

DISPARITY: THE NORMATIVE AND EMPIRICAL FAILURE OF THE FEDERAL GUIDELINES

Albert W. Alschuler*

INTRODUCTION.....	85
I. THE NORMATIVE NATURE OF DISPARITY	87
II. THE DISPARITIES THAT THE GUIDELINES CREATE.....	89
III. THE DISPARITIES THAT THE GUIDELINES WERE INTENDED TO CORRECT	95
A. <i>Judicial Variation</i>	96
B. <i>Geographic Variation</i>	100
C. <i>Race, Ethnicity, and Gender</i>	102
IV. UNDER THE RADAR: THE SOURCES OF DISPARITY.....	106
A. <i>Judges</i>	106
B. <i>Defense Attorneys</i>	111
C. <i>Probation Officers</i>	111
D. <i>Law Enforcement Officers</i>	112
E. <i>Prosecutors</i>	112
CONCLUSION	116

INTRODUCTION

When viewed from any coherent normative perspective, the Federal Sentencing Guidelines have failed to reduce disparity and probably have increased it. Even on paper, these Guidelines often fail to treat like offenders alike, and the Guidelines are worse in practice than on paper. The luck of the judicial draw appears to determine the sentences offenders serve as much as or more than it did before the Guidelines; the region of the country in which an offender is sentenced now makes a greater difference than it did before the Guidelines; and racial and gender disparities have increased.

Part I of this Article emphasizes that sentencing disparity is a partly

* Julius Kreeger Professor of Law and Criminology, University of Chicago. I am grateful for the perceptive comments of Welsh S. White.

normative rather than an entirely empirical concept. It shows how the U.S. Sentencing Commission's initial evaluation of the Guidelines neglected this fact, proclaiming the Guidelines a success simply because judges in the post-Guidelines period came closer to following them than judges did before there were guidelines to apply.

Part II considers the disparities created by the Guidelines. Guidelines principles that appear plausible in some situations may prove nonsensical in others. Moreover, the penalties set by the Sentencing Commission frequently fail to follow a coherent pattern.¹

Part III focuses on the kinds of disparities the Guidelines were designed to prevent—those resulting from the identity of the sentencing judge, the region of the country in which an offender is sentenced, and the offender's race, ethnicity, or gender. It examines the empirical evidence bearing on these questions, particularly that generated by the Sentencing Commission and its staff. As the Commission's studies show, geographic disparity, the unequal treatment of racial and ethnic groups, and disparities between the sentences of women and men have increased in the Guidelines era. The Commission maintains that the amount of disparity attributable to the identity of the sentencing judge has declined, but this claim is unconvincing. Although the

1. The Sentencing Commission has insisted for eighteen years that its initial "Guidelines for most crimes were based on past practices" as revealed by "detailed data drawn from more than 10,000 reports of offenders sentenced in 1985 and additional data from approximately 100,000 more federal convictions." 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 14-15 (2004) [hereinafter U.S. SENTENCING COMM'N, FIFTEEN-YEAR REPORT]. In fact, the Commission considered the average sentences only of offenders in each offense category who had been sentenced to prison, and prior to the Guidelines, more than 40% of all federal offenders received sentences other than imprisonment. Converting the average sentences of *imprisoned* offenders into the Guidelines sentences for *all* offenders would have produced an enormous increase in sentence severity. The Commission guarded against this danger only by "eyeballing" the percentage of offenders in each category who had been sentenced to prison and by reducing Guidelines sentences below the prior averages when this percentage was low. Determining how much to reduce the prior averages was simply a matter of guesswork. This flawed process ultimately produced substantial increases in federal sentences. It replicated prior sentence levels only for crimes so serious that nearly all offenders went to prison even before the Guidelines. See Albert W. Alschuler, *Departures and Plea Agreements Under the Sentencing Guidelines*, 117 F.R.D. 459, 467-68 (1988); see also U.S. SENTENCING COMM'N, SUPPLEMENTARY REPORT ON THE INITIAL SENTENCING GUIDELINES AND POLICY STATEMENTS 27-34 tbl.1(a) (1987); Dissenting View of Commissioner Paul H. Robinson to the Promulgation of Sentencing Guidelines by the United States Sentencing Commission, 5 & n.11 (1987), reprinted in 41 Crim. L. Rep. (BNA) 3174 (1987).

Since its promulgation of the initial Guidelines, the Commission has raised sentences far more often than it has reduced them. Every crime *du jour* appears to prompt an increase in punishment and a press release. See, e.g., Press Release, U.S. Sentencing Comm'n, Commission Tightens Requirements for Corporate and Ethics Programs (May 3, 2004) (proclaiming that the Commission's actions would "lead to a new era of corporate compliance"), <http://www.ussc.gov/PRESS/re10504.htm> (last visited Sept. 14, 2005).

Commission's figures show a small reduction of judge-created disparity in the sentences *initially* imposed, they indicate no reduction of disparity in the sentences offenders ultimately serve.

Prior to the Guidelines, the United States Parole Commission, an agency with guidelines of its own, determined the release dates of prisoners sentenced by judges throughout America. If this Commission succeeded in reducing interjudge disparity even moderately, it almost certainly achieved greater success than that now claimed for the Federal Sentencing Guidelines.² The Guidelines appear to have failed at every job they were designed to do.

Although the Guidelines' failure can be seen in the Commission's statistics, these statistics do no more than skim the surface of the inequalities the Guidelines permit and encourage. Judge-created disparities, for example, are less likely to appear as visible departures from the Guidelines or as differing sentences within authorized Guidelines ranges than as differing *applications* of Guidelines provisions. Researchers do not treat judicial disagreement about the factual and legal questions and the issues of characterization that arise in Guidelines application as "sentencing disparity." Moreover, the Guidelines have vastly increased the sentencing power of prosecutors while reducing the ability of judges to check this power. The final Part of this Article focuses on sentencing disparities that statistical analysis is unlikely to detect or measure—disparities "under the radar" produced by judges, defense attorneys, probation officers, law enforcement officers, and prosecutors.

I. THE NORMATIVE NATURE OF DISPARITY

What counts as sentencing disparity is inescapably normative.³ For many people, the archetype of unequal sentencing is "sentencing by lottery."⁴ In a system of punishment by lottery, however, every offender would be treated like every other who drew the same number. When we say that punishment by lottery is unequal or capricious, we mean that this practice is morally incoherent. Drawing the same number is not the kind of "likeness" we believe should matter. Equality requires the consistent application of a comprehensible normative principle or mix of principles to different cases.

For this reason, evidence that offenders who have committed the same

2. The Commission has data that could resolve this issue more clearly. *See infra* text accompanying note 56. Of course the Parole Commission did not reduce disparity in the decision *whether* to imprison comparable offenders. The Sentencing Commission addressed this issue by authorizing imprisonment for *all* federal offenders and by mandating it for the overwhelming majority of these offenders.

3. The word "disparity" can mean either inequality or difference. In this sentence, I use it to mean inequality. Inequality is another word for "unwarranted" disparity.

4. *See, e.g.*, Jane Barrett, *Sentencing Environmental Crimes Under the United States Sentencing Guidelines—A Sentencing Lottery*, 22 ENVTL. L. 1421 (1992).

crime receive more uniform sentences under a guidelines system than they would have without them does not establish that the guidelines have reduced disparity. Judges in the pre-guidelines period might not have sought to treat everyone who committed the same crime alike. They might have tried to treat offenders of equal moral culpability alike or offenders of equal dangerousness alike or offenders with equal rehabilitative prospects alike. If these judges consistently applied a coherent principle or mix of principles to their cases, researchers could not fairly conclude that the guidelines had reduced disparity. They could conclude only that the guidelines had applied a new set of sentencing principles.

The United States Sentencing Commission sometimes has neglected this fact. In a 1991 report mandated by Congress on the first four years of the Sentencing Guidelines' operation,⁵ the Commission compared cases in four offense categories in which sentences had been imposed in a pre-Guidelines period to what it regarded as comparable cases that arose after implementation of the Guidelines.⁶ The Commission matched cases on the basis of factors the Guidelines deemed relevant to sentencing, such as drug quantity. The Commission then announced that its Guidelines had substantially reduced variation in the sentences judges imposed. The Commission had put the rabbit in the hat, however, by matching cases on the basis of the same factors the Guidelines used to set sentences. The Commission had treated its own view of appropriate sentencing as the measure of equality.⁷

The Commission's study had other methodological flaws,⁸ and after

5. U.S. SENTENCING COMM'N, THE FEDERAL SENTENCING GUIDELINES: A REPORT ON THE OPERATION OF THE GUIDELINES SYSTEM AND SHORT-TERM IMPACTS ON DISPARITY IN SENTENCING, USE OF INCARCERATION, AND PROSECUTORIAL DISCRETION AND PLEA BARGAINING (1991).

6. The Guidelines were not implemented in all federal courts until the Supreme Court upheld their constitutionality in *Mistretta v. United States*, 488 U.S. 361 (1989). Before *Mistretta*, more than 200 federal district judges had held the Guidelines unconstitutional. See U.S. SENTENCING COMM'N, 1989 ANNUAL REPORT 11 (1990).

7. See James M. Anderson et al., *Measuring Interjudge Disparity: Before and After the Federal Sentencing Guidelines*, 42 J.L. & ECON. 271, 280 (1999) (offering essentially this criticism of the Sentencing Commission's report).

8. The Commission excluded cases in which defendants had cooperated with the government, thereby ensuring that a likely source of post-Guidelines disparity would not mar its conclusions. Under the Guidelines, a judge may not reduce a sentence on grounds of "substantial assistance" unless a prosecutor asks him to do so. Then, however, the judge must determine whether and by how much to depart without any guidance. See Patti B. Saris, *Below the Radar Screens: Have the Sentencing Guidelines Eliminated Disparity? One Judge's Perspective*, 30 SUFFOLK U. L. REV. 1027, 1044-53 (1999). Moreover, a judge in the pre-Guidelines period ordinarily determined only the maximum period the Parole Commission could hold an offender, and usually this period was substantially longer than the judge expected the offender to serve. The possibility of release on parole vanished with the Guidelines. Although the Guidelines generally increased the amount of time served by offenders, they often reduced the length of the sentences initially imposed. Greater numerical variation was to be expected in the pre-Guidelines sentences simply because they were

October 2005]

DISPARITY

89

reanalyzing some of the Commission's data, the General Accounting Office rejected the Commission's conclusion that the Guidelines had reduced disparity even in the sentences judges initially imposed.⁹

Unlike the Commission's later studies, its four-year evaluation considered the sentences offenders were expected to serve as well as the sentences judges imposed. In most of the Commission's comparison groups, it failed to show, even by its own measure, a statistically significant reduction of variation in the sentences offenders could expect to serve. The Parole Commission apparently had reduced disparity in the pre-Guidelines period as effectively as the Guidelines did thereafter. Although judges may not have followed the Sentencing Guidelines before they existed, the Parole Commission came surprisingly close to doing so.

