2003


Ian Weinstein
Fordham University School of Law, iweinstein@law.fordham.edu

Follow this and additional works at: http://ir.lawnet.fordham.edu/faculty_scholarship

Part of the Conflicts of Law Commons, Criminal Law Commons, Evidence Commons, and the Jurisprudence Commons

Recommended Citation
Available at: http://ir.lawnet.fordham.edu/faculty_scholarship/417
FIFTEEN YEARS AFTER THE FEDERAL SENTENCING REVOLUTION: HOW MANDATORY MINIMUMS HAVE UNDERMINED EFFECTIVE AND JUST NARCOTICS SENTENCING

Ian Weinstein*

INTRODUCTION

Federal criminal sentencing has changed dramatically since 1988. Fifteen years ago, judges determined if and for how long a defendant would go to jail. Since that time, changes in substantive federal criminal statutes, particularly the passage of an array of mandatory minimum penalties and the adoption of the federal sentencing guidelines, have limited significantly judicial sentencing power and have remade federal sentencing and federal criminal practice. The results of these changes are significantly longer federal prison sentences, as was the intent of these reforms, and the emergence of federal prosecutors as the key players in sentencing. Yet, at the same time, average sentence length appears to be falling slowly as judicial tendency to use the authority granted in the United States Sentencing Guidelines (the “Guidelines”) to mitigate sentences through downward departures appears to be increasing.

This Article begins with the following two-part question: Why are (1) an apparent increase in judicial sentence mitigation through downward departure and (2) declining average sentence length two legacies of the 1980s’ shift to longer prison sentences and increased prosecutorial control of sentence length? Although some might argue these legacies result from judicial resistance to loss of authority or general recognition among judges and prosecutors that federal sentences have become too long, this Article argues that a closer look reveals a much more complex and disturbing picture. Specifically, this Article will demonstrate the presence of two different federal criminal law systems operating in tandem.

The first system is that of narcotics prosecution wherein prosecutorial power often is unchecked and sentences often are unpredictable, but generally are quite harsh. Narcotics sentences have been decreasing steadily for almost ten years—a troubling instability—and there are wide disparities in sentences among similarly-situated defendants. Too many defendants receive sentences that are out of

---

* Associate Professor of Law, Fordham University School of Law. I want to thank my research assistant, Matthew R. Halloran. I was an Associate Attorney with the Federal Defender Division of the Legal Aid Society of New York from 1988 to 1991 and I have been a member of the Criminal Justice Act Panel for the Southern District of New York since leaving that office. Although I strive for analytic objectivity, my views are informed by my experiences representing hundreds of people accused of criminal conduct. Fordham University provided financial support for this project.
proportion to the wrongfulness of their conduct and too few will accept the risks that come with trying to enforce their rights in the face of often overwhelming prosecutorial power.

The second system involves the bulk of non-narcotics federal prosecution. Although there are many difficult questions presented by current prosecutions in fraud, robbery and other federal non-narcotics cases, the sentences in these areas are far more stable and predictable. Although harsh, problems of excessive severity are not as pervasive among sentences in the non-narcotics area, as there are many fewer mandatory minimum cases and judges are a more effective counterweight to instances of prosecutorial overreaching. Federal non-narcotics sentencing may not be the best of all possible systems of criminal sanctions, but it has proved reasonably stable and predictable over time. This Article will argue that much of the stability in non-narcotics sentences is the result of a defensible balance of power among prosecutors, judges and defense lawyers.

Meanwhile, the vast increase in prosecutorial power to control narcotics sentences is at the core of the problems with federal narcotics sentencing. The profusion of new narcotics and gun proscriptions, almost all of which carry mandatory minimum prison sentences, transformed the traditional prosecutorial power to charge into the contemporary prosecutorial power to determine the length of the sentence the defendant will serve. Although the Guidelines shifted sentencing power from judges to prosecutors across the whole range of federal crimes, this Article argues that mandatory minimum statutes have made this shift in power so extreme in narcotics cases that it has damaged the adversary system and wrought real injustice among those sentenced for narcotics crimes.

This Article has two descriptive and four analytic sections. Part I describes the changes that brought us from a system of almost unfettered judicial sentencing discretion to a system in which federal judges have very limited power to control the sentence in a particular case. Part II identifies some major trends in federal sentencing today and discusses some of the ways federal criminal practice has changed in response to changes in sentencing law. Part III describes the structure of federal criminal case processing and sentencing. Part III also analyzes how, given that structure, prosecutorial discretion has become the key force controlling sentencing in cases involving mandatory minimum statutes, and to a lesser degree, the Guidelines cases in general. Part IV analyzes how judicial power and defendants' bargaining power within the adversarial system can function as salutary checks on prosecutorial discretion, ensuring both efficiency and fairness. Part IV argues that we can predict certain patterns in a well-structured system of federal sentencing. Part V analyzes sentencing patterns in non-narcotics cases and argues that sentencing practices are consistent with the framework suggested in Parts III and IV. Part V then concludes that the Guidelines work reasonably well in the non-narcotics cases. Part VI examines sentencing practices and patterns in narcotics cases and discusses why those patterns signal deep flaws in our current system, flaws that can be traced to the mandatory minimum sentencing statutes.
Part VI closes with another criticism of narcotics sentencing, arguing that regardless of the systemic impact, narcotics sentences are simply too long. The Conclusion joins the call for restoring some judicial sentencing power when the next great wave of criminal justice reform sweeps through our political culture.


Before discussing the current trends in federal sentencing, this Article offers a brief overview of how federal sentencing laws changed in the late 1980s and early 1990s in order to fully explicate the currently bifurcated federal criminal law system of non-narcotics cases on the one hand and narcotics cases on the other. This summary may help to situate the current situation for those less familiar with this area and also will signal some of my assumptions to those more familiar with the recent history of federal sentencing.

My first exposure to federal criminal sentencing was during the twilight of the age of unfettered judicial sentencing discretion. In the mid-1980s, federal criminal sentencing was characterized by almost completely individualized and unreviewable judicial decision-making. The length of a federal criminal defendant’s sentence was almost entirely in the control of the district judge, although a negotiated plea agreement could cap the possible sentence with a chosen statutory maximum, and prosecutors and defense lawyers developed local customs to influence their judges. In most cases, however, judges could impose any combina-

1. I practiced criminal law in two jurisdictions that did not have sentencing guidelines, the Superior Court of the District of Columbia and the United States District Court for the Southern District of New York, before the guidelines were in force.

2. Mistretta v. United States, 488 U.S. 361, 363 (1989) (commenting on the extent of pre-Guidelines discretion in the introduction to the decision affirming the constitutionality of the Guidelines). Of course, judicial sentencing discretion was restrained by statutory maximums, but as they are generally quite high, relative to the expected range of punishment, and subject to manipulation by charging multiple counts, they rarely imposed any consequential limits on sentences. Judicial power was also supposed to be moderated by the parole board’s power to release prisoners, but the parole board’s practices were known and accounted for by many judges, particularly after the promulgation of the parole guidelines. See Solem v. Helm, 463 U.S. 277, 285-86 (1983).


4. Early in my career I litigated criminal cases in two jurisdictions without the benefit of sentencing guidelines. In both the Superior Court of the District of Columbia and the U.S. District Court for the Southern District of New York, before the Guidelines were in force, it was common for the prosecutor to agree not to reveal certain
tion of fines and restrictions on liberty from suspending the imposition of sentence, with or without supervision by the probation department, to a sentence of imprisonment at the statutory maximum and the maximum authorized fine. Judges were not required to state reasons for their sentences and a "sentence within statutory limits was, for all practical purposes, not reviewable on appeal." So long as the sentence was within the maximum and procedurally proper, the judge was not subject to review and could not be reversed whether the imposition of sentence for bank robbery was suspended or the defendant was given twenty-five years in the penitentiary.

Consistent with and providing a very limited check on judicial sentencing authority was the parole system. The parole system operated under the Parole Board, which acted under the authority of the executive branch; the Parole Board had the authority to release prisoners after they served a portion of their sentence. Thus, sentences actually served were not completely at the discretion of judges but judges were able to account for parole when they imposed sentence under the old law. Judicial discretion was the key feature of "old law" sentencing, as the

information to the judge (and the judges never directly sought that information) and to agree to take "no position" on defense requests at sentence, which were then frequently granted.

5. Upon entering a judgment of conviction of any offense not punishable by death or life imprisonment, any court having jurisdiction to try offenses against the United States, when satisfied that the ends of justice and the best interest of the public as well as the defendant will be served, may suspend the imposition or execution of sentence and place the defendant on probation for such period and upon such terms and conditions as the court deems best. 18 U.S.C. § 3651 (1984), repealed by Sentencing Reform Act of 1984, Pub. L. No. 98-473, § 212(a)(2), 98 Stat. 1837, 1987 (1987) (applicable to offenses committed prior to Nov. 1, 1987).


7. The non-review doctrine may well have been founded on a mistake, but it became well entrenched by the 1930s. See Kutak & Gottschalk, supra note 6, at 468-72 (tracing history of doctrine of non-review of sentences, suggesting it was founded on a misreading of law, and demonstrating convincingly that courts of appeal occasionally used procedural dress to remand egregious sentences, but arguing for appellate review because doctrine of non-reviewability generally held firm and prevented development of sentencing standards); see also Kevin R. Reitz, Sentencing Guideline Systems and Sentence Appeals: A Comparison of Federal and State Experiences, 91 Nw. U. L. Rev. 1441, 1443-44 (1997) (discussing doctrine of non-review).

8. The system of judicial discretion and individualized sentencing fit with the rehabilitative model of punishment, which required sentences tailored to each individual's particular background and needs. See Stith & Koh, supra note 2, at 227 (discussing rise and fall of rehabilitative model); see also Robert J. Cottrol, Hard Choices and Shifted Burdens: American Crime and American Justice at the End of the Century, 65 Geo. Wash. L. Rev. 506, 506 nn.7-10 (1997) (reviewing MICHAEL TONRY, MALIGN NEGLECT: RACE, CRIME AND PUNISHMENT IN AMERICA (1995), and citing authors on each side of rehabilitation and retribution debate).
MANDATORY MINIMUMS AND NARCOTICS SENTENCING

sentencing system before the Guidelines has come to be known.9

Two related changes, the Guidelines and mandatory minimum sentencing statutes, have remade federal criminal sentencing. Prosecutors today have the discretion and the power to control many federal sentences. The formal, statutory beginning of this change was the Sentencing Reform Act of 198410 ("SRA"), in which Congress mandated the creation of a new commission to develop guidelines. Impetus for the SRA came first from the liberal sentencing reform movement of the 1970s.11 That movement criticized the standardless, unreviewable sentencing regime under the authority of federal judges as lawless and subject to the prejudices and whims of individual judges. Critics focused on extremely harsh sentences for relatively minor offenses and the probability of race and class discrimination in sentencing.12 They urged an increase in the use of alternatives to incarceration, the development of sentencing standards, and appellate review to remedy these problems.13

Insightful commentators offer versions of how that liberal 1970s movement made a 1980s legislative alliance with conservative law-and-order interests who shared the goals of reining in judges and making sentences more uniform.14 Riding the "crime wave" and controlling the legislative process, conservative interests were able to meet the demands of their liberal allies for rules and reviewability while achieving their goals of lengthening criminal sentences and eliminating the rehabilitative model of criminal sentencing.

