait-and-see approach, carefully evaluating the evolution of the new regime over a period of time.⁵⁷

That said, the discussion of uniformity in this article leads me to reiterate and expand upon a proposal I have made elsewhere for restructuring the Guidelines.⁵⁸ *Booker* offers a unique opportunity for Congress and the Commission to revisit the most fundamental structural decisions that were made in crafting the Guidelines. One such decision that has proven flawed was the decision to use a "piecemeal" approach to evaluating offense severity, that is, enhancing sentences on the basis of a large number of narrow, decontextualized, and nominally objective offense characteristics (e.g., did the defendant have a gun, was it discharged, was anyone hurt, how badly, how much money was taken, was a bank targeted).

This piecemeal approach suffers many defects. It tends to exclude softer, but still enormously important, variables, such as mens rea. It discourages consideration and discussion of the moral significance of the offense characteristics that do count, thus draining the sentencing process of any meaning for the participants apart from the bottom-line pronouncement of a sentence length. Moreover, because the Commission cannot possibly anticipate all of the fact patterns to which a guideline will be applied, the piecemeal approach inevitably leads to the overcounting or undercounting of particular variables in particular cases.

Due to its complexity, the piecemeal approach engenders reliability and motivation problems. Its transparency invites manipulation by police and prosecutors (e.g., through fact bargaining or "sentencing entrapment"⁵⁹). Its apparent lack of integrity invites ad hoc congressional tinkering.

As against the piecemeal approach, I have suggested a more open-ended inquiry into such broadly defined sentencing factors as dangerousness and mens rea.⁶⁰ Restructured this way, the Guidelines would assign different weights on an offense-specific basis to the varying degrees to which the sentencing factors are present (e.g., extreme, high, moderate, and low danger created by the defendant's conduct, relative to other defendants who have committed the same type of offense). The Guidelines

would offer several specific examples of fact patterns for each category, but there would necessarily be a certain amount of subjectivity in the determinations, at least at the margins. Additionally, given the breadth of the sentencing factors (e.g., “low danger”), some dissimilar conduct would necessarily be lumped together.

In light of the subjectivity at the margins, as well as the use of a smaller number of broader categories within each offense type, we would effectively abandon the goal of rigorous, nationwide uniformity. On the other hand, it is not clear that rigorous, nationwide uniformity can be achieved in practice. Nor is it clear that nationwide uniformity (as against a community-based model) is even an appropriate goal.

We would gain a more meaningful sentencing process, as judges (or potentially juries) would be effectively invited to contextualize and assess the true significance of relevant offense characteristics. In so doing, we would retain a weaker, less demanding form of uniformity. This is not the same thing as “what the judge had for breakfast.” We would still have guidelines that make real distinctions among defendants, at least where the differences in the blameworthiness of their conduct are obvious and important. While there would be some imprecision around the borders, there would also be relatively clear cases in the heartland of each category.

V. Conclusion

We have a long-standing tradition of subordinating sentencing procedure to sentencing substance. Thus, for instance, in its 1949 decision in *Williams v. New York*, the Supreme Court approved informal sentencing procedures in order to advance the prevailing substantive sentencing goal of that era, rehabilitation.⁶¹ Likewise, the remedy opinion in *Booker* rejected jury-based sentencing because jury procedures would be inconsistent with a prevailing objective of this era, outcome uniformity.

By contrast, the *Booker* merits opinion seems to treat procedural justice at sentencing as an end in its own right, not merely as something to be grudgingly doled out just until it begins to conflict with substantive aims.⁶² “We will do the right thing with procedure,” the merits majority seems to be saying, “and let the substantive chips fall where they may.” This sort of emphasis on procedural justice has influenced the Court’s Eighth Amendment capital sentencing jurisprudence,⁶³ but, until *Apprendi*, it rarely played much of a role in noncapital cases.