II. THE DISPARITIES THAT THE GUIDELINES CREATE

Sentencing researchers typically announce that their goal is to study unwarranted disparity and that unwarranted disparity does not include the differences in punishment authorized by legislatures and sentencing commissions.¹⁰ It is of course a truism that legally authorized sentencing considerations are warranted by law (whether or not they are warranted by common sense), and the researchers have no desire to quarrel with legislatures and sentencing commissions about sentencing policy. From the researchers' perspective, the disparities created by sentencing guidelines are warranted simply because the sentencing commission has said they are.

This perspective, however, cuts off half the action. Adopting the viewpoint of a person of ordinary moral sensibilities rather than of the Sentencing Commission leads quickly to the conclusion that the Sentencing Guidelines have substituted new disparities for old ones. This perspective suggests in fact that the Guidelines have seriously aggravated the problem of disparity.¹¹ Consider, for example, *Chapman v. United States*,¹² in which a defendant maintained that guidelines designed to promote equality violated the constitutional requirement of equal protection.

At the time of the decision in *Chapman*, both the Federal Guidelines and federal mandatory minimum sentencing statutes determined an LSD dealer's sentence by weighing the "mixture or substance" containing the drug. A "hit"

longer. In addition, critics complained that the Sentencing Commission's sample was not large enough to permit meaningful conclusions. The Commission's tests of statistical significance, however, seemed to answer this last objection.

9. GEN. ACCOUNTING OFFICE, CENTRAL QUESTIONS REMAIN UNANSWERED (1992).

10. See, e.g., Paul J. Hofer et al., *The Effect of the Federal Sentencing Guidelines on Inter-Judge Sentencing Disparity*, 90 J. CRIM. L. & CRIMINOLOGY 239, 243-46 (1999).

11. One can reach these conclusions without measuring R-squareds and T-tests. As Yogi Berra noted, "You can observe a lot by just watching."

12. 500 U.S. 453 (1991).

of LSD impregnated in a sugar cube, however, weighs much more than the same hit in a square of blotter paper or a gelatin capsule. Under the Guidelines, a dealer who sold 100 grams of LSD in sugar cubes was sentenced three times more severely than one who sold the same quantity in blotter paper, seven times more severely than one who used gelatin, and nineteen times more severely than one who sold the LSD in pure form. Judge Richard Posner commented, “[T]o base punishment on the weight of the carrier medium makes about as much sense as basing punishment on the weight of the defendant.”¹³ Judges in the pre-Guidelines period may have been quirky, but determining how many years to imprison someone by weighing sugar cubes was madness.¹⁴ In *Chapman*, the Supreme Court construed the term “mixture or substance” to include sugar cubes and held that this construction did not call the Guidelines’ constitutionality into question.¹⁵

Although a post-*Chapman* amendment to the Sentencing Guidelines has eliminated some of the disparities created by weighing “carrier media,”¹⁶ disparities pervade the Guidelines. Mark Osler provides some illustrations:

Under [the Guidelines], a woman who holds just six grams of crack for her own use is assigned a higher offense level than someone who commits criminal sexual abuse of a minor, a man who commits negligent homicide by

13. *United States v. Marshall*, 908 F.2d 1312, 1333 (7th Cir. 1990) (Posner, J., dissenting), *aff’d sub nom.*, *Chapman v. United States*, 500 U.S. 453 (1991).

14. See Albert W. Alschuler, *The Failure of Sentencing Guidelines: A Plea for Less Aggregation*, 58 U. CHI. L. REV. 901, 920 (1991) [hereinafter Alschuler, *The Failure of Sentencing Guidelines*]. In the *United States Court of Appeals for the First Circuit*, a post-*Chapman* opinion by former Sentencing Commissioner and later Supreme Court Justice Stephen G. Breyer compounded the insanity. A cocaine importer had chemically bonded cocaine to the material of his suitcases, and the court held that his sentence should be determined by weighing the suitcases. *United States v. Mahecha-Onofre*, 936 F.2d 623 (1st Cir. 1991). The court, however, did agree to omit the weight of the suitcases’ metal fittings. This ruling might have led defense attorneys to advise their drug-dealer clients that, if DEA agents approached, they should never abandon their drugs in Boston Harbor. If they were ever to do so, federal courts might weigh all the waters of the ocean to determine their sentences.

Courts outside the First Circuit took a different view. The offender with the cocaine-fiberglass suitcases received a sentence at least four years longer because his case was heard in the First Circuit than he would have if the case had been heard in another court of appeals. See Thomas J. Meier, Comment, *A Proposal To Resolve the Interpretation of “Mixture or Substance” Under the Federal Sentencing Guidelines*, 84 J. CRIM. L. & CRIMINOLOGY 377, 402 (1993). The Sentencing Commission eventually took the position of the courts outside the First Circuit. See U.S. SENTENCING GUIDELINES MANUAL § 2D1.1, cmt. n.1 (2004).

15. Empiricists might have pronounced the Guidelines at issue in *Chapman* a great success because offenders with equal weights of mixtures and substances received more uniform sentences under the Guidelines than they would have without them.

16. In cases in which mandatory minimum sentencing statutes are inapplicable, the Guidelines now make LSD sentences dependent on the number of doses possessed or sold. U.S. SENTENCING GUIDELINES MANUAL § 2D1.1, drug quantity tbl.n.H (2004). The weight of the carrier medium still determines the mandatory minimum sentences in LSD cases, however, as well as the Guidelines sentences for most other drugs. *Id.* at n.A.

recklessly shaking a baby to death, a woman caught stealing six million dollars of public money, or the executive who orders employees to dump a truckload of toxic waste knowing that people may die as a result. The crack possessor, in fact, receives the same offense level as that applicable to those who finance terrorist organizations

[I]n child pornography cases, the sentence for an individual who sends a computer image of “virtual” child pornography, made without the use of actual children, would face a sentence twice as harsh as that allowed under the Guidelines for a defendant who actually rapes a child. . . .

[T]he unreported transfer of large amounts of cash is illegal because it may support other outlawed activities such as theft. But . . . the sentence for such a cash transfer, even in the complete absence of proof that there was any underlying illegal activity, leads to greater punishment than that allocated to the woman who actually steals the same amount of money. . . .

[T]he danger of illegally possessed firearms is clear—they may be used for violence. Under the Guidelines, however, the punishment for possessing the weapon . . . is often more severe than that for actually using a gun in a violent crime.¹⁷

Sentencing judges sometimes have departed from the Guidelines to redress the unequal or disproportionate treatment they mandate for codefendants and co-conspirators convicted of the same crime. Appellate courts have nearly always reversed these judges, declaring that disparity between codefendants is not a permissible reason for departure.¹⁸

The offenses most frequently prosecuted in the federal courts are drug crimes (40.4% of the caseload¹⁹) and economic crimes, including white-collar crimes (21.6%²⁰). In both of these offense categories, sentences are driven by measures that often allocate punishments arbitrarily.²¹

Sentencing reformers imagined that guidelines would take into account the

17. Mark Osler, *Indirect Harms and Proportionality: The Upside-Down World of Federal Sentencing*, 74 *Miss. L.J.* 1, 1-6 (2004) (footnotes and parenthetical specifications of offense levels omitted).

18. *See, e.g.*, *United States v. Thurston*, 358 F.3d 51, 73 (1st Cir. 2004); *United States v. Blackwell*, 127 F.3d 947, 951-52 (10th Cir. 1997); *United States v. Wogan*, 938 F.2d 1446, 1449 (1st Cir. 1991) (“[T]he thrust toward equalization of sentences within a single case is not, and by itself cannot be, ‘an aggravating or mitigating circumstance’ as that phrase was employed by Congress . . . and reiterated by the Sentencing Commission. . . .”); *United States v. Joyner*, 924 F.2d 454, 459-61 (2d Cir. 1991); *United States v. Changa*, 901 F.2d 741, 744 (9th Cir. 1990) (affirming a district court’s refusal to equalize the sentences of co-conspirators). *But see* *United States v. Ray*, 930 F.2d 1368, 1372-73 (9th Cir. 1990); *United States v. Nelson*, 918 F.2d 1268, 1273 (6th Cir. 1990); *United States v. Daly*, 883 F.2d 313, 319 (4th Cir. 1989).

19. U.S. SENTENCING COMM’N, FIFTEEN-YEAR REPORT, *supra* note 1, at vi.

20. *Id.* at 47 fig.2.4.

21. The perception of many lawyers and judges is that the Guidelines work more equitably in cases of violent crime, but violent crimes constitute less than 4% of the federal criminal caseload. *Id.* at 68.

same things that judges previously had taken into account but would do so in a more uniform way. The guidelines would rein in outliers, minimize the role of prejudice, and perhaps, if an expert commission saw the need, make some changes in sentencing policy. Otherwise they would leave sentencing pretty much unchanged.

This vision was probably hopeless from the beginning. It rested on the assumption that mental processes can be captured in an algorithm. As the reformers saw it, a commission would need only to determine the components of judges' sentencing decisions and give them the proper weight. Of course, because the commission might not think of everything, the judges should be allowed to depart in truly exceptional cases. It was as though reformers noticed that a reviewer named Ebert often gave more stars to a motion picture than a reviewer named Roeper. The reformers could create a commission to determine how motion picture reviewers made their decisions and prepare a grid to guide them.²²

The belief that detailed, mandatory guidelines could largely duplicate the process by which pre-Guidelines judges made sentencing decisions failed to consider how the human mind differs from a sentencing commission and how decisions made *ex ante* differ from those made *ex post*. In drafting general guidelines, the Sentencing Commission lists aggravators when doing so "seems like a good idea at the time." Of course a criminal should receive a harsher sentence if he employs a weapon, and of course he should receive a harsher sentence if he abducts his victim. Most criminals who abduct their victims use weapons, however, and the Commission may not notice that the two aggravators together produce a far greater enhancement than it has prescribed for a criminal who inflicts a life-threatening injury.²³ The Sentencing Commission cannot fully foresee how its sentencing factors will overlap, interact, and compare with one another. A well-functioning human being who simply assesses "desert" in one case after another is likely to avoid the Commission's moral errors and inequalities. Some algorithm may be hidden in this person's mind, but if so, it is too complex and elusive to be brought to the surface and written down. Moreover, a large number of Guideline factors can itself generate inequalities when judges differ in applying these factors and their differences cumulate rather than cancel each other out.²⁴ Finally, the belief that the Guidelines could largely replicate the process by which pre-Guidelines judges made their decisions overlooked the inability of language to capture recognized differences. Describing in general terms the appropriate

22. If the film-review analogy seems inapt because not much is at stake in film reviews (except for actors and movie moguls), consider the extent to which detailed, mandatory guidelines could solve the problem of faculty disparity in grading seminar papers.