The passage of the SRA created the United States Sentencing Commission (the "Commission"). Congress directed the Commission to draft the Guidelines.15 The Guidelines16 replaced a system of judicially determined discretionary sentencing with a system under which judges apply a complex set of rules keyed to the

---

9. Of course prosecutors played a very important role. In many cases and in many jurisdictions it was common for judges to take cues from the prosecutor and attend carefully to his or her recommendations, but the final authority rested with the sentencing judge.


11. Stith & Koh, supra note 2, at 230-33 (recounting early history of sentencing reform); see also Reitz, supra note 7, at 1447 (noting "reformist vision" of early movement to change sentencing).

12. See MARVIN E. FRANKEL, CRIMINAL SENTENCES 12 (1973) (criticizing discretionary sentencing and the way his colleagues on the federal bench exercised their power); see also Stith & Koh, supra note 2, at 227-30 (analyzing growing debate over rehabilitative sentencing during early 1970s).

13. Norval Morris, another important voice for sentencing reform, advocated appellate review of sentences to develop a "common law of sentencing." See Reitz, supra note 7, at 1447 (quoting Norval Morris, Toward Principled Sentencing, 37 Md. L. Rev. 267, 284 (1977)).

14. See Stith & Koh, supra note 2, at 232-48 (tracing results of political compromise between Sens. Kennedy and McClellan); see also David Yellen, What Juvenile Court Abolitionists Can Learn From the Failures of Sentencing Reform, 1996 Wis. L. Rev. 577, 585-90 (using development of Guidelines to illustrate how sentencing reform politics inevitably lead to harsher punishments).


defendant's conduct and criminal history to determine a specific sentencing range—typically a minimum number of months of imprisonment and a maximum number of months of imprisonment—within which the judge must pick a specific sentence of imprisonment. The sentencing ranges under the Guidelines are set out in the sentencing table, which simultaneously increased sentences for many classes of offenses.

The Guidelines are a complex system and represent many compromises, but one clear result of their adoption is that they have completely altered the balance of sentencing power between judges and prosecutors in at least two major ways. First, the Guidelines made the prosecutors' decisions about how much to investigate a given case and about which particular criminal statute to use to charge a defendant, crucial to the ultimate sentence. As this Article will discuss below, in many cases prosecutors can and do bargain with defendants over both the charges to which he or she will plead guilty and the facts upon which the prosecutor will argue the sentence should be based. These two decisions—charges and facts—can have significant impact upon the defendant's sentence. Second, and more fundamentally, the Guidelines replaced unreviewable judicial discretion with a system of rules subject to appellate review. The availability of appellate review has given prosecutors and defense lawyers a set of legally enforceable claims to replace the arguments and signals that lawyers only could hope would persuade judges deciding cases under the old sentencing system.

The most fundamental current constraint on judicial authority is that judges are bound, with certain limited exceptions, to impose sentences within the range determined by the Guidelines. That range is determined by the charges brought against the defendant and some aspects of the conduct underlying the charges. The legal requirement that a judge impose a specific sentence is subject to the exercise of discretion in two areas. The first area of discretion occurs in about two-thirds of


17. Stephen Breyer, The Federal Sentencing Guidelines and the Key Compromises Upon Which They Rest, 17 Hofstra L. Rev. 1, 8-31 (1988) (discussing six compromises and arguing that only one involved interests trading, while others are rooted in practical needs and unresolved theoretical issues).

18. Several of the most frequently used Guidelines, including those for theft, fraud, and narcotics turn, to an important degree, on the amount of money or drugs involved in the case. See U.S.S.G. §§ 2B1.1, 2D1.1. In some cases, a prosecutor may choose to end an investigation and limit the amount of money or drugs involved, or to pursue an investigation in an effort to increase the sentence.

19. Sentence guidelines systems can be offense- or conduct-based. In a pure offense-based system, the sentence turns only on the offense of conviction. In a pure conduct-based system, the sentence accounts for all the defendant's bad conduct, regardless of the offense of conviction. In the latter system, a robbery defendant who also sold drugs could be sentenced for the drugs after being convicted of robbery. In the former he could not. The drafters of the Guidelines adopted a "modified real offense system," under which all conduct similar or relevant to the conduct of conviction is considered in determining sentence. See Julie R. O'Sullivan, In Defense of the U.S. Sentencing Guidelines Modified Real-Offense System, 91 NW. U. L. Rev. 1342, 1342-49 (1997) (defending fundamental choice of modified, as opposed to pure, charged offense sentencing as necessary to impose sentences meaningfully related to underlying wrongful conduct).
cases, wherein judges determine that the sentence must be within the Guidelines range. Once that determination is made, the judge retains the unreviewable authority to decide at what level in the applicable range to sentence the defendant. The second area of discretion concerns the remaining one-third of all cases, which involve downward departures or sentences imposed below the otherwise applicable Guidelines range. About half of the cases in the downward departure group are sentenced outside the Guidelines at a lower level than that prescribed by the Guidelines because the defendant provided substantial assistance in the investigation or prosecution of others. Only the prosecutor can trigger these departures; however, once the government requests a substantial assistance departure, the extent of the departure is up to the judge.

The other half of the downward departure cases involve non-substantial-assistance downward departures from the Guidelines. In these cases, judges impose sentences below the otherwise applicable Guidelines range because the judges found that in the case at hand "there exists an aggravating or mitigating circumstance of a kind, or to a degree, not adequately taken into consideration by the Sentencing Commission in formulating the guidelines that should result in a sentence different from that described." These exercises of downward departure

20. See infra Part II for a more detailed discussion of the numbers of sentences imposed within and outside of the applicable Guidelines ranges.

21. 18 U.S.C. §§ 3742(a)(3) & (b)(4) (1994) (authorizing appeal of any sentence below or above the range, but not within the lawfully applicable range). The maximum sentence, or top of the range, cannot exceed the minimum sentence, or bottom of the range, by more than the greater of twenty-five percent or six months. 28 U.S.C. § 994(b)(2) (2002). That choice is clearly defined and is all the nonreviewable discretion left to federal judges in sentencing in all the cases that are sentenced within the applicable guideline range.


23. There were substantial assistance departures in 17.9% of all sentenced cases in 2000. Id.


26. There were non-substantial assistance downward departures in seventeen percent of all sentenced federal cases in FY 2000. 2000 FED. SENT. SRCBK., supra note 22, at tbl.26.

27. 18 U.S.C. § 3553(b) (2003). Commentators have debated whether § 3553(b) is the only statutory provision authorizing departure, although courts—including the Supreme Court in Koon—have not taken up the invitation to use other provisions. See, e.g., Judy Clarke & Gerald McFadden, Departures from the Guideline Range: Have We Missed the Boat, or Has the Ship Sunk?, 29 AM. CRIM. L. REV. 919, 921-31 (1992) (reading other sections of
power are subject to review.  

Another set of laws, the mandatory minimum sentencing provisions that also came into effect in the late 1980s, significantly constrain judicial authority to depart downward in the narcotics cases to which they apply. In 1986, while the Commission drafted the Guidelines, Congress began another set of sentencing reforms with the passage of the Anti-Drug Abuse Act of 1986 (the Anti-Drug Abuse Act), the first in a long series of mandatory minimum sentencing statutes that have undercut the Guidelines.

The Anti-Drug Abuse Act’s provisions included statutory mandatory minimum penalties for those who trafficked in, imported, or possessed specified amounts of particular narcotics. Statutory mandatory minimums are, as their name suggests, statutory provisions that require judges to impose not less than the specified sentence upon conviction. Typically the statutes permit harsher but not lesser sentences. The decision to charge a defendant with a statute that carries a mandatory minimum is in the sole, unreviewable discretion of the prosecutor. Once a defendant has been convicted of an offense carrying a statutory mandatory minimum sentence, there are only two ways the judge can mitigate the defendant’s punishment. If the defendant is a first offender, and meets other conditions, he or she may be eligible for a modest sentence reduction through the “safety valve” provision. If the defendant does not meet those conditions, or seeks the possibility of a greater sentence reduction, cooperation in the investigation or prosecution of others, commonly known as snitching or ratting, offers the only

SRA as granting independent and broader authority to depart); Michael Gelacek et al., Departures Under the Federal Sentencing Guidelines: An Empirical and Jurisprudential Analysis, 81 MINN. L. REV. 299, 317-31 (1996) (arguing that § 3553(b) is the only statutory authority for departures).

28. For discussion of appellate practice and review of these cases in the Courts of Appeals, see infra Part VI.


30. The most important mandatory minimum sentence provisions are codified at 21 U.S.C. §§ 841, 844 & 860 and 18 U.S.C. § 924(c) (1997). There are mandatory minimum sentences for offenses other than narcotics sentences but, as the Sentencing Commission noted in 1991, of the sixty mandatory minimum statutes then on the books, the four relating to particular instances of narcotics trafficking, importation and possession and weapons offenses committed in the course of another crime accounted for the vast bulk of the convictions and sentences. U.S. SENTENCING COMM’N, SPECIAL REPORT TO CONGRESS: MANDATORY MINIMUM PENALTIES IN THE FEDERAL CRIMINAL JUSTICE SYSTEM 11 (1991) [hereinafter MANDATORY MINIMUM REPORT] (noting that four statutes setting mandatory minimum penalties for certain drug trafficking, drug importation, drug possession, and firearms offenses accounted for ninety-four percent of all statutory mandatory minimum sentences at the time of report).

31. 18 U.S.C. § 924(c) (2002) offers a counter example in its requirement that a defendant convicted of using or carrying a firearm during or in relation to any crime of violence or drug trafficking be sentenced to a five-year term of imprisonment, not more or less, to run consecutive to any other sentence imposed.

32. This flows from the general charging discretion prosecutors enjoy. See United States v. Armstrong, 517 U.S. 456, 463 (1996) (discovery on a selective prosecution claim is only required after a threshold showing that others of a different race who were otherwise similarly-situated were not prosecuted and citing cases supporting the wide charging discretion of prosecutors). Armstrong also explains that “selective-prosecution claim is not a defense on the merits to the criminal charge itself, but an independent assertion that the prosecutor has brought the charge for reasons forbidden by the Constitution.” Id.

33. 18 U.S.C § 3553(f) (2002); Sentencing Reform Act of 1984 (including safety valve provision).
Although a defendant must first decide that he or she wants to cooperate with the government, once he or she attempts to offer assistance, the decision whether or not to accept the cooperation and whether or not to make the motion that gives the judge the power to depart on the basis of cooperation once again is in the sole and unreviewable discretion of the prosecutor. While they are constrained by norms and policies, in a very real sense federal prosecutors control the application of the federal mandatory minimum sentencing statutes.

Thus, it is clear that federal criminal sentencing changed dramatically between 1985 and the turn of the twenty-first century. Several trends came together to create a complex body of overlapping laws. Prior to the reforms there was only unreviewable judicial discretion and general norms channeling that discretion. Today, federal criminal practitioners and defendants must wade through a complex set of laws to analyze the impact of the Guidelines on a case, understand and evaluate the potential for departure from the Guidelines, and account for statutory mandatory minimums that may overlap with or supercede the Guidelines in a particular case. It is a complex structure that has unintended and unfortunate consequences that will be analyzed in the remainder of this Article.