We should welcome this renewed appreciation for procedural justice. At the same time, we should recognize that the vision of procedural justice embodied in *Apprendi* and its progeny is a rather parsimonious one, focusing on jury trial rights for fact-finding in certain limited circumstances. In ways that have not yet been well explained by the Court, these rights seem to arise from concerns about individual liberty in the face of state power and the democratic legitimacy of sentencing decisions.⁶⁴ Faintly visible in the Court’s reasoning may be a broader imperative that sentencing procedures embody respect for the essential

human dignity of criminal offenders. This sort of dignitary agenda has been more manifest in capital cases,⁶⁵ where the Court has also appropriately emphasized that sentencing procedures should affirm the individual worth of victims.⁶⁶

In any event, the Court’s growing interest in procedural justice at sentencing invites fresh thinking about the subject. I would suggest that the analysis begin with this question: What procedures will assure both offenders and victims that they have had a fair chance to tell their sides of the story to an attentive decision maker who actually cares about what they have to say?⁶⁷ In evaluating potential legislative responses to *Booker*, for instance, I think it would be helpful to know how offenders and victims perceive the fairness of sentencing juries. To the extent that juries can contribute to perceptions of fairness at sentencing, they ought to receive serious consideration. Of course, this and other innovations become more feasible and attractive once we recognize that uniformity is not a categorical imperative at sentencing, but only one goal to be pursued along with many others, including procedural justice.

Notes

¹ *United States v. Booker*, 125 S. Ct. 738, 761 (2005).

² *Id.* at 760–61.

³ *Id.*

⁴ Recall that the term “uniformity” is here used to encompass not just different treatment for the differently situated but also similar treatment for the similarly situated. See *infra* Part I.

⁵ This is clear, for instance, in Breyer’s discussion of the contrasting Smith/Jones and Johnson/Jackson hypotheticals. See *infra* text accompanying notes 7–9.

⁶ See, e.g., Stephen Breyer, *The Federal Sentencing Guidelines and the Key Compromises Upon Which They Rest*, 17 HOFSTRA L. REV. 1, 13 (1988).

⁷ 125 S. Ct. at 760.

⁸ *Id.* at 761.

⁹ *Id.*

¹⁰ 18 U.S.C. § 3553(a).

¹¹ See *Ewing v. California*, 538 U.S. 11, 36–37 (2003) (Breyer, J., dissenting) (agreeing with use of proportionality analysis for determining constitutionality of prison sentences and identifying key variables as real-offense conduct and circumstances, criminal history, and length of prison term).

¹² See Michael M. O’Hear, *Sentencing the Green-Collar Offender: Punishment, Culpability, and Environmental Crime*, 95 J. CRIM. L. & CRIMINOLOGY 133, 154–59 (2004).

¹³ See, e.g., Paul H. Robinson & John M. Darley, *The Utility of Desert*, 91 NW. U.L. REV. 453, 492–493 (1997) (arguing that weight of prior convictions should be, at most, “rather modest” when determining desert).

¹⁴ Breyer, *supra* note 6, at 17–18.

¹⁵ *Id.* at 13–14.

¹⁶ *Id.* at 18–20.

¹⁷ *Harris v. United States*, 536 U.S. 545, 570 (2002) (Breyer, J., concurring) (“Mandatory minimum statutes are fundamentally inconsistent with Congress’ simultaneous effort to create a fair, honest, and rational sentencing system through the use of Sentencing Guidelines.”).

¹⁸ See UNITED STATES SENTENCING COMMISSION, SPECIAL REPORT TO CONGRESS: COCAINE AND FEDERAL SENTENCING POLICY 1 (1997).