23. See Michael O'Hear, *The Myth of Uniformity*, 17 FED. SENT'G REP. 249, 252 (2005).

24. See *id.*

influence of situational and personal characteristics on sentences is often impossible. Quantifying harms, however, seems easy. Just count the stolen dollars and weigh the drugs. The Federal Sentencing Guidelines became crime tariffs mostly because the Sentencing Commission found it easier to write them that way.

The quantity of unlawful drugs possessed by a dealer and the amount of money stolen by a thief usually are relevant to the sentences they should receive. A thief who reaches into a till, however, generally takes whatever is there, be it fifty dollars or five thousand. Moreover, when the Guidelines sweep in drugs and losses tenuously connected to the crimes of which offenders have been convicted, sentences driven by weights and measures can be especially arbitrary. A later Part of this Article discusses sentencing in white-collar crime cases.²⁵ The remainder of this Part discusses drug crimes.²⁶

A drug courier often does not know what drug is inside the package she carries, let alone how much of that drug there is. The courier may be a woman traveling with children who was recruited partly because she seemed unlikely to fit a drug-courier profile or to be searched thoroughly by a customs agent. After being apprehended by a diligent customs agent, this courier might be taken to court to play a game called "Sentencing Guidelines" or "Wheel of Fortune." In this game, the host opens the sealed packages of the courier-contestants and weighs their contents. The weights determine which contestants win twenty-year, all-expense-paid visits to Leavenworth, Kansas, and which receive lesser prizes.

To determine the sentence of a crop duster who knowingly sprayed a field of marihuana, a court now must count the plants in the field and multiply by 100 grams. The Guidelines declare, "In the case of an offense involving marihuana plants, treat each plant, regardless of sex, as equivalent to 100 G of marihuana."²⁷

What constitutes a plant, however, has been the subject of extensive litigation.²⁸ In one case, a marihuana grower with very bad timing made 502 small cuttings from larger plants shortly before his arrest.²⁹ To support his contention that the appropriate standard was "viability," he offered expert testimony that most of his cuttings would not have survived and that "[a]

25. See *infra* text accompanying notes 101-11.

26. Cf. JUDICIAL CONFERENCE OF THE UNITED STATES, 1995 ANNUAL REPORT OF THE JUDICIAL CONFERENCE OF THE UNITED STATES TO THE U.S. SENTENCING COMMISSION 2 ("The Judicial Conference . . . encourages the Commission to study the wisdom of drug-sentencing guidelines which are driven virtually exclusively by the quantity or weight of the drugs involved.").

27. U.S. SENTENCING GUIDELINES MANUAL § 2D1.1, drug quantity tbl.n.E (2004).

28. See, e.g., *United States v. Silvers*, 84 F.3d 1317 (10th Cir. 1996) (holding that plants need not be alive); *United States v. McMahon*, 935 F.2d 397 (1st Cir. 1991) (holding that unusable plants about to be discarded must be included).

29. *United States v. Eves*, 932 F.2d 856 (10th Cir. 1991).

cutting becomes a plant when it develops a root system sufficient to allow the cutting to maintain open stomas so that it can exchange gas and provide for energy requirements.”³⁰ The Tenth Circuit rejected the grower’s contention, declaring, “If a cutting has a root ball attached it will be considered a plant.”³¹ The Sentencing Commission later endorsed the Tenth Circuit standard,³² noting that “this issue arises frequently.”³³

In the cases of drug couriers and others, the Sentencing Guidelines may yield disparate sentences even when they direct a judge to weigh only the drugs the offender himself possessed at the time and place of his crime. The Guidelines, however, frequently require courts to weigh drugs possessed at other times and places and drugs possessed by other people.

In constructing the Guidelines, the Sentencing Commission used a principle it called “real-offense sentencing.”³⁴ This term had come into use as a matter-of-fact recognition that a legal system dependent on plea bargaining frequently fails to convict offenders of their real crimes. When a bargain has enabled an offender to avoid conviction for his most serious crimes, a judge must sentence him within the limits applicable to the less serious crimes to which he has pleaded guilty. Before the Guidelines, scholars and judges debated whether, in selecting the sentence for these crimes, the judge should take account of the offender’s “real” crimes.

The Sentencing Commission extended the concept of “real-offense sentencing” much more broadly—to everything it called “relevant conduct.” It defined relevant conduct to include, among other things, “all acts and omissions . . . that were part of the same course of conduct or common scheme or plan as the offense of conviction.”³⁵ Under this provision, courts often weigh drugs the prosecutor never charged as part of the offender’s “real offense” and even drugs the jury acquitted the offender of possessing.³⁶ Relevant conduct also includes “all reasonably foreseeable acts and omissions of others in furtherance of . . . jointly undertaken criminal activity.”³⁷

The elasticity of terms like “reasonably foreseeable,” “same course of conduct,” and “common scheme or plan” gives judges some practical discretion in deciding which drugs to weigh. The Sentencing Commission notes, “Evidence from field research suggests that . . . ambiguity in the rule, and reluctance to . . . subject defendants to . . . severe penalties . . . limits the rule’s

30. *Id.* at 858.

31. *Id.* at 860.

32. U.S. SENTENCING GUIDELINES MANUAL § 2D1.1, cmt. n.17 (2004).

33. *Id.* app. C, amend. 518.

34. See David Yellen, *Reforming the Federal Sentencing Guidelines’ Misguided Approach to Real-Offense Sentencing*, 58 STAN. L. REV. 267 (2005) (in this Issue).

35. U.S. SENTENCING GUIDELINES MANUAL § 1B1.3(a)(2) (2004).

36. See *United States v. Watts*, 519 U.S. 148 (1997).

37. U.S. SENTENCING GUIDELINES MANUAL § 1B1.3(a)(1)(B) (2004).

October 2005]

DISPARITY

95

application.”³⁸ When judges exercise their discretion in differing ways, unequal sentences occur. More troubling disparities arise, however, when there is no play in the rule and judges must apply it.

In San Francisco, a steadily employed, forty-nine-year-old dockworker with no criminal record accommodated a friend by driving him to a drug transaction. Because this friend sold more than fifty grams of crack cocaine, a federal judge was required to sentence the dockworker to ten years.³⁹ The next case before the judge might have involved an otherwise identical offender with a less high-rolling friend. In this case, the judge would have imposed a substantially smaller sentence.

A small-time street dealer was arrested in New York with two vials of crack. Then the dealer’s supplier was apprehended with 586 additional vials in his hat.⁴⁰ Another dealer’s supplier might have been arrested with only ten vials; a third dealer’s supplier might not have been arrested; and a fourth dealer’s supplier (or the supplier’s supplier) might have been found with a warehouse full of cocaine. To people of ordinary moral sensibilities, sentencing the four street dealers to very different prison terms would seem arbitrary. Federal court lawyers and judges, however, get used to it.

III. THE DISPARITIES THAT THE GUIDELINES WERE INTENDED TO CORRECT

Every assessment of sentencing disparity rests on a normative judgment, and there are no uncontroversial criteria for sentencing. There is, however, a useful way for dispassionate social scientists to study sentencing disparity. Rather than accept the Guidelines’ own standards as a baseline, researchers can assess the influence on sentencing of clearly *inappropriate* circumstances. Federal sentencing reformers emphasized three things that, in their view, should not determine sentences—the identity of the sentencing judge, the region of the country in which an offender is sentenced, and the offender’s race, ethnicity, or gender.

38. U.S. SENTENCING COMM’N, FIFTEEN-YEAR REPORT, *supra* note 1, at 144. Paul Hofer, the Sentencing Commission’s Senior Research Associate, and two co-authors add:

Some judges may . . . adopt a result-oriented approach that begins with the sentence they wish to impose and works backwards to identify the facts leading to that result. It may not even be necessary to ignore facts. Many findings required by the guidelines are sufficiently subjective to afford significant discretion to those who wish to use it. . . . The relevant conduct guideline, on which all other guideline calculations rest, is notoriously complicated and subject to differing interpretations.

Hofer et al., *supra* note 10, at 258-59.

39. See *New Drug Law Leaves No Room for Mercy*, CHI. TRIB., Oct. 5, 1989, § 1A, at 28.

40. See *United States v. Joyner*, 924 F.2d 454 (2d Cir. 1991).

A. Judicial Variation

The Sentencing Commission notes, “The legislative history of the [Sentencing Reform Act of 1984] clearly shows . . . that different treatment by different judges was the chief problem the Act was designed to address, as well as regional differences in sentencing.”⁴¹

The first post-Guidelines study to focus on how much sentencing outcomes depended on the identity of the sentencing judge is also the study whose methodology is the easiest for nonprofessionals to understand. Joel Waldfogel examined sentences in three federal districts in which cases were assigned randomly to judges and every judge appeared to have a comparable mix of cases over time.⁴² Waldfogel determined the length of the average sentence imposed in each district before and after the Guidelines and then how much each individual sentence deviated from the average. He found that the judges who imposed below-average sentences before the Guidelines imposed below-average sentences thereafter. The judges who formerly imposed above-average sentences still imposed them too.

Waldfogel averaged the amount by which all sentences in each court deviated from the court’s overall average to determine the “mean absolute deviation.” If the Sentencing Guidelines had reduced the influence of the judge’s identity on sentencing—that is, if they had brought judges closer together—this measure of disparity should have declined following implementation of the Guidelines. Waldfogel found, however, that judicial disparity increased significantly in two of the three districts studied and remained essentially unchanged in the third. In the District of Connecticut, the mean deviation increased from 4.2 months before the Guidelines to 9.9 months after; in the Southern District of New York, it rose from 5.8 to 10.4 months; and in the Northern District of California, it rose insignificantly from 4.2 to 4.4 months.⁴³

Critics of the Federal Sentencing Guidelines warned of the evils the Guidelines would produce, but none of them suggested (or imagined) that the Guidelines would *increase* interjudge disparity. Three later studies did report some reductions of disparity in the sentences judges imposed, but the reductions ranged from small to negligible.

41. U.S. SENTENCING COMM’N, FIFTEEN-YEAR REPORT, *supra* note 1, at 11.

42. Joel Waldfogel, *Aggregate Inter-Judge Disparity in Federal Sentencing: Evidence from Three Districts*, 4 FED. SENT’G REP. 151 (1991).

43. Because Waldfogel’s post-Guidelines period began before the Supreme Court’s decision in *Mistretta v. United States*, 488 U.S. 381 (1989), his post-Guidelines figures might have included some cases not sentenced under the Guidelines. *See supra* note 6. Paul Hofer and his co-authors also criticized Waldfogel for not taking account of “overall changes in the lengths of sentences for some types of crimes under the guidelines.” Hofer et al., *supra* note 10, at 279. Without these changes, however, the increase in disparity might have been greater. *See supra* note 8.