II. FEDERAL SENTENCING AT THE TURN OF THE CENTURY

One of the interesting stories that might be told about federal criminal sentencing at the turn of the century is that the prosecutors and judges who enforce our sentencing laws do not use them as aggressively as the Commission and Congress intended. More than one third of the sentences imposed are lower than (departures from) the Guidelines sentence. The rate of downward departure has increased every year. Similar trends can be found in state guidelines systems.  

34. Cooperation has both a statutory basis, see 18 U.S.C. 3553(e) (2000), which exempts cooperating defendants from the statutory mandatory minimum, and a matching Guidelines provision, U.S.S.G. § 5K1.1 (Substantial Assistance in the Investigation or Prosecution of Others), that permits judges to depart from the otherwise applicable Guidelines range.

35. Weinstein, supra note 25 at 585-91 (analyzing case law insulating prosecutorial decisions to accept or reject cooperation from judicial review).

36. The drafters of the Guidelines predicted that although courts have the legal freedom to depart, "they will not do so very often." U.S. SENTENCING COMM'N, 2002 GUIDELINES MANUAL ch. 1, pt. A(4)(b) (Guidelines Resolution of Disputed Issues, Departures) [hereinafter GUIDELINES MANUAL].

37. In 2000, there were downward departures in 34.9% of all cases. 2000 FED. SENT. SRCBK., supra note 22, at tbl.27.

38. See infra notes 161-73 and accompanying text for details on rates of departures.

39. Data from two state court systems is consistent with these general observations. Minnesota has a system of presumptive guidelines, similar in broad outline to the federal system, and shows some roughly similar overall patterns. Since the reform of that system in 1989, which made sentences longer and more predictable, sentences trended up until 1994 and then began a slow, and unsteady, decline, now about seven percent from the peak. Discretionary mitigating departures have also increased slowly and steadily and predominate over discretionary enhancements. See Minnesota Sentencing Guidelines Commission homepage, available at http://www.msgc.
The rate of strict compliance with the Guidelines is even lower than these statistics suggest. Prosecutors mitigate many sentences with charge and sentence fact bargaining, such prosecutorial bargaining is not directly observable in the sentencing data. What should we make of all this leniency in a sentencing system designed to encourage longer sentences?

The short answer is that the statistics show only apparent mitigation in many cases. While there are downward departures in more cases as time goes on, average sentence length has remained the same for most offenses. Rates of mitigation are independent of average sentence length because prosecutors and judges use their discretion to keep overall sentence length the same, while they


40. Defining terms such as "strict compliance," "evasion," and "circumvention" is a difficult task in the Guidelines area. I take the perceptive but open textured definition offered by Schulhofer and Nagel as a starting point. They note that a sentence may appear to fall within the range determined by the judge but still evade the Guidelines, while a sentencing involving a departure may be in strict compliance with the Guidelines. I join them in examining "the frequency with which cases result in a sentence different from the one that would be imposed if the entire Guidelines system were respectfully followed." Schulhofer & Nagel, supra note 3, at 1290 (claiming that, strictly speaking, Guidelines generally lead to more consistent sentencing but allow broad prosecutorial discretion). Application of this standard is imprecise and requires one to make judgments.

41. Charge bargaining occurs when prosecutors agree to accept a plea of guilty to less serious charges than those for which the defendant would have gone to trial. It is the bargaining at the heart of the plea bargaining process. Sentence fact bargaining is more peculiar to the Guidelines and involves an agreement between the prosecution and defense that the defendant will be sentenced on the basis of agreed-upon facts which will, if accepted by the court, place the defendant in a lower sentencing range than he would likely have faced after trial. For a discussion of these mechanisms and their persistence under the Guidelines, see Schulhofer & Nagel, supra note 3 (describing persistence of bargaining under Guidelines).

42. There is, however, strong evidence that prosecutors use charge and fact bargaining to control sentences. For a discussion of this phenomenon and some of the relevant literature, see infra notes 123-24 and accompanying text.

43. Along with the high rate of departures, continued increase in that rate is also contrary to the Commission's vision for the Guidelines. Departures were intended to address factors the Commission had not yet considered, but that could be studied and added to the Guidelines over time. See Guidelines Manual ch. 1, pt. A(4)(b). Thus, departures should have decreased as the system was refined over time. In fact, the Commission has made very few changes in response to departure practices.

44. This group, which will be also be referred to as "non-narcotics cases," includes fraud, theft and violent offenses. It includes about forty percent of all federal cases in 2000. 2000 Fed. Sent. Srcbk., supra note 22, at fig.A. This paper does not analyze immigration offenses, which accounted for about twenty percent of all cases in 2000. Id. Immigration cases are concentrated in the Southwest and pose their own special problems. See Frank O. Bowman, Ill & Michael Heise, Quiet Rebellion II: An Empirical Analysis of Declining Federal Drug Sentences Including Data from the District Level, 87 Iowa L. Rev. 477, 560 (2002) (observing that "U.S. Attorneys of the Mexican border districts have consciously seceded from the Guidelines regime, declaring unilaterally that local conditions entitle them to disregard national law"). I also do not address special areas of federal enforcement, such as antitrust, national defense, or civil rights, which make a very small but important part of the federal criminal justice system.
increase individualization of sentences through increased departures. In narcotics cases, however, average sentence length has been declining slowly and steadily for ten years, although narcotics sentences remain quite harsh. In this group of cases, declining sentence length is accompanied by increasing rates of downward departure. Because narcotics cases account for about forty percent of all federal prosecutions, narcotics cases overwhelm all the other offenses in the combined sentencing statistics, which show declining overall sentences because of the impact of the large group of narcotics sentences.

When we look at the narcotics and non-narcotics cases separately, we see that there is a more complicated story behind the appearance of prosecutorial and judicial sentence mitigation. It is a story of balance and imbalance among three forces: (1) prosecutorial discretion; judicial discretion; and (3) the bargaining power of criminal defendants. For most of the non-narcotics cases, prosecutors, judges and defendants have reached a rough equilibrium. First, prosecutors exercise significant control of sentence length through charging decisions and use sentencing phase mitigation to give the appearance of mitigation to induce pleas. Second, judges use their sentencing discretion in a similar fashion to help manage their dockets. Third, although defendants participate in the criminal process under compulsion, they are willing to plead guilty and accept these deals in a reasonably predictable manner. This Article thus argues that the current pattern of stable sentence length and slowly increasing rates of sentence mitigation represents one reasonable balance of forces in our criminal justice system.

Narcotics cases, however, are different. In narcotics cases, the harsh mandatory

45. Several mechanisms may keep average sentences stable as rates of downward departure increase. Charging discretion is probably the single most important factor. Prosecutors gradually charge more serious offenses to offset more departures. Judges can also moderate the magnitude of the departures to offset their overall impact and may balance departures with sentencing adjustments that do not appear in the departure statistics. For a discussion of these mechanisms see infra note 176 and accompanying text.

46. See Bowman & Heise, supra note 44, at 485 fig.1.


48. Overall downward departures have increased in narcotics cases. However, a rise in judicial departures has more than counterbalanced a trend of decreasing prosecutor-controlled substantial-assistance downward departures. See infra notes 151-54 and accompanying text.

49. In 2000, 22,253 defendants were sentenced for narcotics trafficking offenses. 2000 FED. SENT. SRCBK., supra note 22, at fig.A. That group made up forty-one percent of all sentenced defendants in 2000.

minimum statutes distort sentencing, hindering efficient and just enforcement of the narcotics laws. Mandatory minimum laws transformed prosecutorial discretion from a tool to allocate scarce enforcement resources into the driving force of the federal criminal process. Unfortunately, prosecutorial discretion is ill-suited for that role. We have lost the efficiency benefits of investigative and charging discretion; rather, prosecutorial discretion has become the unchecked power that allows individual prosecutors to choose among a vast range of punishments with almost no review, causing unwarranted disparity and excessive punishment.

Disparity results when prosecutors have unchecked power to choose among an array of applicable statutes with very different sentencing consequences. The pressure to individualize sentences leads to vastly different punishments for apparently similarly-situated defendants. The negative effects extend beyond the problem of disparity and compromise other important values. In a system in which power is unchecked and trial is no longer a realistic option, evidence is rarely tested, prosecutorial power is not challenged, and normative questions go unasked. Our American devotion to limited power and checks and balances should make us worry about a system in which the fairness of sentences is in the eye of the prosecutor.

Furthermore, whatever the systemic impact, federal narcotics sentences simply are too long. Sentences of fifteen or twenty years and longer have become all too common for defendants who play relatively minor roles in large drug conspiracies. There is little evidence that we are winning the war on drugs this way but a great deal of evidence that we are destroying the lives of thousands of people.

Justice Breyer's comments bolster the view that narcotics sentences under the mandatory minimum laws are simply too harsh and that they are at the core of the

51. A straightforward example is a cooperating defendant who receives a two-year sentence when he testifies against an equally culpable co-defendant, who then receives a fourteen-year sentence (and would have been convicted without the cooperator's testimony).

52. The power of the American prosecutor traditionally has been limited by a range of factors. See, e.g., William T. Pizzi, Understanding Prosecutorial Discretion in the United States: The Limits of Comparative Criminal Procedure as an Instrument of Reform, 54 OHIO ST. L.J. 1325, 1336-52 (1993) (describing how range of informal and indirect controls, including checks inherent in adversary system, elections and localism). Professor Daniel Richman has noted the tendency to check even within the law enforcement establishment itself: "After all, the entire American criminal Justice system is characterized by an almost instinctive embrace of fragmented authority, with the tensions between police and prosecutors, attorneys general and district attorneys seen as a virtue, rather than a vice." Daniel Richman, Federal Criminal Law, Congressional Delegation, and Enforcement Discretion, 46 UCLA L. REV. 757, 807-08 (1999) (arguing that Congress exercises greater control over federal law enforcement through funding, oversight hearings and other bureaucratic mechanisms than by limiting or tailoring substantive statutes, but these controls only work in some areas of enforcement and tend to be motivated less by policy preferences and more by the desire to curb executive power).