- ¹⁹ See Frank O. Bowman, III, *Pour Encourager Les Autres?: The Curious History and Distressing Implications of the Criminal Provisions of the Sarbanes-Oxley Act and the Sentencing Guidelines Amendments That Followed*, 1 OHIO ST. J. CRIM. L. 373, 376 (2004).
- ²⁰ PROTECT Act, Pub. L. No. 108-21, §§ 401(a), (i), 117 Stat. 650 (2003).
- ²¹ See O'Hear, *supra* note 9, at 220-222 (describing "missing culpability factors" in environmental guidelines); Douglas A. Berman, *The Model Penal Code Second: Might "Film Schools" Be in Need of a Remake?*, 1 OHIO ST. J. CRIM. L. 163, 166 (2003) (noting general criticism that guidelines downplay mens rea considerations).
- ²² See O'Hear, *supra* note 12, at 225-30, 266-67.
- ²³ The relevant extortion guideline would appear to be U.S.S.G. § 2B3.2.
- ²⁴ The relevant robbery guideline would appear to be U.S.S.G. § 2B3.1.
- ²⁵ In 2002, fewer than 1 percent of cases involved an upward departure. U.S. SENTENCING COMM'N, 2002 SOURCEBOOK OF FEDERAL SENTENCING STATISTICS 53 (2004).
- ²⁶ From the standpoint of uniformity, this residual discretion contemplated by the Guidelines is a two-edged sword: while it may be used to bring more aspects of the offender's conduct to bear at sentencing, it may also be used to dilute the significance of conduct and record by authorizing the consideration of offender characteristics and other considerations that do not relate to uniformity.
- ²⁷ See 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, 1539 (1997).
- ²⁸ I assume here, for reasons discussed in Part I above, a definition of uniformity based exclusively on record and offense conduct. Admittedly, a consistently applied plea discount might arguably fall within the parameters of a broader vision of uniformity, in which sentences are appropriately adjusted on the basis of any consideration that relates to any legitimate purpose of punishment. A plea discount, of course, advances the purpose of encouraging guilty pleas. One may, of course, question whether the encouragement of guilty pleas is a legitimate purpose of punishment. Even accepting this premise, however, the acceptance provision runs into difficulties from a uniformity standpoint. First, while it may operate as such in most cases, the acceptance provision is *not*, strictly speaking, a plea discount; rather, it is a vaguely worded sentencing adjustment that embodies conflicting purposes and has been subject to conflicting judicial interpretations. *Id.* at 1510-12. Second, to the extent that the acceptance provision does embody a plea discount, the operation of this plea discount in practice is subject to inconsistencies among prosecutors in plea-bargaining practices. Indeed, the Feeney Amendment recently enhanced the importance of prosecutorial discretion in the implementation of § 3E1.1 by requiring a prosecutor's motion as a predicate to the award of a full three-point reduction.
- ²⁹ See, e.g., *United States v. Galvez-Barríos*, 2005 WL 323703 (E.D. Wis. 2005) (finding that fast-track programs create unwarranted disparities). See also U.S. SENTENCING COMM'N, FIFTEEN YEARS OF GUIDELINES SENTENCING: AN ASSESSMENT OF HOW WELL THE FEDERAL CRIMINAL JUSTICE SYSTEM IS ACHIEVING THE GOALS OF SENTENCING REFORM xiii (2004) (finding substantial assistance departures, among several Guidelines provisions thought to promote disparity, to be responsible for greatest amount of variation in sentence lengths).
- ³⁰ See, e.g., *United States v. Smith*, 2005 WL 549057 (E.D. Wis. 2005) (crack); *United States v. Galvez-Barríos*, 2005 WL 323703 (E.D. Wis. 2005) (illegal reentry).
- ³¹ In 2002, about 60 percent of sentences were at the Guidelines minimum, while only about 10 percent were at the Guidelines maximum. U.S. SENTENCING COMM'N, *supra* note 25, at 59.
- ³² O'Hear, *supra* note 12, at 226.
- ³³ Vincent L. Broderick, *Local Factors in Sentencing*, 5 FED. SENT. REP. 314, 314 (1993).
- ³⁴ Michael M. O'Hear, *National Uniformity / Local Uniformity: Reconsidering the Use of Departures to Reduce Federal-State Sentencing Disparities*, 87 IOWA L. REV. 721, 742-43 (2002).
- ³⁵ R. Barry Ruback & Jonathan Wroblewski, *The Federal Sentencing Guidelines: Psychological and Policy Reasons for Simplification*, 7 PSYCHOL. PUB. POL'Y & L. 739, 764-65 (2001).
- ³⁶ See *id.* at 765-66 ("For example, if each of 5 decisions independently leading to a sentence recommendation (e.g., use of a weapon, degree of injury) had a reliability of .90, the final decision arguably could have a reliability as low as (.90)⁵ or .59. A reliability of .90 is considered good, whereas a reliability of .59 is not.").
- ³⁷ *Id.* at 765 (discussing P. B. Lawrence & P. J. Hofer, *An Empirical Study of Relevant Conduct Guidelines*, 4 FED. SENT. REP. 330 (1992)).
- ³⁸ *Id.* at 768-69.
- ³⁹ *Id.* at 768.
- ⁴⁰ *Id.* at 768-69.
- ⁴¹ Stephen J. Schulhofer & Ilene H. Nagel, *Plea Negotiations Under the Federal Guidelines: Guideline Circumvention and Its Dynamics in the Post-Mistretta Period*, 91 NW. U. L. REV. 1284, 1285 (1997) (finding that Guidelines circumvented in 20-35 percent of cases).
- ⁴² Frank O. Bowman III & Michael Heise, *Quiet Rebellion II: An Empirical Analysis of Declining Federal Drug Sentences Including Data From the District Level*, 87 IOWA L. REV. 477, 479-480 (2002).
- ⁴³ See U.S. SENTENCING COMM'N, *supra* note 29, at xii ("[A] variety of evidence developed throughout the guidelines era suggest that the mechanisms and procedures designed to control disparity arising at the presentencing stages are not all working as intended and have not been adequate to fully achieve uniformity of sentencing.").
- ⁴⁴ See, e.g., *id.* at xiii (discussing nationwide survey of judges indicating that judges infrequently go behind plea agreements to examine underlying conduct).
- ⁴⁵ See *id.* at xiii.
- ⁴⁶ Breyer, *supra* note 6, at 2.
- ⁴⁷ Under an advisory system, of course, judges may mitigate the false-positive problem by deviating from the Guidelines on a case-by-case basis. It is unclear, however, how much freedom judges will have to do so in the new regime. Nor is it clear how willing judges will be to use whatever freedom they have to address these concerns.
- ⁴⁸ 2005 WL 161223, *1 (E.D. Wis. 2005).
- ⁴⁹ *Id.* at *2.
- ⁵⁰ *Id.*
- ⁵¹ 2005 WL 273168, *1 (D. Utah 2005).
- ⁵² *Id.* at *5.
- ⁵³ *Id.* at *2.
- ⁵⁴ *Id.* at *13-14.
- ⁵⁵ See, e.g., Bowman & Heise, *supra* note 42, at 556; U.S. SENTENCING COMM'N, *supra* note 29, at 98, 100-01 (discussing studies). But see *id.* at 101-02 (reporting new study concluding that district differences account for only 2.8 percent of variation).
- ⁵⁶ See Michael M. O'Hear, *Localization and Transparency in Sentencing: Reflections on the New Early Disposition Departure*, 27 HAMLINE L. REV. 358, 360-365 (2004).
- ⁵⁷ See, e.g., American Bar Association, Criminal Justice Section, Report to the House of Delegates: Recommendation 1 (2005).
- ⁵⁸ O'Hear, *supra* note 12, at 266-269.