Abigail Payne examined interjudge disparity in three federal district courts. She concluded that the disparity attributable to differences among the judges before the Guidelines was small—less than 5% of the total variation in sentences.⁴⁴ A consistent finding of the post-Guideline studies has been that interjudge variation before the Guidelines was much smaller than sentencing reformers evidently believed it to be. The sentencing reformers relied in significant part on differences in judicial responses to hypothetical cases, and real-world sentencing is apparently different.⁴⁵ Although the level of interjudge disparity declined in two of the districts studied, Payne described the decline as negligible.⁴⁶

A distinctive feature of Payne's study was its effort to distinguish the influence of the Sentencing Guidelines from that of mandatory minimum sentencing legislation that became effective at about the same time. Like the Sentencing Guidelines, mandatory minimum sentences can reduce disparity—for example, by requiring all judges to impose more severe sentences than any of the judges would have chosen if allowed to consider cases on their merits. Payne studied drug crimes and property crimes separately because, although the Sentencing Guidelines applied to both sorts of crime, the mandatory minimums applied only to the drug crimes. In two of her three districts, Payne found no reduction of disparity in property-case sentences. She inferred that the overall reduction of disparity was attributable mostly to the mandatory minimums.

A public defender, an economist, and a prominent critic of the Federal Sentencing Guidelines conducted the study most supportive of the hypothesis that the Guidelines had reduced interjudge disparity.⁴⁷ These authors reported

44. A. Abigail Payne, *Does Inter-Judge Disparity Really Matter? An Analysis of the Effects of Sentencing Reforms in Three Federal District Courts*, 17 INT'L REV. L. & ECON. 337, 338 (1997).

45. See, e.g., Joel Waldfogel, *Does Inter-Judge Disparity Justify Empirically Based Sentencing Guidelines?*, 18 INT'L REV. L. & ECON. 293 (1998); see also Hofer et al., *supra* note 10, at 296. Sentencing reformers regularly cited the study of differing judicial responses to case reports described in ANTHONY PARTRIDGE & WILLIAM B. ELDRIDGE, *THE SECOND CIRCUIT SENTENCING STUDY: A REPORT TO THE JUDGES OF THE SECOND CIRCUIT* (1974). This study showed that judges do disagree substantially about some cases.

46. The Sentencing Commission's Senior Research Associate and his co-authors objected that "like Waldfogel, Payne used data from a transitional period, which included many non-guideline cases mixed with cases sentenced under the guidelines, thus diluting any effect of the guidelines." Hofer et al., *supra* note 10, at 281. This criticism was unfounded. In Payne's post-Guidelines cases, sentences were imposed between January 1, 1989, and December 31, 1991. Payne, *supra* note 44, at 345. Because the Supreme Court's decision in *Mistretta* came on January 18, 1989, only a very small number of non-Guidelines cases could have found their way into her post-Guidelines group.

47. James M. Anderson, Jeffrey R. Kling & Kate Stith, *Measuring Interjudge Sentencing Disparity: Before and After the Federal Sentencing Guidelines*, 42 J.L. & ECON. 271 (1999). Kate Stith's comprehensive criticism of the Guidelines appears in KATE STITH & JOSÉ A. CABRANES, *FEAR OF JUDGING: SENTENCING GUIDELINES IN THE FEDERAL COURTS* (1998).

that the expected difference in the sentence lengths of two judges receiving comparable cases was sixteen to eighteen percent in the pre-Guidelines period. This difference fell to somewhere between eight and thirteen percent after the Guidelines were implemented.⁴⁸ In temporal terms, the expected interjudge difference dropped from 4.9 months before the Guidelines to 3.9 months after.⁴⁹

The study conducted by Paul Hofer and other members of the Sentencing Commission's staff examined more courts and more cases than any other study. In the first part of this study, the Hofer group focused on the nine cities in which at least three judges had imposed sentences in both the pre-Guidelines and post-Guidelines periods. (These periods were separated by ten years.) The researchers reported that, before the Guidelines, only 2.32% of the variation in sentences was explained by the identity of the sentencing judge. This figure fell to an even smaller 1.24% in the post-Guidelines period, a reduction of 1.08%.⁵⁰ In the pre-Guidelines period, offenders could expect the identity of the sentencing judge to make a difference of about 7.87 months in their sentences. With the Guidelines in effect, the expected difference in sentence attributable to the judge fell to 7.61 months—0.26 months less.⁵¹ In other words, in the average case, the Guidelines might have reduced interjudge sentencing disparity by about a week.

A second part of the Hofer study focused on the forty-one cities that had three or more judges on the bench during both the pre-Guidelines and post-Guidelines periods (whether or not they were the same judges). In the pre-Guidelines period, the identity of the sentencing judge explained 2.40% of the variation in sentences. This figure fell to 1.64% in the post-Guidelines period—a 0.76% reduction.⁵²

In temporal terms, the variation among judges in the forty-one-city study increased in the post-Guidelines period—from 8.89 months before the Guidelines to 9.69 months after, an increase of approximately three weeks.⁵³ The explanation for this apparent anomaly (a larger temporal variation attributable to the judge but a smaller percentage of total variation attributable to the judge) is that the sentences judges imposed in the post-Guidelines period were more severe than the sentences judges imposed before the Guidelines. A larger number of months was therefore a smaller percentage of the total average sentence, and a larger numerical variation in months was also a smaller variation in the percentage of variation explained.

A judge who is more severe than his colleagues in white-collar crime cases

48. Anderson et al., *supra* note 47, at 303.

49. *Id.* at 294.

50. Hofer et al., *supra* note 10, at 287.

51. *Id.* at 287-88 & tbl.1.

52. *Id.* at 289.

53. *Id.* at 290 tbl.2.

may be more lenient in drug cases. The sentencing disparities resulting from this judge's idiosyncrasies could disappear in the analysis described above—an analysis focusing on what Hofer and his colleagues called the “primary judge effect.” The researchers therefore also sought to study what they called “offense type by judge interaction.” They reported, however, that “[t]he offense type by judge interaction fell in the forty-one-city analysis but actually increased in the nine-city analysis.”⁵⁴

The Hofer group did not attempt to separate the effects of the Sentencing Guidelines from those of mandatory minimum sentences. It seems likely, however, that most or all of the reported reduction in interjudge disparity was attributable to the mandatory minimums rather than the Guidelines. By far the largest reductions in interjudge disparity occurred in the two offense categories affected by the mandatory minimums—drugs and firearms. Reductions were smaller in fraud and larceny cases, and interjudge disparity actually increased substantially in immigration and robbery cases.⁵⁵

All of the post-Guidelines studies focused on the sentences judges imposed rather than the sentences offenders actually served. The reported reductions in disparity in the sentences imposed were small, and if the Parole Commission reduced disparity at all in the pre-Guidelines period, its leveling was likely to have matched or exceeded that accomplished by the Guidelines. The Sentencing Guidelines do not appear to have lessened disparity at all in the variable that matters most to offenders: how much time they must serve.

Despite its flaws, the Sentencing Commission's 1991 report did address whether the Guidelines had reduced disparity in the time offenders served. The Hofer group and the Commission's *Fifteen-Year Report* did not consider this question although the data were apparently at hand. The Hofer group reported that it had “conducted some analyses using as outcome the expected time to be served” and that it calculated the time to be served for the pre-Guidelines period by simulating the parole-release guidelines.⁵⁶ After this pregnant announcement, however, Hofer and his co-authors said no more.

A reader of the Sentencing Commission's *Fifteen-Year Report* would not learn that, according to one study by the Commission's staff, interjudge disparity decreased by only one week per case. This reader also would not learn that, according to another staff study, this disparity actually increased by three weeks per case. Instead, the reader would discover statements like this:

Rigorous statistical study both inside and outside the Commission confirm [sic] that the guidelines have succeeded at the job they were principally designed to do: reduce unwarranted disparity arising from differences among judges. . . . [T]he “primary judge effect” was reduced by approximately one third to one half with the implementation of the guidelines, and “interaction

54. *Id.* at 297.

55. *Id.* at 293 tbl.3.

56. *Id.* at 312; *see also id.* at 287 n.125.

effects” have been reduced even more substantially.⁵⁷

The Commission’s assertion to the contrary notwithstanding, the best judgment is that the Sentencing Guidelines have failed at the job they were principally designed to do. The small reduction in interjudge disparity reported by the Commission seems attributable mostly to the mandatory minimum sentences rather than the Guidelines, and the reduction accomplished by both devices together was probably less than the Parole Commission achieved in the pre-Guidelines period. The Sentencing Commission’s 258-box sentencing grid appears to have lived in vain.

B. Geographic Variation

Judge Patti Saris was a staff attorney for the Senate Judiciary Committee at the time of the Sentencing Reform Act. She recalls, “The rallying cry was: why should a bank robber in California get a different sentence from a bank robber in Texas?”⁵⁸ Congressional committees cited differences in the average sentences imposed for federal crimes in different federal districts.⁵⁹

Similarly arresting differences among federal districts exist today.⁶⁰ To revisit Judge Saris’s example, the average sentence for robbery in the Northern District of Texas in fiscal year 2002 (142 months) was almost twice what it was in the Central District of California (72 months). The robbery sentence in the Central District of Texas was more than twice the average robbery sentence in the Southern District of New York (62 months). The average drug-trafficking sentence in Manhattan (89 months) was twice the average drug-trafficking sentence in Brooklyn (44 months). The average drug-trafficking sentence in the Southern District of California was 23 months, one-third of what it was in the Northern District of California (74 months) and one-fourth of the average drug-trafficking sentence in the Central District of California (95 months).

Differences like these established the need for sentencing reform in 1984. They apparently do not prove anything today. The Sentencing Commission cautions against giving weight to uncontrolled comparisons because “regional variations do not necessarily indicate *unwarranted* disparity.”⁶¹ As the Commission emphasizes, the mix of cases within an offense category can differ from one district to the next.

57. U.S. SENTENCING COMM’N, FIFTEEN-YEAR REPORT, *supra* note 1, at 140.

58. Saris, *supra* note 8, at 1028.

59. *See, e.g.*, S. REP. NO. 98-225, at 41 & 45 tbl.1 (1984), *reprinted in* 1984 U.S.C.C.A.N. 3182, 3224, 3228 tbl.1.

60. The source of all the figures that follow is U.S. SENTENCING COMM’N, FEDERAL SENTENCING STATISTICS BY STATE, DISTRICT, AND CIRCUIT: OCTOBER 31, 2001, THROUGH SEPTEMBER 30, 2002, <http://www.ussc.gov/JUDPACK/JP2002.htm> (last visited Sept. 16, 2005).