53. This is both a normative and practical problem. It is simply wrong to impose overly harsh, disparate sentences and it is inefficient. At the very least, widely varying sentences for similar offenses are unlikely to be an optimal distribution for deterrence and, as I have argued elsewhere, investigative and charging discretion exercised without feedback does not enjoy the benefit of a self-regulating system. Weinstein, supra note 25 (arguing that all the costs of cooperation are externalized and prosecutors have no incentives to use this tool efficiently).
general problem with narcotics sentences. Justice Breyer, concurring in *Harris v. United States*,\(^54\) noted that “[m]andatory minimum statutes are fundamentally inconsistent with Congress’ simultaneous effort to create a fair, honest and rational system...”\(^55\) Citing remarks by two other Supreme Court Justices, a Senator, and academic commentary, Justice Breyer observed that mandatory minimum sentences increase disparity, result in unfair sentences, transfer power to prosecutors, and encourage subterfuge.\(^56\)

However, the larger problem with mandatory minimum statutes is their basic incompatibility with the fundamental structures that undergird our criminal justice

---

55. See id.
56. Concurring in the Court’s judgment that *Apprendi* does not apply to all statutory mandatory minimum cases, Justice Breyer wrote:

In saying this, I do not mean to suggest my approval of mandatory minimum sentences as a matter of policy. During the past two decades, as mandatory minimum sentencing statutes have proliferated in number and importance, judges, legislators, lawyers, and commentators have criticized those statutes, arguing that they negatively affect the fair administration of the criminal law, a matter of concern to judges and to legislators alike. See, e.g., Remarks of Chief Justice William H. Rehnquist, NAT. SYMPOSIUM ON DRUGS AND VIOLENCE IN AMERICA 9-11 (June 18, 1993); Kennedy, *Hearings before a Subcommittee of the House Committee on Appropriations*, 103d Cong., 2d Sess., 29 (Mar. 9, 1994) (mandatory minimums are “unprudent, unwise and often an unjust mechanism for sentencing”); Breyer, *Federal Sentencing Guidelines Revisited*, 14 CRIM. JUSTICE 28 (Spring 1999); Hatch, *The Role of Congress in Sentencing: The United States Sentencing Commission, Mandatory Minimum Sentences, and the Search for a Certain and Effective Sentencing System*, 28 WAKE FOREST L. REV. 185, 192-96 (1993); Schulhofer, *Rethinking Mandatory Minimums*, 28 WAKE FOREST L. REV. 199 (1993); Raeder, *Rethinking Sentencing and Correctional Policy for Nonviolent Drug Offenders*, 14 CRIM. JUSTICE 1, 53 (Summer 1999) (noting that the American Bar Association has opposed mandatory minimum sentences since 1974).

Mandatory minimum statutes are fundamentally inconsistent with Congress’ simultaneous effort to create a fair, honest, and rational sentencing system through the use of Sentencing Guidelines. Unlike Guideline sentences, statutory mandatory minimums generally deny the judge the legal power to depart downward, no matter how unusual the special circumstances that call for leniency. See *Melendez v. United States*, 518 U.S. 120, 132, 135 L. Ed. 2d 427, 116 S. Ct. 2057-133 (1996) (Breyer, J., concurring in part and dissenting in part); cf. *Koon v. United States*, 518 U.S. 81, 95, 135 L. Ed. 2d 392, 116 S. Ct. 2035-96 (1996). They rarely reflect an effort to achieve sentencing proportionality—a key element of sentencing fairness that demands that the law punish a drug “kingpin” and a “mule” differently. They transfer sentencing power to prosecutors, who can determine sentences through the charges they decide to bring, and who thereby have reintroduced much of the sentencing disparity that Congress created Guidelines to eliminate. U.S. SENTENCING COMM’N, SPECIAL REPORT TO CONGRESS: MANDATORY MINIMUM PENALTIES IN THE FEDERAL CRIMINAL JUSTICE SYSTEM i-iv, 31-33 (1991) (SENTENCING REPORT); see also Schulhofer, *supra*, at 214-20. They rarely are based upon empirical study. See Rehnquist, *supra*, at 9-10; Hatch, *supra*, at 198. And there is evidence that they encourage subterfuge, leading to more frequent downward departures (on a random basis), thereby making them a comparatively ineffective means of guaranteeing tough sentences. See SENTENCING REPORT 53.

Id.
system: prosecutorial discretion, overcriminalization, and the counterbalancing force of the adversarial model along with the Sixth Amendment right to trial. If we view sentencing as the final phase in the process of criminal law enforcement rather than an end in itself, it is apparent that harsh mandatory sentencing cannot work in our system. In short, criminal statutes designed to overcriminalize in the service of discretionary enforcement cannot and should not be rigidly and uniformly enforced. The high maximum sentences and profusion of crimes and theories of liability are intended to give prosecutors a wide range of choices and to promote efficiency. We cannot rigidly enforce these laws in every case to which they apply because there are simply too many cases. We should not try to enforce these laws in every case to which they apply, as justice will not be served by rigidly applying very broad laws. Yet the alleged benefits of mandatory minimum sentencing statutes—certainty, greater deterrence, and a version of fairness in which everyone gets exactly the same punishment without regard to individual circumstance—would accrue only if the mandatory minimum laws could be enforced rigidly against every wrongdoer. Because we cannot punish every wrongdoer, we should not attempt to punish some wrongdoers excessively. Moreover, experience demonstrates that in fact we are nowhere close to applying mandatory minimum laws uniformly, even to those wrongdoers who are apparently similarly-situated.

Importantly, the structural incompatibility of the Guidelines and statutory mandatory minimum sentences is not merely theoretical. The problems play out in the everyday enforcement of federal criminal law. As we will see, there are important differences between the sentencing practices for offenses that do not carry mandatory minimum sentences and those that do. Those differences can be seen in statistics and case studies that show federal criminal law working reasonably well with the Guidelines in many instances where the mandatory minimums do not apply. Sentencing for most categories of federal crime is characterized by stable average sentence lengths and slowly increasing apparent

57. The Sentencing Commission has been making a version of this argument for years. See generally MANDATORY MINIMUM REPORT, supra note 30 (noting incompatibility between Guidelines and mandatory minimum sentencing statutes). The Supreme Court reaffirmed its commitment to prosecutorial discretion to invoke mandatory minimum statutory penalties in Harris, by holding that most sentencing factors that do not raise the sentence above the statutory maximum need not be alleged in the indictment and proven beyond a reasonable doubt to a jury or fact-finding judge. Although the case speaks predominantly in the language of judicial discretion, it notes that “the political system may channel judicial discretion . . . by requiring defendants to serve minimum terms after judges make certain factual findings.” Harris, 563 U.S. at 545. In narcotics cases, factual findings about the quantity of drugs involved become relevant when prosecutors charge statutory subsections that turn on the quantity of narcotics. See infra notes 76-77 and accompanying text.

58. Overcriminalization describes the American practice of legislative adoption of broad proscriptions which permit prosecutors to use their charging discretion to select enforcement priorities. See infra Part III for a fuller discussion of overcriminalization.

59. See infra Part IV for a discussion of how enforcing the rights of criminal defendants can counterbalance and check prosecutorial discretion and power.
mandatory minimums and narcotics sentencing

mitigation. That combination is a sign of balance among important systemic forces. Narcotics sentences, wherein mandatory minimums are at play, however, are not stable over time and sentencing practices are unstable. A closer look reveals an area rife with excessive punishment and injustice.

Problems of lack of stability in narcotics sentences are of recent origin and are the direct result of the changes in the past fifteen years discussed above.\(^6\) The steady march of prosecutorial discretion, coupled with harsh statutory mandatory minimum penalties shifted the focus of federal criminal practice. A focus on sentencing litigation, as opposed to litigating Fourth-, Fifth-, and Sixth-Amendment issues in pretrial motions or litigating guilt or innocence at trial, dominates federal criminal law practice to a degree unimaginable since the late 1980s. Of course, guilty pleas have long been the norm; however, from the late 1960s to the mid 1980s, pretrial litigation was the strategic focus of many cases.\(^6\) Motions to suppress evidence and to challenge indictments were important strategic devices for many defense lawyers. Such motions were used to learn about the evidence, to educate the judge about the favorable aspects of the case, and to gain bargaining leverage.

In the late 1980s, the balance of power shifted among defense lawyers, prosecutors and judges. Mandatory minimum sentences began to have real bite while suppression motions became harder to win as the Supreme Court stepped back from the rights revolution of the mid 1960s and 1970s. Thus, pretrial motions faded in strategic importance. This shift manifested not only in defendants' diminishing likelihood of winning pretrial substantial relief in the form of motions, but also in the fact that the tactical advantages to motion practice faded as judges lost sentencing power to prosecutors. There was less to be gained by showing the judge the weakness of the prosecutor's case and more to be lost by challenging the prosecutor. Most importantly, prosecutors came to realize just how powerful a club they wielded as discretion and authority shifted to them from the judges. By the mid 1990s, as illustrated in Part III, federal prosecutors could realistically threaten many defendants with very long sentences. In that environment, most will not gamble on a trial, pretrial motion, or any litigation that would jeopardize defen-

---

\(^6\) For a discussion of the strategic use of motion practice, see generally Harry Subin, Chester Mirsky & Ian Weinstein, Federal Criminal Practice: Prosecution and Defense § 16.3 (Examination of Witnesses-Tactical Considerations) (1992) (discussing examining witnesses at a hearing on pretrial motions and citing A. Amsterdam, Trial Manual for the defense of Criminal Cases §§ 223, 265-266 (1984)).

\(^6\) This shift did not happen all at once but began in the late 1980s, as mandatory minimum narcotics statutes began to reappear. For the history of these penalties, see Linda Drazga Maxfield & John H. Kramer, U.S. Sentencing Comm'n, Substantial Assistance: An Empirical Yardstick Gauging Equity in Current Federal Policy and Practice 11 (1998). At first the statutes did not apply to conspiracy charges and were frequently bargained away. By the early 1990s, the mandatory minimum statutes were expanded and could be applied to many narcotics cases.
dants' limited chances for the sentence mitigation the prosecutor controls. At the same time, the SRA created the now-vibrant field of sentencing litigation. Before the SRA and the Guidelines, sentencing discretion was almost entirely unguided and unreviewable. There were virtually no legal claims that could be raised at sentencing and no real prospect of appellate review. Today, thousands of sentencing appeals are heard each year in the United States Courts of Appeals, and sentencing litigation under the Guidelines is a major feature of federal criminal practice.

However, it was not this way fifteen years ago and it need not be this way forever. We can return to a better balance of power between the prosecutor and the defendant and step away from a system in which we pay all too little attention to whether a defendant should be punished and focus much of our resources on determining how long the punishment should be. It is past the time to heed the lessons of injustice and defendants' lack of bargaining power and to improve on and reform the tremendous changes to sentencing practices of the late 1980s.

III. HOW CHOOSING WHOM TO INVESTIGATE AND WHOM TO CHARGE CAME TO CONTROL NARCOTICS SENTENCING

Analysis of American criminal justice begins with a look at overcriminalization and prosecutorial discretion. Our legislatures have long criminalized much more conduct than can be sanctioned effectively. Overcriminalization is a central structural feature of American criminal justice and one of its great traditional strengths. It permits local officials effectively to allocate and reallocate enforcement and prosecutorial resources to respond to changing circumstances by adjusting priorities within the authority granted by the broad legislative criminalization, or overcriminalization. It is within this framework that American police and prosecutors have been granted virtually unreviewable authority to allocate.
investigative and prosecutorial resources.69

Meanwhile, until the late 1980s, the federal prosecutor's discretion to determine whom to investigate and when and what charges to bring was counterbalanced in sentencing by judges, who determined the actual sentence to be imposed. This separation of control between the power to charge and the power to sentence was a significant check on prosecutorial power. Three related and well-recognized trends of the last twenty years have greatly eroded the separation between charging and sentencing and have shifted power both to charge and to sentence to prosecutors, particularly in federal narcotics cases. The first trend is accelerating overcriminalization. Federal authorities now have jurisdiction over a greater range of conduct and can bring charges from an ever-expanding menu of more specific and serious charges. The second trend is the emergence of mandatory minimum sentencing statutes that have turned the power to charge into the power to dictate the sentence. Charging and sentencing are much more tightly bound than they were under the old law. A third trend, increasingly harsh sentences, has made the increase in discretion more significant because prosecutors not only have more control over cases, but they also wield a bigger stick when they exercise that authority.