- ⁵⁹ See Sandra Guerra, *The New Sentencing Entrapment and Sentence Manipulation Defenses*, 7 FED. SENT. REP. 181 (1995).
- ⁶⁰ O'Hear, *supra* note 12, at 265-72.
- ⁶¹ 337 U.S. 241, 246-251 (1949).
- ⁶² See, e.g., *United States v. Booker*, 125 S. Ct. 738, 755-56 (2005) (merits majority) ("We recognize . . . that in some cases jury factfinding may impair the most expedient and efficient sentencing of defendants. But the interest in fairness and reliability protected by the right to a jury trial . . . has always outweighed the interest in concluding trials quickly.").
- ⁶³ See, e.g., *Lockett v. Ohio*, 438 U.S. 586, 605 (1978) (plurality) (holding that capital defendants entitled to offer any mitigat-

ing evidence at sentencing in light of the "need for treating each defendant in a capital case with that degree of respect due the uniqueness of the individual").

- ⁶⁴ *Booker*, 125 S. Ct. at 752 (merits majority) (discussing jury trial right as based on desire to prevent government oppression and preserve civil and political liberties).
- ⁶⁵ See, e.g., *Lockett*, 438 U.S. at 605.
- ⁶⁶ See, e.g., *Payne v. Tennessee*, 501 U.S. 808, 823-24 (1991).
- ⁶⁷ My proposal for restructuring the Guidelines, which makes relevant a broader range of considerations that bear on just punishment, would be quite consistent with this approach to procedural justice.