61. U.S. SENTENCING COMM’N, FIFTEEN-YEAR REPORT, *supra* note 1, at 100 (emphasis in original).

In the same article in which they reported on interjudge disparity, the Commission's research staff presented findings on geographic variation. In the staff's nine-city study, this variation accounted for only 1.03% of the total variation in sentences prior to the Guidelines. This figure grew to 3.59% after the Guidelines went into effect.⁶² In the pre-Guidelines period, an offender could expect the jurisdiction in which his case arose to make a difference of about 5.24 months in his sentence. With the Guidelines in effect, the expected difference in sentence attributable to the jurisdiction in which a case arose increased to 12.94 months.⁶³

In the staff's forty-one-city study, the court in which a case was heard accounted for 1.81% of the total variation in sentences in the pre-Guidelines period. It accounted for 5.64% of this variation in the post-Guidelines period. In temporal terms, the expected variation per case was 7.70 months before the Guidelines and 17.97 months after.⁶⁴

The Sentencing Commission repeatedly characterized the reported reduction in interjudge disparity from 2.32 to 1.24% as a 50% reduction (which it almost was). By the same measure, geographic disparity more than tripled after implementation of the Guidelines.⁶⁵

The Commission's *Fifteen-Year Report* did not declare that geographic disparity had tripled under the Guidelines. On a page of the *Report* in which a boldface box proclaimed that "[t]he best and most recent statistical analyses indicate that the guidelines have significantly reduced interjudge disparity compared to the preguidelines era,"⁶⁶ the Commission noted delicately, "The available evidence suggests that regional disparity remains under the guidelines, and some evidence suggests it may even have increased among drug trafficking offenses."⁶⁷

The Hofer group's findings are subject to two interpretations. On the one hand, the numbers are small enough that one could shrug them off and proclaim the Guidelines a bust. The Guidelines do not seem to have made much difference in the things social science researchers can measure. On the other hand, the numbers are large enough that one could reasonably take them seriously. On this view, the reported increase in geographic disparity swamped the reported reduction in interjudge disparity so that the Guidelines were an even bigger bust. The most important lesson of the Hofer group's findings may not be that *geographic* disparity increased but rather that *disparity* increased. The actions of some participants in the criminal justice system—prosecutors,

62. Hofer et al., *supra* note 10, at 289.

63. *Id.* at 288 tbl.1.

64. *Id.* at 290 tbl.2.

65. To present variations in small numbers as large percentages may be a rhetorical trick, but it is the Sentencing Commission's trick.

66. U.S. SENTENCING COMM'N, FIFTEEN-YEAR REPORT, *supra* note 1, at 94.

67. *Id.*

judges, or whoever else really determined sentences—had caused more arbitrary variation in criminal sentences than occurred before the Guidelines.

When an increase in regional disparity is accompanied by a reduction in interjudge disparity, one might infer that differing judicial attitudes could not explain the increased geographic variation. The Hofer group's methodology, however, does not directly support this inference. Because Hofer and his colleagues measured judicial disparity only within jurisdictions, their findings do not preclude judicial responsibility for the increased geographic variation.

The best judgment, however, is that some reduction in judge-produced disparity was more than offset by an increase in prosecutor-produced disparity. The Hofer group's data were gathered before the introduction of "fast-track" programs that now generate huge geographical variations in drug and immigration sentences.⁶⁸ The increased regional variation that Hofer and his colleagues discovered, however, was attributable mostly to drug cases. They commented:

In addition to the discretion they have always employed to bring and dismiss charges or to make sentence recommendations, the last ten years have witnessed the creation of new tools by which prosecutors can control sentencing. These include mandatory minimum statutes that limit judges' discretion, motions for departure based on a defendant's substantial assistance in the prosecution of others, which are solely in the hands of prosecutors, and factual stipulations accompanying plea agreements, which under the guidelines have a direct and predictable impact on the guideline range applicable to a case. The mechanisms designed to regulate this discretion . . . may not be capable of preventing prosecutorial decisions from reintroducing unwarranted disparity.⁶⁹

C. Race, Ethnicity, and Gender

In the Sentencing Reform Act of 1984, Congress mandated that the Sentencing Guidelines be "entirely neutral as to the race, sex, national origin, creed, and socioeconomic status of offenders."⁷⁰ Before the Guidelines became effective, blacks and whites convicted in the federal courts were imprisoned at exactly the same rate (54%). Because nearly all defendants convicted of immigration violations were imprisoned and these defendants were overwhelmingly Latino, Latinos went to prison at a higher rate (69%). The average maximum sentences of the whites, Latinos, and blacks sent to prison were similar—50 months for whites, 52 months for Latinos, and 53 months for

68. Fast-track programs are described in text at *infra* notes 99-100.

69. Hofer et al., *supra* note 10, at 300. Stephanos Bibas's article in this Issue explores questions of regional disparity in greater depth and detail. See Stephanos Bibas, *Regulating Local Variations in Federal Sentencing*, 58 STAN. L. REV. 137 (2005) (in this Issue).

70. 28 U.S.C. § 994(d) (2005).

October 2005]

DISPARITY

103

blacks.⁷¹

The picture in the federal courts is different today: 78% of all white offenders, 79% of all black offenders, and 88% of all Latino offenders are imprisoned.⁷² The average sentences of the imprisoned whites and Latinos remain similar—46 months and 44 months respectively. The average sentence of blacks is 83 months.⁷³

On a Sentencing Commission graph showing the estimated time to be served by sentenced offenders over time, three lines depict the sentences of blacks, whites, and Latinos.⁷⁴ These lines start at the same level, but with the arrival of the Guidelines and new mandatory minimum sentencing legislation, the line for blacks begins a skyrocketing ascent. A racial gap appears at about the time most federal courts implemented the Guidelines and the mandatory minimum sentences. The gap widens abruptly as the Guidelines go into effect in the remaining courts. The line for black offenders reaches its high point in the mid-1990s with blacks imprisoned for periods twice as long as they were before the Guidelines. Then the line declines slightly. The lines showing the sentences of white and Latino offenders begin at the same level as the line for blacks but increase far less dramatically. The racial gap that arrived with the Guidelines has widened enormously.

This racial gap is largely, but not entirely, driven by legally relevant considerations and by one legally relevant consideration in particular. Eighty-one percent of the defendants prosecuted for crack cocaine offenses in the federal courts are black.⁷⁵ In the Anti-Drug Abuse Act of 1986, Congress distinguished for the first time between crack and powder cocaine and mandated more severe sentences for crack cocaine. It did so by requiring 100 times more powder cocaine than crack to trigger the same mandatory minimum sentences.⁷⁶

71. See Douglas C. McDonald & Kenneth E. Carlson, *Why Did Racial/Ethnic Sentencing Differentials in Federal District Courts Grow Larger Under the Guidelines?*, 6 FED. SENT'G REP. 223, 223 (1994).

72. BUREAU OF JUSTICE STATISTICS, SOURCEBOOK OF CRIMINAL JUSTICE STATISTICS 2002, at 419 tbl.5.20 (2004).

73. *Id.* at 420 tbl.5.21.

74. U.S. SENTENCING COMM'N, FIFTEEN-YEAR REPORT, *supra* note 1, at 116 tbl.4.2. Note that the sentences described in the preceding paragraph were maximum sentences. This paragraph discusses changes in the sentences that offenders were expected to serve.

75. *Id.* at 132.

76. 21 U.S.C. § 841(b)(1)(A)(iii)-(B)(iii) (2005). The two drugs are pharmacologically similar. Because crack is usually smoked rather than inhaled through the nose, however, it provides more mind-bangs per buck and is probably more addictive. Many of the beliefs about crack that were common in the 1980s now appear to be myths. For example, using crack during pregnancy poses no greater risk of producing birth defects than using other drugs, including powder. See U.S. SENTENCING COMM'N, FIFTEEN-YEAR REPORT, *supra* note 1, at 132. Some of the crack myths of the 1980s (including the belief that men crazed by the drug had superhuman strength) might have been influenced by views of the people thought to use it.

The U.S. Sentencing Commission treated Congress's mandatory minimum sentences as its starting point. Inserting the drug quantities specified by the mandatory minimums into a seventeen-level gradation of quantities, the Guidelines imposed many sentences above the mandatory minimums. Today "[a]bout 25 percent, or eighteen months, of the average expected prison time of 73 months for drug offenders . . . can be attributed to guideline increases above the mandatory minimum penalty levels."⁷⁷

Congress may not have intended its mandatory minimums to generate the range of penalties the Guidelines prescribed. Congress believed that the quantities for which it mandated five-year terms identified "managers of retail traffic" while those triggering ten-year terms identified "manufacturers or the heads of organizations."⁷⁸ Congress apparently had in mind a two-tier sentencing system.

Although decisions of the Sentencing Commission aggravated the effects of the crack/powder disparity, the Commission proposed legislation and submitted a Guidelines amendment in 1995 to eliminate the disparity. For the first time in the Guidelines' history, Congress and the President rejected a Guidelines amendment approved by the Commission.⁷⁹ Two weeks before President Clinton signed the legislation restoring the 1-to-100 crack/powder ratio,⁸⁰ he observed in a speech at the University of Texas:

[B]lacks are right to think something is terribly wrong . . . when almost one in three African American men in their 20s are either in jail, on parole, or otherwise under the supervision of the criminal justice system—nearly one in three. And that is a disproportionate percentage in comparison to the percentage of blacks who use drugs in our society. Now, I would like every white person here and in America to take a moment and think how he or she would feel if one in three white men were in similar circumstances.⁸¹

Since its unsuccessful effort in 1995, the Sentencing Commission has submitted two other proposals to reduce the crack/powder disparity, but Congress has not acted.⁸²

The racial gap in federal sentences cannot entirely be explained by the 1-to-100 crack/powder ratio and other legally relevant variables. The Sentencing Commission reported, for example, that although blacks constitute 48% of the offenders who appear to qualify for a mandatory firearms enhancement in drug

77. U.S. SENTENCING COMM'N, FIFTEEN-YEAR REPORT, *supra* note 1, at 54.

78. *Id.* at 48 (quoting H.R. REP. NO. 845, pt. 1, at 16-17 (1986)). At least with respect to crack, these beliefs were clearly erroneous. Five grams of crack, the amount needed to trigger a five-year mandatory minimum sentence, is "the amount a heavy user might consume in a weekend." *Id.* at 132.

79. See Pub. L. No. 104-38, 109 Stat. 334 (1995).

80. President Clinton's Statement on Signing Legislation Rejecting U.S. Sentencing Commission Recommendations, 2 PUB. PAPERS 1700 (Oct. 30, 1995).

81. President Clinton's Remarks at the University of Texas at Austin, 2 PUB. PAPERS 1600, 1602 (Oct. 16, 1995).

82. U.S. SENTENCING COMM'N, FIFTEEN-YEAR REPORT, *supra* note 1, at 51.

cases, they constitute 64% of the offenders who receive it.⁸³ In addition, prosecutors seek “substantial assistance” departures for blacks and Latinos less often than for whites, and this disparity persists when researchers do their best to control for legally relevant variables.⁸⁴ The disparity in substantial assistance departures may reflect the lesser ability of blacks and Latinos to provide information useful to prosecutors, the greater reluctance of blacks and Latinos to provide this information (because of their greater loyalty to co-offenders or their greater fear of reprisals), or the prosecutors’ racial favoritism. When minority defendants do receive substantial assistance departures, the departures they receive are smaller than those received by whites.⁸⁵

More generally, the Sentencing Commission reports that, after controlling for legally relevant variables, black men are 20% more likely to be imprisoned for drug offenses than are white men, and Latino men are 40% more likely. When sent to prison, the sentences of black and Latino men for drug crimes are likely to be about 10% (7 months) higher than those of whites.⁸⁶

A gender gap in federal sentences preceded the Guidelines. The time served by men in federal prisons before the Guidelines exceeded that served by women by about nine months or 50%. In the years since the Guidelines were implemented, the gender gap has grown. The time served by men increased 96% after the Guidelines while that served by women increased 75%. Men now serve 51 months on average and women 28.⁸⁷ The previous nine-month gender gap has grown to 23 months.