A. Accelerating Overcriminalization in a World of Harsh Mandatory Minimum Statutes

1. Accelerating Overcriminalization

Criminal codes get longer as every societal ill is met with a new crime. William Stuntz estimates that the federal criminal law, once viewed as an area of limited and specialized application, proscribes more than 2000 different acts, from securities fraud to carjacking.70 State codes display a similar pattern.71 Prosecutorial discretion works to allocate resources in a system of overcriminalization.72 We

69. There are supposed to be political limits on this authority. Prosecutors who abuse their power, or use it ineffectively, will lose office when either they, or their executive bosses, are voted out of office. These political controls are also supposed to give defendants some leverage against prosecutorial power. They, or their counsel, can make a public issue of abuse of power to invoke these political controls. The power of these controls has received critical attention. See Sara Sun Beale, What's Law Got To Do With It? The Political, Social, Psychological And Other Non-Legal Factors Influencing The Development Of (Federal) Criminal Law, 1 BUFF. CRIM. L. REV. 23, 42-53 (1997) (cataloguing public opinion influences on criminal sentencing); Richman, supra note 52 (arguing that Congress exercises more control over federal prosecutorial discretion than is usually recognized through hearings, informal contacts, control of high level appointments and other mechanisms).

70. Stuntz, supra note 66, at 514-15.

71. Id.

72. For an interesting look at prosecutorial discretion, see generally Pizzi, supra note 52 (using comparative law to explore the role of the American prosecutor and arguing that reforms on civil law model cannot work because of our commitment to adversary justice, harsh criminal statutes and political arrangements).
recognize that there is more wrongdoing than the system effectively can address and pass many criminal laws for at least two reasons. First, the mere proscription of conduct, even if unenforced, will discourage some wrongdoing by those who choose to obey the law without coercion. Second, we want to give the people with the best access to information sufficient tools to deal with current problems. If there is a rash of bank robberies, that crime should get more attention. When public concern moves to fraud, so too can the resources. The essential idea behind our system of discretion and overcriminalization is that criminals should know that their conduct may be punished at any time, even if we cannot punish everyone all the time.

Of course, as we criminalize more conduct, prosecutors can exercise more authority. This is particularly true when a criminal code offers more ways to charge any one given instance of bad conduct. Federal narcotics laws offer an example of how potential charges for a given misdeed have expanded in recent years. In the early 1980s, a defendant arrested for distributing narcotics faced the possibility of being charged with the substantive offense of distribution and, if others were involved, as is almost always the case in narcotics distribution, with conspiracy. The narcotics laws changed dramatically in the late 1980s and early 1990s. Today, that same defendant faces those two possibilities, with mandatory minimum sentencing if the quantity of drugs is sufficient, as well as a range of charges for money laundering, use of a gun in connection with a narcotics offense, distributing narcotics near a school or truck stop, being a repeat

73. The strategic use of the power to investigate and charge a variety of offenses may first have come to public attention in the prosecution of Al Capone. When federal authorities could not build a traditional murder or bootlegging case, they prosecuted that notorious gangster for tax evasion. He served his sentence at Alcatraz and died in prison just the same as if he had been sentenced for his more serious transgressions.

74. 21 U.S.C. § 841 (2002) ("Except as authorized by this subchapter, it shall be unlawful for any person knowingly or intentionally ...

75. 21 U.S.C. § 846 (2002) ("Any person who attempts or conspires to commit any offense defined in this subchapter ... ").

76. For a brief history of the anti-drug sentencing reforms of the 1980s, see Michael A. Simons, Departing Ways: Uniformity, Disparity And Cooperation In Federal Drug Sentences, 47 VILL. L. REV. 921, 928-30 (2002).

77. 21 U.S.C. § 841(b) (2002). The mandatory minimum provisions were added in 1986. See supra Part II.

78. The money laundering statute, 18 U.S.C. § 1956 (2002), was added to the federal code in 1986, Anti-Drug Abuse Act of 1986, Pub. L. 99-570, Title XIII, § 1352(a), 100 Stat. 3207 (1986), and has been repeatedly amended since then.

79. 18 U.S.C. § 924(c) (2002). This statute imposes a five-year mandatory sentence, consecutive to any other sentence, for the use of a weapon in connection with a narcotics offense.

80. 21 U.S.C. § 860 (2002). This statute, formerly codified as 21 U.S.C. § 845a, was part of the Anti-Drug Abuse Act of 1986, see supra notes 29-30, and has been repeatedly amended. It currently imposes higher maximum sentences and a mandatory minimum, regardless of the quantity of drugs on defendants convicted of offenses occurring at or near schools, playgrounds, youth centers, video arcades and swimming pools. Id.

81. 21 U.S.C. § 849 (2002). This statute was passed in 1994, Pub. L. 91-513, Title XII, § 409, Title XVIII, § 180201(b)(1), 108 Stat. 2046 (1994), and doubles or triples the maximum sentence for defendants convicted of a narcotics offense at or near a truck stop or safety rest area.
narcotics offender, running a continuing narcotics enterprise, investing drug profits, and distributing to, or employing, someone under eighteen years of age in a narcotics offense. Most narcotics defendants could be charged in at least a half a dozen different ways, each of which would carry vastly different sentencing implications.

2. The Role of Reviewable Guidelines and Mandatory Minimum Statutes

As federal narcotics statutes expanded, the sentencing reforms of the 1980s gave prosecutors more power to control sentences. The introduction of mandatory minimum sentences, and to a lesser degree, the Guidelines, gave control over federal narcotics sentencing to prosecutors. The importance of this shift cannot be overstated. Before the Guidelines were widely enforced and mandatory minimum sentences got serious in the early 1990s, federal criminal sentencing was characterized by almost completely individualized and unreviewable judicial sentencing discretion. The length of a federal criminal defendant’s sentence was almost entirely in the control of the district judge, although a negotiated plea agreement could cap the possible sentence with a chosen statutory maximum, and prosecutors and defense lawyers developed local customs to influence their judges. In most cases, however, judges could impose any sentence from an

82. 21 U.S.C. § 841(b) (2002). The statute doubles the maximum sentence and the mandatory minimum sentence for offenders with prior narcotics convictions. There are strict procedural requirements for the invocation of the enhanced mandatory minimums for prior offenders, codified at 21 U.S.C. § 851 (2002).


85. For an illustrative example of the range of charging options and the sentencing consequences of different charging strategies, see infra Part III.B.

86. The development of the Guidelines and the proliferation of mandatory minimum statutes in the late 1980s and early 1990s is discussed above in Part I.

87. Mistretta v. United States, 488 U.S. 361, 363 (1989) (commenting on extent of pre-Guidelines discretion in introduction to decision affirming constitutionality of Guidelines). There was also the restraint of the very weak doctrine of proportionality review, see Solem v. Helm, 463 U.S. 277, 284-85 (1983), but this was a very vague boundary.

88. The Guidelines have changed, but not eliminated, sentence manipulation through plea agreements. Schulhofer & Nagel, supra note 3 (arguing that Guidelines have made charging decisions more consistent, but prosecutors retain a considerable degree of sentencing discretion). For discussions of pre-Guidelines plea bargaining, see generally Alschuler, supra note 3 (criticizing plea bargaining); Utz, supra note 3 (defending plea bargaining).

89. Variations in local practice are a persistent and important feature of our legal system. When I practiced in the District of Columbia Superior Court, it was common for prosecutors to explain on the record that they would “take no position” on negotiated sentencing matters. In some cases, however, the prosecutor would “consent to” or “join” the defense request. Judges understood that in the first case the defendant was getting the benefit of a bargain and many judges would honor the agreement absent special circumstances. In the second set of situations most judges consistently honored what they understood as the prosecutor’s affirmative request. In the Southern District of New York the practice more often involved a tacit or explicit agreement not to reveal some sentencing facts to the judge or the probation officer, each of whom in turn tended not to seek information beyond the matters presented by the parties.
unsupervised suspended sentence to the statutory maximum. As discussed above, a prosecutor could bring a single charge of distribution and the defendant could receive twenty years in prison or the prosecutor could bring a one hundred-count indictment charging every conceivable crime and the defendant could receive a suspended sentence.

As discussed above, the Guidelines and the mandatory minimum statutes dramatically shifted the power to control the length of a federal defendant’s sentence. Prosecutors are now the dominant force in determining the length of the defendant’s sentence, particularly in narcotics cases. The prosecutor’s unreviewable decision to charge or not to charge statutory mandatory minimum provisions, prior offender provisions and other related offenses (such as money laundering related to a narcotics offense) essentially determine the length of the sentence. Judges have no authority over those charging decisions, so defendants have nowhere else to look for mercy but to the prosecutor.

3. Harsh and Unfair Sentencing

As more statutes gave prosecutors more power to control sentencing, the sentences also grew much longer and increasingly unfair. One explicit goal of the sentencing reforms of the 1980s was to increase the severity of federal sentences, particularly for narcotics offenses. This goal has been achieved. Between 1980 and 1995, the average narcotics sentence imposed by federal judges almost doubled, from about forty months to eighty months, despite the abolition of parole. As defendants were typically paroled after serving about forty per cent of the sentences imposed under the old law, the time defendants actually served more than tripled between 1980 and 1995. Although incomplete data prevents comparison of actual sentences served going back to 1980, actual time in prison for narcotics offenses tripled from twenty-seven months in 1984 to seventy-nine months in 1993. Sentences for violent crimes doubled in that period and sentences for firearms offenses nearly tripled.

90. For the authority to suspend sentence under the old law, see supra note 12.
91. Stith & Koh, supra note 2, at 284-87 (summarizing evidence that legislative compromises that made passage of SRA possible were intended to increase the severity of federal sentences under Guidelines).
92. For the most thorough discussion of average sentence lengths in narcotics cases, see Frank O. Bowman, III & Michael Heise, Quiet Rebellion? Explaining Nearly a Decade of Declining Federal Drug Sentences, 86 IOWA L. REV. 1043, 1067-74 (2002); Bowman & Heise, supra note 44; see also Hofer & Semisch, supra note 47.
93. Hofer & Semisch, supra note 47 at 16, fig.7.
94. Sentences increased by the imposition of mandatory minimums on most narcotics offenses and imposing guideline ranges that raised sentences for almost all offenses relative to the pre-Guidelines averages. The Guidelines drafter also had a consistent preference for sentences of imprisonment, rather than probationary sentences. Probationary sentences for narcotics and violent offenses have virtually disappeared. Probation only and probation and alternative sentences accounted for about eighteen percent of all pre-Guideline narcotics sentences and about five percent of all Guideline narcotics sentences. Non-jail sentences accounted for about twelve percent of all pre-Guidelines sentences for violent offenses and about five percent of all Guidelines sentences for violent offenses. For white-collar offenses the percentage of straight probation sentences has
Another goal of sentencing reform was to decrease unwarranted disparity among sentenced defendants. Of course, there remains very significant debate about the difference between warranted and unwarranted disparity. In practice, however, sentence length is somewhat more predictable today. Differences in sentence length from defendant to defendant turn less on relative culpability and more on the scope of the overall offense in which the defendant was involved. As part of the Guidelines’ quest for sentencing factors easily reduced to quantifiable and therefore readily compared scales, the central determinant of sentence length in drug cases is the weight of the narcotics reasonably foreseeable to the defendant who participates in a deal or conspiracy. Thus, a defendant who does no more than help to unload a truck with a ton of cocaine starts at the same sentence level as those who arranged the shipment or who negotiated the sale or purchase of the drugs. Although the Guidelines permit some differentiation among participants, a first-time offender paid one hundred dollars to unload those boxes can face a ten-year mandatory minimum and a guideline sentence of 151 to 188 months following trial. Proof that the defendant who unloaded the boxes also

95. Congress sought reasonable uniformity in sentencing by narrowing the wide disparity in sentences imposed for similar criminal offenses. GUIDELINES MANUAL ch. 1, intro., A.3 (The Basic Approach).


97. This framework also applies to cases involving fraud or theft, where the amount of money the defendant stole, or attempted to steal, drives the Guidelines. Thus, an attempt to use worthless bonds with a face amount of $10,000,000 to secure a loan can result in a sentence range of sixty-three to seventy-eight months for a first time offender, in a case in which the loan was never made and there was no loss of money. The actual theft of $15,000 would make a first time offender eligible for probation or a sentence of six to twelve months. Neither sentence involves a mandatory minimum and would be subject to a number of potential sentence adjustments and departures. Reasonable people can disagree about which case is more serious, but the Guidelines equate willingness to harm, planning to harm, agreeing to harm, and actually harming as equally culpable. See U.S.S.G. § 2B1.1. See generally Frank O. Bowman, III, The 2001 Federal Economic Crime Sentencing Reforms: An Analysis and Legislative History, 35 IND. L. REV. 5 (2001) (analyzing the amended economic crimes guideline and discussing actual, intended and potential harms).


100. The base offense level would be thirty-eight. U.S.S.G. § 2D1.1. Assuming such a defendant received a four level adjustment as a minimal participant, id. § 3B1.2, his offense level would be thirty-four and the sentence range would be 151-188. GUIDELINES MANUAL ch.5, pt. A (Sentencing Table).
drove another participant to a meeting and made a supportive comment could increase the range to 235 to 293 months. A defendant with a significant criminal history who is the central figure in a plan to sell one kilogram of cocaine could face a sentencing range of seventy-eight to ninety-seven months after trial. Justice is not served by relying on the quantity of drugs as the central element in allocating punishment.

Federal narcotics sentencing is now characterized by harsh and often inequitable sentences. Discussion of the relative increase over the past fifteen years obscures the real impact of the changes. The average time actually served by federal narcotics defendants increased by a factor of three from 1984 to 1992, but defendants’ life spans did not change. Their families do not age any more slowly. Thus, the real-life impact for those who face these sentences and their families must be considered and cannot be overstated.

B. The Scope of Prosecutorial Power to Control Narcotics Sentences

These three trends—accelerating overcriminalization, increasing use of mandatory minimum statutes, and harsh penalties—give prosecutors enormous control over narcotics sentencing. An example illustrates this point: Consider Mr. X, a defendant arrested in Manhattan for possession of five grams of crack cocaine. Assume that Mr. X has a prior state conviction for sale of a small amount of heroin and that his co-defendant Mr. Y kept a gun in a drawer in the kitchen of the apartment in which the two were arrested. Suppose the government alleges that Mr. X was a courier for Mr. Y and that Mr. Y had been observed at the arrest location for six months and is believed to sell about fifty grams of crack a week.

Mr. X could be charged under three different subsections of 21 U.S.C. § 841, which prohibits narcotics trafficking. Under § 841(b)(1)(C), which is silent as to the quantity of drugs, the maximum penalty would be twenty years in prison and there would be no mandatory minimum. If charged in this way, Mr. X’s sentence would be up to the judge, within the strictures of the Guidelines, including the

101. This assumes that same defendant receives no adjustment for role in the offense and is sentenced based on an offense level of thirty-eight. Id.
102. The base offense level in this case would be would be twenty-six. U.S.S.G. § 2D1.1. Assuming this defendant received no adjustments and was in criminal history category III, id. § 4A1.1 (Criminal History Category), his sentence range would be seventy-eight to ninety-seven months. GUIDELINES MANUAL ch. 5, pt. A (Sentencing Table).
103. See generally David Cole, As Freedom Advances: The Paradox Of Severity In American Criminal Justice, 3 U. PA. J. CONST. L. 455, 458 (2001) (noting that United States imposes far harsher sentences than other Western countries, and further noting that this disparity is most pronounced with respect to property and drug crimes).
104. Hofer & Semisch, supra note 47, at 17.
106. The bottom of the Guidelines base offense level for five grams of crack, before any upward or downward adjustments or increases for criminal history, is sixty-three months, see U.S.S.G. § 2D1.1 & GUIDELINES MANUAL
possibility of departure. The sentence could reasonably range from two to ten years.\textsuperscript{107}

Another charging possibility is an indictment under § 841(b)(1)(B).\textsuperscript{108} In this case the charging language would specify that Mr. X possessed with intent to distribute five grams or more of a mixture or substance containing cocaine base. Conviction under this subsection would carry a forty-year maximum sentence but, more importantly, conviction would trigger a five-year mandatory sentence and would put any sentence below five years beyond the reach of judicial discretion.\textsuperscript{109}

The third possibility under § 841 is a charge under subsection (b)(1)(A).\textsuperscript{110} The prosecutor could charge that Mr. X possessed fifty grams or more of crack on the theory that he was engaged in this activity for an extended period of time. Conviction of this charge would carry a maximum sentence of life in prison and a ten-year mandatory minimum.\textsuperscript{111}

Mr. X could also be charged with committing this offense within 1000 feet of a school zone (only two places in all of Manhattan and the Bronx have zones that are not within 1000 feet of a school zone), invoking a one-year mandatory minimum.\textsuperscript{112} The government could also bring an indictment charging the use of a gun in a narcotics offense,\textsuperscript{113} which would carry an additional five-year mandatory

\begin{footnotes}
\item[107] Starting at offense level twenty-six for five grams of crack, U.S.S.G. § 2D1.1, and criminal history category II, id. § 4A1.1 (Criminal History Category), a judge would likely grant the three-level downward adjustment for acceptance of responsibility, id. § 3E1.1 (Acceptance of Responsibility), and could reduce the offense level four more points for minimal role, id. § 3B1.2. The range would then be thirty-three to forty-one months before considering possible grounds for downward departure. Practices vary widely, but sentences such as this are not uncommon in at least one district with which I have personal experience.
\item[109] There is, however, an exception under the first offender safety-valve. See supra note 33. Of course cooperation, at the option of the prosecutor, see 18 U.S.C. 3553(f) & U.S.S.G. § 5K1.1, is the most common device for mitigating the impact of statutory mandatory minimum sentences.
\item[111] In my experience, the two neighboring federal districts in New York City offer an example of contrasting uses of this statute and the similarly structured proscription on importation of narcotics. 21 U.S.C. § 960 (2002). The United States Attorney's Office for the Eastern District of New York, which has jurisdiction over Kennedy International Airport, manages its very heavy docket of narcotics importation cases by bringing indictments that do not allege any particular quantity of drugs. These cases do not involve statutory mandatory minimum sentences. That office routinely agrees to downward adjustments and the sentences imposed are often significantly lower than a strict reading of the Guidelines would suggest. Across the East River, the United States Attorney's Office for the Southern District of New York routinely indicts couriers on charges carrying mandatory minimums and negotiates to preclude defendants' requests for many of the same downward adjustments. Sentences for couriers carrying the same weight of drugs are usually two to four years in Brooklyn but often eight to ten years in Manhattan.
\item[113] 18 U.S.C. § 924(c) (2002). This fact pattern presents proof issues, as the government would have to show that Mr. X had more than general knowledge of the gun. See United States v. Santeramo, 45 F.3d 622, 624 (2d Cir. 1995), citing with approval United States v. Powell, 929 F.2d 724 (D.C. Cir. 1991) (holding that accomplice must "know to a practical certainty" that the principal would be carrying a gun). See generally Tyler B. Robinson, A
sentence to run consecutive with any other sentence. Also possible is an indictment charging Mr. X as a second offender using a gun in relation to a narcotics crime involving five grams or more of cocaine base, in which case the mandatory minimum would be fifteen years.\textsuperscript{114}

A separate area of prosecutorial control of sentencing involves advocacy around the issue of relevant conduct. Under the Guidelines,\textsuperscript{115} the prosecutor might argue at sentencing that Mr. X should be held accountable for the total quantity of drugs sold by Mr. Y, regardless of the language of the indictment. That would raise the amount of drugs involved to over a kilogram of crack and the corresponding guideline range from about five years to about fifteen years. Because this would be a Guidelines issue as opposed to a mandatory minimum statute question, a sentence enhanced by relevant conduct could be moderated by judicial discretion through adjustments and departures. The prosecutor, however, chooses among these options for controlling the sentence within the statutory and/or Guidelines schemes.

Some may argue that I have overstated the reach of prosecutorial power;\textsuperscript{116} rather, I acknowledge that such power does have bounds. Among the most commonly discussed constraints are prosecution guidelines or policies that limit or channel discretion.\textsuperscript{117} For example, Justice Department policy, as set out in what

---

\textsuperscript{114} He would be subject to a ten-year mandatory minimum as a second time narcotics offender under 21 U.S.C. § 841(b)(1)(B) (2002) and a five-year mandatory consecutive sentence under 18 U.S.C. § 924(c) (2002).

\textsuperscript{115} U.S.S.G. § 1B1.3 (Relevant Conduct). This section permits a judge to sentence the defendant on the basis of all the conduct in a common scheme or plan. In narcotics prosecutions a defendant may stand convicted of a particular substantive act and be sentenced on the basis of the total quantity involved in the course of drug dealing, so long as the sentence does not exceed the statutory maximum. See Edwards v. United States, 523 U.S. 511, 513-14 (1998) (concluding that judge determines whether conduct is part of same course of conduct). See generally Paul M. Secunda, Cleaning Up the Chicken Coop of Sentencing Uniformity: Guiding the Discretion of Federal Prosecutors Through the Use of the Model Rules of Professional Conduct, 34 Am. Crim. L. Rev. 1267 (1997) (discussing how relevant conduct rules may be used by prosecutors to exercise discretion).

\textsuperscript{116} For commentators who argue against, or seek to moderate, the claim that sentencing power has shifted to prosecutors, see Frank O. Bowman, III, The Quality of Mercy Must be Restrained, and Other Lessons in Learning to Love the Federal Sentencing Guidelines, 1996 Wis. L. Rev. 679, 724-26 (arguing that the shift is not as great as some suggest); James B. Burns et al., We Make the Better Targets (But the Guidelines Shifted Power from the Judiciary to Congress, not from the Judiciary to the Prosecution), 91 Nw. U. L. Rev. 1317, 1318 (1997) (arguing that Guidelines have shifted power from Judiciary to Congress, rather than from Judiciary to Executive and prosecutors are under the control of Executive branch).