Unlike the growing racial gap in federal sentences, the increasing gender gap cannot be largely explained by statutory innovations like the crack/powder disparity or other legally relevant variables. The Sentencing Commission reported that, after controlling for relevant variables, men were twice as likely to be imprisoned for drug crimes as women. Prison sentences in drug cases and other cases were twenty-five to thirty percent longer for men. Women received more substantial downward departures.⁸⁸

The Sentencing Commission’s *Fifteen-Year Report* observes, “Part of the more lenient treatment may arise . . . from differences between the genders that are relevant to sentencing but not well captured by the available data.”⁸⁹ The *Report* mentions that women may take less-central roles in criminal activity than men and may be “more instrumental in raising their children than their

83. *Id.* at 90.

84. See LINDA DRAZGA MAXFIELD & JOHN H. KRAMER, SUBSTANTIAL ASSISTANCE: AN EMPIRICAL YARDSTICK GAUGING EQUITY IN CURRENT FEDERAL POLICY AND PRACTICE 13-14 & n.30 (1998) (indicating that blacks are 8% to 9% less likely than whites to receive substantial assistance departures and that Latinos are 7% less likely).

85. *Id.*

86. U.S. SENTENCING COMM’N, FIFTEEN-YEAR REPORT, *supra* note 1, at xiv-xv.

87. *Id.* at 127-28 & fig.4.9.

88. *Id.* at 127-28.

89. *Id.* at 128.

male counterparts.”⁹⁰ The Guidelines, however, provide that “family ties and responsibilities are not ordinarily relevant in determining whether a departure may be warranted”⁹¹ and add that a “defendant’s role in the offense is relevant in determining the applicable guideline range . . . but is not a basis for departure from that range.”⁹² The Sentencing Commission’s *Fifteen-Year Report* apparently took a different view. Like the Guidelines’ effort to restrict interjudge and geographic disparity, the Guidelines’ effort to limit racial, ethnic, and gender disparity appears to have been a bust.

IV. UNDER THE RADAR: THE SOURCES OF DISPARITY

An earlier Part of this Article noted that legislatures and sentencing commissions can themselves be sources of sentencing disparity. This Part focuses on other participants in the criminal justice system and the disparities they produce. Much disparity flies beneath the radar, safe from detection by statistical analysts.

A. Judges

The Feeney Amendment to the Prosecutorial Remedies and Tools Against the Exploitation of Children Today (PROTECT) Act of 2003⁹³ was prompted by statistics. Senator Hatch explained why this amendment limited and sometimes forbade judicial departures from the Sentencing Guidelines:

[C]ourts, unfortunately, have strayed further and further from [the Guidelines’] system of fair and consistent sentencing . . . [D]uring the period 1991 . . . to the year 2001, the number of downward departures—in other words, soft-on-crime departures, excluding those requested by the Government for substantial assistance and immigration cases along the Southwest border—has steadily climbed. In 1991, the number of downward departures was 1,241 and rose by 2001 to a staggering total of 4,098. This chart shows the rate of downward departures has increased over 100 percent during this period . . . and nearly 50 percent over the last 5 years alone.⁹⁴

According to the Justice Department, downward departures of the sort described by Senator Hatch increased from 9.7% of all cases in 1996 to 14.7% in 2001.⁹⁵ In 2001, however, 40% of all departures for reasons other than

90. *Id.* at 128-29.

91. U.S. SENTENCING GUIDELINES MANUAL § 5H1.6 (2004).

92. *Id.* § 5H1.7.

93. Pub. L. No. 10-821, 117 Stat. 650 (codified in scattered sections of 18 U.S.C.).

94. 149 CONG. REC. S5113, S5115 (daily ed. Apr. 10, 2003). Representative Tom Feeney maintained that judges were “arbitrarily deviating from the sentencing guidelines . . . based on their personal biases and prejudices, resulting in wide disparity in sentencing.” 149 CONG. REC. H2403, H2423 (daily ed. Mar. 27, 2003).

95. See Max Schanzenbach, *Have Federal Judges Changed Their Sentencing Practices? The Shaky Empirical Foundations of the Feeney Amendment*, 2 J. EMPIRICAL

October 2005]

DISPARITY

107

substantial assistance were initiated by the Justice Department itself, mostly as a result of plea agreements.⁹⁶ Although the number of departures increased, their magnitude declined so that sentences did not become more lenient.⁹⁷ Judges appointed by Republicans were as likely to depart downward as judges appointed by Democrats.⁹⁸

Geographic variations in departure rates suggest significant sentencing disparities. The General Accounting Office (GAO) found that, after controlling for offense and offender characteristics, downward departures in drug cases were 6.87 times more likely in the Ninth Circuit than in the Eighth.⁹⁹ The Sentencing Commission noted, however, that “fast-track” programs created by prosecutors in districts with high drug and immigration caseloads accounted for most of the disparity found by the GAO. These programs offer sentences far below Guidelines levels to defendants who waive everything—the right to challenge their indictments, the right to discovery, the right to seek the suppression of illegally obtained evidence, the right to trial, the right to appeal, and the right to seek post-conviction relief.¹⁰⁰

The Commission excluded “fast-track” cases from its own analysis. Using the Eighth Circuit figures as a baseline, it concluded:

The odds of receiving a downward departure for mitigating circumstances remain over three times higher in the Ninth Circuit than in the Eighth, almost three times higher in the Second, and two times higher in the DC circuit. In the Fifth Circuit, on the other hand, the odds of departure are just 17 percent that of the Eighth Circuit.¹⁰¹

The power to depart and the power to select a sentence within the Guidelines range are the most visible aspects of the sentencing discretion judges possess, but they are not the most important ones. To indicate more likely sources of judge-created sentencing disparities, this Part describes some sentencing issues in *United States v. Segal*, a case in which I am one of the defendant’s counsel.

Michael Segal owned and managed a large insurance brokerage. His company received premiums from its customers, which it paid to insurance carriers as the premiums became due, ordinarily thirty to forty-five days after they were received. Illinois insurance regulations required the agency to keep these premiums in a premium fund trust account (PFTA) in the interim. The

LEGAL STUD. 1, 12 (2005).

96. U.S. SENTENCING COMM’N, FIFTEEN-YEAR REPORT, *supra* note 1, at 111.

97. Schanzenbach, *supra* note 95, at 32-33.

98. *Id.* at 39.

99. U.S. GOV’T ACCOUNTING OFFICE, GAO-04-105, FEDERAL DRUG OFFENSES: DEPARTURES FROM SENTENCING GUIDELINES AND MANDATORY MINIMUM SENTENCES, FISCAL YEARS 1999-2001 app. III, at 50 (2003).

100. *See* *United States v. Ruiz*, 536 U.S. 622, 625-26 (2002) (describing the fast-track program in the Southern District of California).

101. U.S. SENTENCING COMM’N, FIFTEEN-YEAR REPORT, *supra* note 1, at 112.

government alleged that Mr. Segal was responsible for shortages in his agency's PFTA. At one point, according to the government, the PFTA contained \$30 million less than it should have. The government did not contend that Mr. Segal or his company ever failed to pay a premium when it was due, and no carrier or customer ever complained of any loss.¹⁰²

After Mr. Segal was convicted of mail fraud, wire fraud, and racketeering, the court ordered him to forfeit his company, worth tens of millions of dollars, and to forfeit \$30 million in addition. Although Mr. Segal is a first offender, the government maintains that the Guidelines dictate a prison sentence of between thirty years and life. Mr. Segal maintains that the Guidelines make him eligible for a sentence of less than three years. Here are some of the issues in dispute:

1. The Sentencing Guidelines set sentences in fraud cases primarily on the basis of the financial loss caused by the fraud.¹⁰³ Whether the PFTA deficit caused no loss or a loss of \$30 million is the largest point of disagreement. Thousands of pages of reported decisions consider the meaning of loss under the Guidelines, and the amount of loss appears to be disputed in nearly all white-collar crime cases.

2. Even if the PFTA deficit is treated as the measure of loss, the amount of this deficit is in dispute. A government witness testified that the deficit should be measured by consolidating the accounts of all of the insurance brokerages Mr. Segal owned, offsetting the deficit of one company's PFTA with the surplus in another's. A consolidated accounting would cut the deficit in half.

3. The Guidelines provide that the loss shall be reduced by whatever restitution the offender made before his offense was detected, and they define the time of detection as "the earlier of the time the offense was discovered by a victim or government agency or the time the defendant knew or should have known that the offense was detected or about to be detected by a victim or government agency."¹⁰⁴ Before Mr. Segal was charged with a crime or regulatory violation, his company eliminated the deficit. The government contends, however, that this restitution came too late. A witness reported a conversation in which Mr. Segal mentioned an anonymous letter the Illinois Department of Insurance had received concerning the deficit. The Director of the Illinois Department of Insurance testified, however, that neither he nor, as far as he knew, any other member of the department had received such a letter. Determining whether the Department of Insurance received an anonymous letter—and if so, when—could make a difference of twenty-five years in Mr. Segal's sentence.

102. Federal prosecutors rarely act as de facto state insurance commissioners. They apparently took that role in this case because Mr. Segal was reputed to possess useful information about prominent politicians.

103. U.S. SENTENCING GUIDELINES MANUAL § 2B1.1, cmt. n.3(A) (2004) (defining loss as pecuniary harm).

104. *Id.* § 2B1.1, cmt. n.3(E)(i) (internal numbering omitted).

October 2005]

DISPARITY

109

4. The government also contends that Mr. Segal should have known that his offense was about to be discovered. Some employees of Mr. Segal's company, after complaining to him about the PFTA deficit, resigned to work for a competitor. The government contends that Mr. Segal should have realized that the employees who had complained to him would also complain to the authorities. A judicial determination of what Mr. Segal "should have known" also could make a twenty-five-year difference in his sentence.

5. The government contends that Mr. Segal's restitution was insufficient because he replenished the PFTA with borrowed funds rather than with his own, a circumstance that Mr. Segal maintains makes no difference.

6. Mr. Segal and the government disagree about what version of the Guidelines to apply. The Sentencing Commission substantially increased the penalties for white-collar crimes in amendments to the Guidelines that became effective in November 2001. The government contends that the fraudulent scheme continued after the date of these amendments. Mr. Segal maintains that the crime of mail fraud is complete when an offender places an item in the mail with the intent to defraud and that all of the mailings alleged by the government occurred before the Guidelines amendments.