\textsuperscript{117} For a discussion of the impact of charging guidelines, see David Robinson, Jr., The Decline and Potential Collapse of Federal Guideline Sentencing, 74 Wash. U. L.Q. 881, 905-08 (1996) (arguing that change in Department of Justice charging guidelines from the more restrictive Thornburgh Memorandum to the more lenient Reno Bluesheet introduced another source of unwarranted disparity). Those directives appear to have little impact beyond their occasional service as post hoc justifications. For a general, if dated, discussion of the problems with developing and enforcing policies to control prosecutorial discretion, see generally Norman Abrams, Internal Policy: Guiding the Exercise of Prosecutorial Discretion, 19 UCLA L. Rev. 1 (1971) (arguing that enforcing prosecutorial guidelines will always be limited by prosecutorial discretion, the very phenomenon sought to be regulated). For a more sympathetic view of the utility of prosecution guidelines and policies, see
has come to be known among line prosecutors and defense lawyers as the Thornburgh Memorandum,118 requires prosecutors to charge all the "most serious readily provable crimes."119 This important policy directive, along with strong traditions and institutional habits, channels important aspects of prosecutorial discretion;120 however, the Thornburgh Memorandum does not control the central force of prosecutorial discretion in the American criminal justice system.

The investigative decision is largely untouched by policies that regulate charging. Among the group of all wrongdoers, some are never the subject of investigation.121 Some investigations will reveal only a portion of the culpable conduct, while others continue until every bad act is uncovered. One cannot claim that a given defendant has been treated similarly to others who engaged in the same conduct but were never investigated or were not investigated with the same vigor. Yet charging policies cannot address that problem. Even among those similarly charged, it is often impossible to determine whether additional investigation would reveal other wrongful conduct. For example, most narcotics dealing must involve some element of money laundering, but the government has no obligation to investigate that offense in every narcotics case.

Even in well-investigated and carefully-charged cases, the high demands of legal proof require prosecutors to exercise discretion in making charging deci-

Michael Simons, Prosecutorial Discretion and Prosecution Guidelines: A Case Study in Controlling Federalization, 75 N.Y.U. L. Rev. 893, 963 (2000) (arguing that guidelines controlling when a case should be federally prosecuted can address some of the dangers of federalization of state crimes).


120. For further discussion of traditional and institutional constraints on prosecutorial discretion, see Richman, supra note 52; Richard S. Frase, The Decision to File Federal Criminal Charges: A Quantitative Study of Prosecutorial Discretion, 47 U. Chi. L. Rev. 246, 291-99 (1980) (observing that individual prosecutors' offices have an interest in maintaining a consistent and uniform charging policy); Sarah J. Cox, Prosecutorial Discretion: An Overview, 13 Am. Crim. L. Rev. 383, 414 (1976) (noting that perhaps the greatest constraint on prosecutorial discretion is the prosecutor's desire to win); see also Elisabeth Alden Bressee, Prosecutorial Discretion, 75 Geo. L.J. 859, 861-62 (1987) (arguing that prosecutors should be constrained by desire to avoid violating defendants' constitutional rights).

121. Some are never investigated because they are unknown to the government. Comparison between this group and those who are prosecuted does not say much about discretion, as very little is exercised in the decision not to investigate unknown people. There is, however, a significant group of people the federal government could intelligently target for investigation if it had the resources. Concerns about political pressures not to pursue vigorously some classes of wrongdoers are legitimately raised in the white-collar area but resources are clearly a limiting factor there and in other classes of cases. Even in narcotics prosecution, the government has intelligence about many offenders against whom it has not yet developed legally sufficient evidence. Every defense lawyer who has represented a cooperator has heard lists of drug dealers recited. Some of them are clearly known to the government but many remain in the queue, awaiting investigation that may never come.
sions. Provability is a murky and moving target. How much time should investigators and prosecutors invest to develop evidence of another narcotics transaction against a defendant who already faces a stiff sentence? Should a prosecutor reinterview a questionable witness in order to bring an additional count? How much time should be invested to permit reasonable certainty about an issue that would only increase an already lengthy sentence? These are exactly the resource allocation issues at the heart of prosecutorial discretion. The doctrine is a powerful and ostensibly simple way of giving those with the best access to information the ability to use that information to determine how to enforce the law with regard to a particular defendant.

The notion that prosecutors do not really have or do not really exercise this discretion blinks at the reality of our criminal justice system. Efforts to hide behind policies suggesting that every defendant, or every wrongdoer, is prosecuted as severely as the law allows denies the central wisdom of the system of unreviewable investigatory and charging discretion and, if accepted, would leave very important questions about allocation of resources unexamined. The bottom line is that a federal prosecutor has the practical power to select among a wide range of sentences. In the example of Mr. X, one set of choices would result in a minimum sentence beyond judicial control that could range from five to fifteen years. Other choices would permit the judge to impose a sentence between very minimal time in jail and life. Combinations of these charges could result in a sentence determined by a combination of judicial and prosecutorial power that ranged between the lightest and heaviest possible sentences. Although prosecution guidelines and office policies are an important framework, the daily pressures of enforcement and litigation, filtered through the norms of a particular United States Attorney's Office, affect the crucial details of daily exercise of federal prosecutorial discretion.

IV. FORCES THAT MODERATE: TAKING ADVANTAGE OF DISCRETION

Prosecutorial discretion, overcriminalization, and harsh sentences are only one side of the equation. Malefaction, in all its odious forms, stands on the other side. The trick is to figure out how to use the almost plenary power to fairly and efficiently battle crime. The mere fact that society has given to prosecutors the power to imprison many defendants for very long terms does not mean, in either the practical or normative sense, that prosecutors should do so in every case.

A. Why Prosecutors Need to Make the Right First Offer

Prosecutors can use their ever-expanding range of unreviewable options in a

122. The actual power rests with the person who approves charging decisions in any particular United States Attorney's Office. Line assistants do not make these choices alone, although they can influence decisions in many ways.
variety of ways. Prosecutors might choose the charge that most exactly fits the
conduct at issue and so honor the legislature’s authority to define criminal
behavior. Of course, exact fit also carries proof burdens, so prosecutors might
choose the most general offense available to ease their stiff evidentiary burden.
Prosecutors might also consider strategic factors in choosing charges and try to
maximize the number of rapid pleas to reasonably serious and appropriate charges.
If most cases come down to a negotiation between the prosecutor and the
defendant to determine the price (length of prison sentence) at which the defendant
will sell his plea, then the criminal complaint is the prosecutor’s initial offer.
Negotiation theory suggests that in this two-party, adversarial negotiation involv-
ing a zero sum or distributive issue, the best result will be realized by making as
extreme an initial offer as the party can reasonably defend. This is just what
many prosecutors will do, using their charging discretion to initiate the case with
charges that carry the potential for a very lengthy sentence.

The defendant’s first response therefore is an extreme counteroffer: the manda-
tory interposition of a plea of “not guilty” at the time of arraignment. That is an
offer of zero in response to the high initial demand of a relatively long sentence. Of
course, one important difference between the plea negotiation and buying a car, for
instance, is that one party to this negotiation cannot easily walk away from the
negotiation. The defendant who cannot gain an acquittal will be compelled to pay a
price for his conduct and can only exercise some control over that price in two
ways. He can either plead guilty in exchange for some sentencing concession from
the prosecutor or he can appeal to a third party—the court—to moderate the
sentence. But the court’s power to moderate the sentence is greatly constrained by
the mandatory minimums, leaving many defendants at the mercy of the prosecutor.

Perhaps that is not such a bad thing. Many narcotics defendants are stunned
when just hours after their arrest, they learn from their defense lawyers that the
government has witnesses, tape recordings, and drug seizures that will likely put
the defendants in prison for ten, twenty, or more years. For many, the most
stunning aspect of the system is the fact that the sentences are mandatory and only
the prosecutor has the authority to alter the outcome. Of course, the system may
want defendants to be stunned. Initiating a prosecution with strong evidence and
convincing threats of extremely harsh punishment encourages rapid, low-cost
guilty pleas and cooperation against other potential defendants. Both rapid

123. In 2000, 95.5% of all federal convictions were secured by a plea of guilty. 2000 FED. SENT. SRCBK., supra
note 22, at fig.C. For some of the most insightful arguments against the system, by one of the most persistent
critics of plea bargaining, see generally Stephen J. Schulhofer, Plea Bargaining as Disaster, 101 YALE L.J. 1979
(1992) (arguing for elimination of plea bargaining). For discussions of pre-Guidelines plea bargaining, see
generally Alschuler, supra note 3 (criticizing plea bargaining as irrational, unfair, and reaching results unrelated to
proper objectives of criminal prosecution); Utz, supra note 3 (defending plea bargaining).

negotiations to resolve whether a given party gets more or less of a limited good often result in strategically
competitive negotiations in which extreme first offers are positively correlated with the final result).
resolution and cooperation can lead to a much more efficient use of scarce law enforcement resources.\textsuperscript{125}

But some offers may be too harsh and also inefficient. The threat of a twenty- or thirty-year sentence can induce an extreme response, such as a demand for trial. In order to produce credible support, some extreme threats may also require extra work. A prosecutor must consider the marginal cost of increasing punishment. A given defendant may quickly plead guilty to a five-year sentence. If the prosecutor holds out for the longer sentence that would be imposed after trial, perhaps fourteen years, the case could require litigation. A fully-litigated case requires tremendously more time and effort than a plea. Although it is only an estimate, a plea might require five hours of additional work after indictment,\textsuperscript{126} while a fully-litigated trial and appeal may require hundreds of hours.\textsuperscript{127} There are of course other economic constraints in the system, such as judicial time and prison space. From the prosecutor's point of view, however, the first five years of imprisonment are very cheap and the last nine years are extremely expensive.

Economic concerns like prison space and trial costs are, of course, only small pieces of the resource allocation question.\textsuperscript{128} Some relatively small number of all the cases must be tried to keep the threat of trial credible and to induce pleas in the remaining cases. There is also a significant question about how much of a benefit is necessary to encourage the right number of pleas to the right charges for the correct sentences. Sentences imposed after trial and plea should be high enough to serve the deterrent and expressive functions of the law, to encourage cooperation and pleas of guilty, and still to be fair and equitable. Criteria for these goals are impossible to specify and even a working definition is difficult to achieve.

We rely upon a system of bargaining to achieve a balance among these goals: deterrence, expression, cooperation, guilty pleas, justice, and equity. In theory,

\textsuperscript{125} However, as we will see, there is no evidence that we are managing these forces to maximize efficiency. More importantly, efficiency is not the only goal of the criminal justice system.

\textsuperscript{126} My informed guesstimate is that the case would require no more than five additional hours of prosecutorial time and virtually no additional investigation. The primary prosecutorial work would involve negotiating the deal, customizing the standard plea agreement, attending the plea, reviewing the pre-sentence report, and appearing at the sentencing.

\textsuperscript{127} Litigation through trial and sentencing may demand an additional 100 to 200 hours of prosecutorial time and perhaps sixty additional hours of investigator time, compared to a plea of guilty. Pretrial motions could reasonably require five to ten hours of research and drafting and three to six hours of hearings. Pretrial preparation might require five to ten hours of witness preparation, involving both prosecutors and agents, and ten or more hours of document and tape review and other preparation. An average smaller narcotics trial might require five to eight court days, requiring at least fifty to eighty hours of time, often from two prosecutors and the case agent. Litigating sentencing issues could require an additional five to ten hours. Most trial convictions are appealed and that process involves anywhere from fifteen to several hundred additional hours of prosecutor time, depending on the complexity of the legal issues.