7. The government contends that Mr. Segal's sentence should be enhanced because his offense involved sophisticated means.¹⁰⁵ Mr. Segal maintains that there is nothing sophisticated about withdrawing money from a bank account.¹⁰⁶

8. The government argues that Mr. Segal's sentence should be enhanced because he derived more than one million dollars from a financial institution.¹⁰⁷ Mr. Segal argues that an insurance brokerage does not fall within the Guidelines' definition of "financial institution."

9. The government seeks an enhancement for abuse of trust.¹⁰⁸ Mr. Segal maintains that because the definition of his offense required an abuse of trust, this enhancement would constitute double counting.

10. The government also contends that Mr. Segal's sentence should be enhanced because he was the leader of a criminal activity that involved five or more participants or was otherwise extensive.¹⁰⁹ Mr. Segal replies that no evidence shows that the company employees identified by the government ever did anything criminal.

105. *See id.* § 2B1.1(b)(9).

106. A judge can determine whether an offender's means were sophisticated with virtually no risk of reversal by an appellate court. *Compare* United States v. Rettenberger, 344 F.3d 702 (7th Cir. 2003) (holding that simulating an illness to defraud an insurance company is "sophisticated"), *with* United States v. Kaufman, 800 F. Supp. 648 (N.D. Ind. 1992) (holding that maintaining fraudulent and nonfraudulent sets of books is not "sophisticated").

107. U.S. SENTENCING GUIDELINES MANUAL § 3B1.1 (2004).

108. *Id.*

109. *Id.* § 3B1.1(a).

11. Finally, the government maintains that Mr. Segal's sentence should be enhanced because he obstructed justice.¹¹⁰ The government alleges that Mr. Segal made a false statement during a two-hour interrogation immediately after his arrest. Mr. Segal denies both that his statement was false and that it could have obstructed the government's investigation. The government dismissed the false-statement count in Mr. Segal's indictment shortly before his case was submitted to the jury, but its dismissal of the false-statement charge does not preclude it from resurrecting this allegation as a sentencing enhancement.

This list presents a simplified description of some of the issues posed by the Sentencing Guidelines in Mr. Segal's case. Of course, Mr. Segal is the same person and his crime remains the same crime regardless of how these issues are resolved.¹¹¹ Resolving them would not advance anyone's understanding of what punishment Mr. Segal deserves.

If a dozen judges were to throw away the Guidelines and consider the punishment Mr. Segal deserves, they might well disagree. None of them, however, would make twenty-five years of imprisonment depend on whether the Illinois Department of Insurance received an anonymous letter the day before or the day after Mr. Segal replenished his company's PFTA. None of them would worry about whether a \$30 million PFTA deficit should be characterized as no loss or as a \$30 million loss. None of them would think it important to decide whether an insurance brokerage should be called a financial institution. All twelve judges might agree that the issues briefed and argued in Mr. Segal's case have little bearing on the sentence he should receive.

The asserted justification for spending pages, hours, legal-research fees, and sweat on the sorts of legalisms illustrated by the Segal case is that doing so promotes equality in sentencing. The considerations that the Guidelines make decisive may be artificial, but artificiality is preferable to the unrestricted play of personality.

The premise of this argument seems flawed. Few of the issues presented in Mr. Segal's case are so clear that every straight-shooting judge would resolve them the same way. In resolving many of these issues, a judge would almost certainly be influenced by his own view of the seriousness of Mr. Segal's crime. There is no reason at all to suppose that a dozen judges asked to decide Mr. Segal's case under the Sentencing Guidelines would be less likely to disagree than a dozen judges asked to determine his sentence without the benefit of the Guidelines. Because the Guidelines force yes-or-no choices with large consequences, the disparity in sentences might be greater. When judges

110. *See id.* § 3C1.1.

111. It seems wasteful (indeed, almost unreal) to labor over Guidelines issues of characterization now that the decision in *United States v. Booker*, 125 S. Ct. 738 (2005), has made the Guidelines advisory. *Booker*, however, requires the sentencing judge to calculate the Guidelines sentence. In addition, the judge who will determine Mr. Segal's sentence is a Vice Chair of the Sentencing Commission.

October 2005]

DISPARITY

111

disagree in resolving what appear to be issues of fact and law, however, people no longer call their disagreements “sentencing disparity.” The judges’ resolutions of these issues become part of the controls from which disparity is measured.

B. *Defense Attorneys*

A lawyer in a case like Mr. Segal’s might have noted that his client’s allegedly fraudulent conduct extended beyond the effective date of the current Sentencing Guidelines and might have failed to challenge the government’s assumption that the current version of the Guidelines apply. This lawyer might thereby have cost his client ten years in the penitentiary. No state and, indeed, no other jurisdiction in the world makes an offender’s sentence as dependent on the quality of his counsel as do the federal courts under the Guidelines. Finding what is relevant to a case in the 629-page Guidelines Manual, the 1100 pages of appendices explaining Guidelines provisions, and the endless judicial decisions interpreting the Guidelines takes a very good lawyer, and not every federal defendant has one. The disparate performances of attorneys generate sentencing disparities that fly beneath the radar.

C. *Probation Officers*

The official who will first pass judgment on the complex legal issues posed by Mr. Segal’s case has no legal training at all. Before the Guidelines, probation officers conducted investigations of the background and conduct of offenders and prepared presentence investigation reports concerning the correctional treatment these offenders should receive. The Guidelines made the traditional function of these officers superfluous, but jobs for the boys are jobs for the boys. In the Guidelines regime, probation officers became unlicensed lawyers, law clerks, and Guidelines specialists.¹¹²

Probation officers seem no less likely than judges to vary in their application of the Guidelines. They may be more likely. Pamela Lawrence and Paul Hofer asked forty-six randomly selected probation officers to apply the relevant conduct Guideline to four offenders who had participated in the same drug conspiracy. They found substantial variation. For the least culpable of the offenders, the probation officers’ calculations ranged from one to five years in

112. The decision in *United States v. Booker*, 125 S. Ct. 738 (2005), made the Sentencing Guidelines advisory and directed judges to shape sentences in accordance with the general objectives specified in 18 U.S.C. § 3553(a) (2005). *Booker* appeared to allow probation officers to resume their historic role, but at least in the Northern District of Illinois, they do not want to. They insist that they are Guidelines specialists, and they do not discuss § 3553(a) objectives in their reports. If a defendant wishes to present an argument based on § 3553(a), he must present it directly to the sentencing judge.

prison.¹¹³

D. Law Enforcement Officers

Because drug sentences under the Guidelines are determined by drug quantities, undercover agents can increase an offender's sentence by persuading him to purchase more of a drug or by continuing to sell to him until he passes a sentencing threshold.¹¹⁴ Agents may be able to manipulate other circumstances affecting the offender's sentence as well, such as the presence of a gun and the location of the crime.¹¹⁵ Some courts have recognized "sentencing entrapment" as a justification for reducing sentences in cases of outrageous law enforcement conduct,¹¹⁶ but others have not.¹¹⁷ The manipulation of sentencing exposure by law enforcement officers appears to be a significant source of disparity.¹¹⁸

E. Prosecutors

In 1989, I wrote that although the Federal Sentencing Guidelines were likely to increase the power of prosecutors, "[g]uilty plea rates are currently so high that even substantial increases in prosecutorial bargaining power cannot yield great increases in these rates."¹¹⁹ I was wrong. Guilty pleas, which accounted for 87% of all federal convictions in the years before the Guidelines, account for 97% today.¹²⁰ Marc Miller remarks that the Guidelines have "achieved the virtual elimination of criminal trials in the federal system."¹²¹

The Sentencing Guidelines and mandatory minimum sentences have set the stage for the "good-cop, bad-cop" stratagem on a grand scale. Congress and the Sentencing Commission play the bad cops, threatening the accused with harsh

113. Pamela B. Lawrence & Paul J. Hofer, *An Empirical Study of the Application of the Relevant Conduct Guideline 1B1.3*, 4 FED. SENT'G REP. 330 (1992).

114. See, e.g., *United States v. Rosen*, 929 F.2d 839 (1st Cir. 1991).

115. See U.S. SENTENCING COMM'N, FIFTEEN-YEAR REPORT, *supra* note 1, at 82.

116. See *United States v. Barth*, 990 F.2d 422 (8th Cir. 1993).

117. See *United States v. Williams*, 954 F.2d 668 (11th Cir. 1992).

118. See Eric P. Berlin, *The Federal Sentencing Guidelines' Failure To Eliminate Sentencing Disparity: Governmental Manipulations Before Arrest*, 1992 WIS. L. REV. 187; Andrew G. Deiss, Comment, *Making the Crime Fit the Punishment: Prearrest Sentence Manipulation by Investigators Under the Sentencing Guidelines*, 1994 U. CHI. LEGAL F. 419 (1994); Jeffrey L. Fisher, Comment, *When Discretion Leads to Distortion: Recognizing Pre-Arrest Sentence-Manipulation Claims Under the Federal Sentencing Guidelines*, 94 MICH. L. REV. 2385 (1996).

119. Albert W. Alschuler, *The Selling of the Sentencing Guidelines*, in THE U.S. SENTENCING GUIDELINES: IMPLICATIONS FOR CRIMINAL JUSTICE 49, 91 n.4 (Dean J. Champion ed., 1989).

120. U.S. SENTENCING COMM'N, FIFTEEN-YEAR REPORT, *supra* note 1, at 30.

121. Marc L. Miller, *The Foundations of Law: Sentencing Equality Pathology*, 54 EMORY L.J. 271, 277 (2005).

October 2005]

DISPARITY

113

treatment. The prosecutor takes the part of the good cop, promising to protect the defendant if he abandons the right to trial and cooperates. Substantial sentencing discretion remains except for defendants who resist the prosecutor's will.

The Sentencing Commission's policy statements admonish judges to reject plea agreements that undermine the Guidelines.¹²² One study concluded, however, that "key participants in the sentencing process were generally unfamiliar with . . . the policy statements."¹²³ The Sentencing Commission observes that "[j]udicial review of plea agreements pursuant to the policy statements . . . appears to be very limited."¹²⁴ In each of the years between 1997 and 2002, the reason judges gave most often or next-most often for departing from the Guidelines was "pursuant to plea agreement."¹²⁵ Most plea agreements, moreover, do not require departure from the Guidelines. Bending the Guidelines is enough. Federal Rule of Criminal Procedure 11(c)(1)(C) provides:

[T]he plea agreement may specify that an attorney for the government will . . . agree that a specific sentence or sentencing range is the appropriate disposition of the case, or that a particular provision of the Sentencing Guidelines, or policy statement, or sentencing factor does or does not apply (such a recommendation or request binds the court once the court accepts the plea agreement).¹²⁶

In 2003, a memorandum by Attorney General John Ashcroft declared, "Any sentencing recommendation made by the United States in a particular case must honestly reflect the totality and seriousness of the defendant's conduct and must be fully consistent with the Guidelines and applicable statutes and with the readily provable facts about the defendant's history and conduct."¹²⁷ Like similar directives from Attorneys General before him, Ashcroft's apparent restriction of plea bargaining made headlines. On the surface, it appeared to be a significant reform. The memorandum's exceptions, however, largely swallowed its rule. Prosecutors could bargain when their cases were weak, when they could get useful information from a defendant, when immigration and drug cases clogged the courts, when enhancements would remove any incentive for a defendant to plead guilty, or when the prosecutors

122. U.S. SENTENCING GUIDELINES MANUAL § 6B1.2 (2004).

123. U.S. SENTENCING COMM'N, FIFTEEN-YEAR REPORT, *supra* note 1, at 87 (citing Diane Marks, Prosecutorial Discretion in Connecticut: A Lesson in Indirect Constraint, Paper Presented at the Yale Law School Symposium on Federal Sentencing Guidelines (2002)).

124. U.S. SENTENCING COMM'N, FIFTEEN-YEAR REPORT, *supra* note 1, at 144.

125. *Id.* at 92.

126. FED. R. CRIM. P. 11(c)(1)(C).

127. Memorandum from John Ashcroft, Attorney General, to All Federal Prosecutors Regarding Departmental Policy Concerning Charging Criminal Offenses, Disposition of Charges, and Sentencing (Sept. 22, 2003), http://www.usdoj.gov/opa/pr/2003/September/03_ag_516.htm (last visited Sept. 16, 2005).

could get written approval from their superiors. The prosecutors apparently could bargain whenever doing so would advance their interests. The only thing they could not consider was the merits of a case.

I am told that, in one federal district, the United States Attorney circulated the Ashcroft memorandum to the other lawyers in his office with a note saying that it appeared to require no change in “what this office is doing already.” In this district, prosecutors and defense attorneys often bargained directly about how much time a defendant would serve and then determined what Guidelines applications would implement their bargains. The Sentencing Commission notes that prosecutors have viewed directives like Attorney General Ashcroft’s with skepticism and that “nationwide DOJ policies . . . were less determinative of prosecutorial conduct than internal U.S. Attorney’s office policies.”¹²⁸

Together with mandatory minimum sentencing legislation, the Guidelines have added to the prosecutors’ armaments. They have given prosecutors more leverage partly by increasing the severity of criminal sentences. The Sentencing Reform Act provides, “The sentencing guidelines prescribed under this chapter shall be formulated to minimize the likelihood that the Federal prison population will exceed the capacity of the Federal prisons,”¹²⁹ but the number of inmates in federal prisons has more than quadrupled since 1985.¹³⁰ This growth in the federal prison population is not attributable to increased crime rates but is attributable in part to larger federal caseloads.¹³¹ It is also partly attributable to the fact that a higher proportion of convicted offenders go to prison.¹³² What has driven the explosion in the prison population even more than these developments, however, is the greater length of prison terms. The Sentencing Commission notes that “federal offenders sentenced in 2002 will . . . spend about twice as long in prisons as did offenders sentenced prior to passage of the [Sentencing Reform Act].”¹³³

Congress allows departures from both the Guidelines and mandatory minimum sentences for “substantial assistance in the investigation or prosecution of another person,” but it allows these departures only when prosecutors request them.¹³⁴ When a prosecutor seeks a substantial assistance departure, the bottom is the limit. No statute or guideline constrains the extent of the defendant’s reward. Departures for substantial assistance occur in about

128. U.S. SENTENCING COMM’N, FIFTEEN-YEAR REPORT, *supra* note 1, at 84.

129. 28 U.S.C. § 994(g) (2005).

130. BUREAU OF JUSTICE STATISTICS, SOURCEBOOK OF CRIMINAL JUSTICE STATISTICS 2002, at 487 tbl.6.12 (2004).

131. U.S. SENTENCING COMM’N, FIFTEEN-YEAR REPORT, *supra* note 1, at 76.

132. *See id.* at vi (“The use of imprisonment spiked in the early years of guidelines implementation and then resumed a long gradual climb, reaching 86 percent of all offenders by 2002, about 20 percent higher than it had been in the preguidelines era.”).

133. *Id.* at 46.

134. *See* 18 U.S.C. § 3553(e) (2005); U.S. SENTENCING GUIDELINES MANUAL § 5K1.1 (2004).

October 2005]

DISPARITY

115

17% of all cases and other departures in about 18% more.¹³⁵ Substantial assistance departures, however, are larger and account for twice as much variation in federal sentences (4.4% of all variation versus 2.2%).¹³⁶ The frequency of substantial assistance departures varies greatly from one district to the next.¹³⁷

Congress provided another escape route from mandatory minimum sentences in a “safety-valve” provision for low-level drug offenders. This escape route is open, however, only for offenders who cooperate with prosecutors.¹³⁸

As initially approved, the Sentencing Guidelines offered a two-level reduction for “acceptance of responsibility.”¹³⁹ The Sentencing Commission later provided another one-level reduction for “timely” notification of an intent to plead guilty.¹⁴⁰ In the PROTECT Act of 2003, Congress prohibited judges from granting a reduction for timely notification of an intent to plead guilty unless a prosecutor requested it.¹⁴¹

In several federal jurisdictions with high immigration and drug caseloads, prosecutors created what they called “fast-track” programs offering sentences far below Guidelines levels to defendants who waive almost all of their procedural rights.¹⁴² In the PROTECT Act, Congress authorized departures from the Guidelines of no more than four levels to implement these programs.¹⁴³

Prosecutors control whether mandatory minimum sentences will be imposed. In a recent case, a twenty-two-year-old defendant was arrested on two occasions for possessing both drugs and a firearm. Although he had no criminal record before these arrests, his conviction of the second offense required the court to impose a mandatory minimum sentence of twenty-five years, which the offender would be required to serve after he completed his first sentence. When Judge Myron H. Thompson imposed the total sentence of forty years required by section 924 of Title 18, he called this sentence “draconian.” He noted that not only would the offender’s child grow up without a father but his grandchildren, if he had any, would be teenagers or young adults before he was

135. U.S. SENTENCING COMM’N, FIFTEEN-YEAR REPORT, *supra* note 1, at xiii. In some jurisdictions, substantial assistance and other departures are so frequent that only a minority of offenders receive Guidelines sentences. *See* United States v. Green, 346 F. Supp. 2d 259, 275 n.53 (D. Mass. 2004).

136. U.S. SENTENCING COMM’N, FIFTEEN-YEAR REPORT, *supra* note 1, at 102.

137. *Id.* at 103.

138. 18 U.S.C. § 3553(f)(5) (2005).

139. U.S. SENTENCING GUIDELINES MANUAL § 3E1.1(a) (2004).

140. *Id.* § 3E1.1(b).

141. PROTECT Act of 2003, Pub. L. No. 108-21, § 401(g)(2), 117 Stat. 650, 671.

142. *See supra* text accompanying notes 99-100.

143. PROTECT Act, § 401(m)(2)(B), 117 Stat. at 675.

released.¹⁴⁴

Judge Thompson and other judges are required to impose the mandatory minimum sentences specified by section 924, but prosecutors have a choice. The Sentencing Commission reports that, after the exercise of prosecutorial discretion in charging and plea bargaining, only 20% of the offenders who used firearms to commit drug crimes received the mandatory sentences that section 924 prescribes, and offenders who carried firearms without using them received the section 924 enhancements even less often.¹⁴⁵

It is difficult to believe that Congress cares about sentencing disparity in the slightest. After many of its members ringingly denounced disparity, Congress enacted the PROTECT Act restricting downward departures by judges. The same statute, however, validated “fast-track” programs that effectively jettisoned the Guidelines. The PROTECT Act followed the pattern of the previous twenty years—jeering disparities created by judges while cheering those created by prosecutors. Judge William G. Young remarks, “Enhanced plea bargaining is actually the central goal of the guidelines.”¹⁴⁶ When Senators Thurmond and Kennedy co-sponsored the Sentencing Reform Act, observers wondered which one was selling the farm. Today we know the answer, and it was not Senator Thurmond. Increasing prosecutorial power and the severity of criminal punishments was not the unintended consequence of Guidelines designed to reduce sentencing disparity. Instead, it was the point all along. Disparity was just a code word.¹⁴⁷

CONCLUSION

In a decision requiring courts to weigh sugar cubes, blotter paper, and other carrier media, Judge Easterbrook wrote, “Experience with the guidelines suggests . . . [that every] attempt to make the system of sentences ‘more rational’ carries costs and concealed irrationalities, both loopholes and unanticipated severity. Criminals have neither a moral nor a constitutional claim to equal . . . treatment.”¹⁴⁸

Sentencing guidelines are premised on the view that criminals do have some claim to equal treatment, and except, perhaps, for Judge Easterbrook, everyone I know favors the idea of sentencing guidelines.¹⁴⁹ With Judge

144. *United States v. Washington*, 301 F. Supp. 2d 1306 (M.D. Ala. 2004).

145. U.S. SENTENCING COMM’N, FIFTEEN-YEAR REPORT, *supra* note 1, at 90.

146. *United States v. Green*, 346 F. Supp. 2d 259, 270 (D. Mass. 2004).

147. Some people, including most of the legal academics who write about sentencing, did imagine that disparity reduction was the name of the game, but these ingénues were easy marks for those who sought a system of harsh penalties controlled by prosecutors in which virtually every defendant, if sane, would bend to the prosecutors’ will.

148. *United States v. Marshall*, 908 F.2d 1312, 1326 (7th Cir. 1990), *aff’d sub nom.*, *Chapman v. United States*, 500 U.S. 453 (1991).

149. I offer my own view of how useful guidelines could be structured in Alschuler,

October 2005]

DISPARITY

117

Easterbrook, however, one may doubt that the quest for perfect parity among rapist-robber-muggers justifies a 629-page guidelines manual with 1100 pages of appendices and more legalisms than *Jarndyce v. Jarndyce*. The effort to promote equality through detailed directions has not worked for the reasons Judge Easterbrook said it could not work.

The Federal Sentencing Guidelines have not reduced disparity between black and white offenders. Instead, the Guidelines and Congress's mandatory minimum sentencing legislation have created and aggravated racial disparities. Under the Guidelines, the gender gap in federal sentences has also widened, and geographic disparities have increased. The Guidelines may have reduced the influence of judicial personality on sentencing but only slightly and only when one looks exclusively to the sentences judges initially impose. The Parole Commission almost certainly did as well as the Guidelines or better in reducing disparity in the sentences offenders actually serve. The artificial rules and standards of the Sentencing Guidelines have created inequalities of their own, and the price of whatever success the Guidelines have achieved in reducing judge-created sentencing disparities has been the burgeoning of prosecutor-created disparities.

The Failure of Sentencing Guidelines, *supra* note 14, at 939-49, and Albert W. Alschuler, *Monarch, Lackey, or Judge*, 64 U. COLO. L. REV. 723, 731-32 (1993).