\textsuperscript{128} Prison space, given its very high cost, is just one example of the range of resources issues that must be balanced. See Robert L. Misner, Criminal Law: Recasting Prosecutorial Discretion, 86 J. CRIM. L. & CRIMINOLOGY 717, 766-71 (1996) (urging tying prosecutorial discretion to availability of prison resources to internalize costs and provide an incentive to maximize strategic law enforcement planning).
prosecutors make the plea offers they think will induce pleas while ensuring appropriate punishment. If the offers are too high and defendants do not plead at rates that serve prosecutors, they will either lower their offers or raise the charges to induce more pleas. Of course, there are many legal constraints on the operation of this market, but experienced criminal lawyers recognize a significant element of the marketplace in plea negotiations.

B. How Differing Agenda Help Resolve Cases

As in many other kinds of negotiations, differing goals among the various players make reaching an agreement easier because each participant focuses on a different element. Prosecutors are most concerned about getting a conviction; it is the statistic that is reported and publicly available and the most common measure of success.\(^\text{129}\) Most prosecutors are relatively less concerned about maximizing the sentence in every case, which makes them more willing to bargain over sentences to achieve a conviction.\(^\text{130}\) The focus on convictions rather than sentence length is rational. Certainty of punishment has more of a deterrent effect than does mere severity of punishment.\(^\text{131}\) Prosecutors, therefore, should put more resources into convicting more people, even if they serve shorter sentences, than into convicting fewer people but achieving harsher punishments.

Many federal criminal defense lawyers,\(^\text{132}\) on the other hand, have a very strong incentive to focus on minimizing punishment. In many federal cases, defendants

\(^{129}\) See Daniel Richman, Old Chief v. United States, 83 Va. L. Rev. 939, 965-69 (1997) (noting that prosecutors seek to maximize convictions and federal prosecutors, in particular, are measured against their peers by conviction rates and their ability to move the docket by securing guilty pleas).

\(^{130}\) Although local cultures vary, in my experience, federal prosecutors will talk about their success at getting convictions but would feel it unseemly to brag about putting defendants in jail for a very long time. One important distinction may be between career prosecutors and those who spend several years as prosecutors before moving on to other legal settings. Career prosecutors may focus more on sentencing than their shorter-term colleagues, who may be concerned about how they integrate into the defense bar once they finish in government service. It is also likely that federal prosecutors in states which have generally longer and harsher state prison sentences will typically reflect that local culture and have more interest in longer sentences. Death penalty cases present another special category of cases in which the particular sentence imposed is often very important to the prosecutor.

\(^{131}\) See Paul H. Robinson & John M. Daley, The Utility of Desert, 91 Nw. U. L. Rev. 453, 458-64 (1997) (reviewing evidence that high sentences have very little deterrent effect because the likelihood of arrest is so low and arguing for a desert-based theory of criminal punishment).

\(^{132}\) Of course retained counsel has somewhat different incentives from appointed counsel and lawyers in white-collar cases face different constraints from many other types of cases. In particular, lawyers seeking private clients may need to protect their reputations as lawyers who will fight hard to get their clients acquitted. It is often bad for business to be able to attract only the recently indicted client who recognizes he or she has no chance of complete exoneration and must focus on mitigating the punishment. See Daniel C. Richman, Cooperating Clients, 56 Ohio St. L.J. 69, 75-77 (1995) (arguing that criminal defense lawyers in private practice often have reputational and related economic interests in not representing cooperators or at least having a reputation for not representing cooperators). Lawyers representing white-collar defendants also face situations in which the collateral consequences of any conviction are greater than the likely criminal sanction. For many wealthy defendants, the loss of business, licenses and reputation may be more significant than a stay in a minimum-security prison and the collateral consequences are virtually never attenuated by cooperation.
face strong evidence and have little hope of escaping conviction. Although virtually all federal criminal defense attorneys I know are vigorous advocates and fight hard at every stage, in many cases there is little hope of exoneration, but there is often real hope of sentence mitigation. Prosecutors cannot negotiate over the polar choice of guilty or not guilty, but they can offer an incentive to plead through apparent sentence reduction. Defendants are often sensitive to even modest reductions in sentence. As studies consistently show, people are quite sensitive to relative change from an established benchmark, in this case the prosecutor’s initial offer. Once the defendant has accepted the possibility of the high end of a given guideline range, any decrease from that sentence takes on greater psychological importance than might be predicted. Prosecutors and defense attorneys have complimentary incentives: prosecutors want to maximize the number of convictions so they offer a sentencing incentive to defendants anxious to minimize their punishment.

Judges, meanwhile, have their own incentives. In the sentencing process, as in other parts of their work, judges are concerned primarily with docket management. Federal judges have few tools to encourage rapid case resolution, so it is particularly important for such judges to use whatever discretion they have to build and maintain a reputation consistent with their case management philosophy. For many judges, the concern with reputation encourages fairly consistent sentencing patterns that serve to develop and to sustain the expectations that encourage desirable behavior. Judges typically exercise moderate and predictable leniency in

133. Much federal law enforcement is proactive, while state enforcement tends to be reactive. In other words, by the time the typical federal defendant is arrested the government has amassed evidence and can present trained law enforcement agents as witnesses. In many state cases, the investigation follows the arrest and relies upon civilian witnesses.

134. 60,255 defendants were “disposed of” in the federal courts in 1996. Dismissals accounted for 7083 of those defendants. Of the remaining 53,172 defendants, 4976 (9.4% of the defendants whose cases were not dismissed) went to trial. Of those 4976 defendants who went to trial, 902 (eighteen percent of all defendants who went to trial) were acquitted and 4074 (eighty-two percent of all defendants who went to trial) were convicted. The remaining 48,196 (90.6% of the defendants whose cases were not dismissed) entered pleas of guilty or nolo contendere. Bureau of Justice Statistics Sourcebook of Criminal Justice Statistics tbl.5.27 (1998) [hereinafter “1996 Fed. Sent. Srbk.”]. The rate of trials in the state courts appears similar, with an average plea rate of 89% reported for felony convictions in state courts in 1994. Id. at tbl.5.52.


136. See generally Amos Tversky & Daniel Kahneman, Introduction, in DANIEL KAHNEMAN, PAUL SLOVIC & AMOS TVERSKY ED., JUDGMENT UNDER UNCERTAINTY: HEURISTICS AND BIASES 14-20 (1982) (discussing cognitive bias which leads us to view given problems in quite different terms depending on the perspective from which the problem is viewed); Ian Weinstein, Don’t Believe Everything You Think: Cognitive Bias in Client Counseling, 9 CLINICAL L. REV. 501 (reviewing research on cognitive bias and suggesting how lawyers should account for this problem in counseling clients) (on file with author).

Thus far, I have described a system in which prosecutors set the range of the sentence through charging decisions intended to maximize convictions and efficiently allocate resources. Defense lawyers seek the lowest sentence possible and judges want to move cases along as quickly as possible. I also grant that each actor in the system is also concerned about justice, equity, individual success, and a host of other factors. Prosecutorial discretion and plea bargaining, however, are at the core of our system. Resource allocation through charging discretion is a very important part of this picture. In this view of the criminal justice system, one can expect that charging and sentencing decisions will have a strategic element in that they will be driven by concerns other than the merits of the particular case.

Part V of this Article argues that the forces of prosecutorial discretion have reached an equilibrium with the systemic imperative to resolve cases efficiently and the defendant’s power to reject or resist plea offers deemed to be too harsh. Part V offers observations that in non-narcotics cases, where generally these forces operate without mandatory minimum sentencing statutes, sentences and sentencing practices are reasonably stable and predictable. Average sentence length for a given offense in the non-narcotics realm tends to remain constant over time while downward departure rates tend to drift slowly and steadily upward. These trends permit the parties to make informed choices, to promote fairness over time, and to strike a balance between efficiency and fairness.

V. SENTENCING IN NON-NARCOTICS CASES: A REASONABLE BALANCE OF STRATEGY AND STABILITY

National sentencing data for non-narcotics cases reveals that sentences for these cases have been stable over time. Patterns of departures and other sentencing practices reveal that prosecutors and judges engage in consistent strategic charging and sentencing to maintain that stability, to encourage pleas and to individualize sentences. This Part first discusses the evidence of the stability of non-narcotics sentences and then analyzes three sentencing practices. The first is the overwhelming systemic preference for downward departures in contrast to the almost total lack of upward departures. The second is the very strong tendency to sentence defendants at either the very lowest or the very highest sentence within the Guidelines range, and not to distribute the sentences evenly within the ranges. The third is the consistent and slow increase in the percentage of cases in which there are downward departures without a drop in average sentence length.

A. The Stability of Non-Narcotics Sentencing Practices

Sentence length, rates of substantial assistance departures, and trial rates for

138. See infra Part V.B for a discussion of how judicial sentencing patterns reflect this tendency.
non-narcotics cases have been stable over the recent past for non-narcotics cases. For white-collar offenses, the second largest group of federal prosecutions after drug cases, the average sentences have hovered around twenty months for the past five years. Similarly, average sentences for violent offenses have increased and decreased within a modest range. This stability is in contrast to the steady decline in narcotics sentences in the same period. The latter will be discussed in Part VI.

Rates of substantial assistance departures have also been stable for non-narcotics cases. Over the past five years, there have been substantial assistance departures in approximately seventeen percent of non-narcotics cases. White-collar offenses tend to have rates lower and more stable than the narcotics offenses; the same is true for robbery, which is the most common violent offense prosecuted. A third indicator of systemic stability is the relatively constant rates of pleas in non-narcotics cases. Although these rates do show some increase over time, the increase is still less than in the narcotics cases. In 1996, 93.1% of the defendants in fraud cases pleaded guilty. In 2000, the comparable figure was 95.5%, and robbery cases show a similar trend. The plea rates in forgery/counterfeiting cases have been consistent at about ninety-seven percent.

139. This group includes fraud, embezzlement, forgery/counterfeiting, bribery, tax offenses, and money laundering. 2000 Fed. Sent. SrCbK., supra note 22, at fig.E n.1. In 2000, there were 10,540 cases in that group, accounting for about eighteen percent of all federal offenses. Id. at fig.A.
140. In 1996 and 1997, the average sentence was 20.2 months. It dipped to 19.2 months in 1998, rose to 20.7 in 1999, and settled back to 20.5 months in 2000. Id. at fig.E. In that same period, narcotics sentences marched steadily down. See infra Part V.
141. This group includes murder, manslaughter, kidnapping, sexual abuse, assault, bank robbery, and arson. 2000 Fed. Sent. SrCbK., supra note 22, at fig.E n.1. In 1996, the average sentence was 106.7 months. It dipped to 104.7 in 1997, rose to 105.7 in 1998, fell to 97.9 in 1999, and rose back to 102 in 2000. Id. at fig.E.
142. See infra Part VI for a discussion of the decline in the average length of sentences in narcotics cases.
143. In 1996, there were substantial assistance departures in 17.4% of the fraud cases. In 1997, the rate was 17%, in 1998 it was 14.6%, in 1999 it was 17.3%, and in 2000 it was 18.8%. 1996 Fed. Sent. SrCbK., supra note 134, at tbl.27.
144. Substantial assistance rates in the next largest group of white-collar offenses after fraud cases, forgery/counterfeiting cases, are: 10.7% in 1996, 12.7% in 1997, 11.5% in 1998, 12.4% in 1999 and 11.5% in 2000. Embezzlement cases have lower, but also steady, rates of substantial assistance, in part because embezzlement tends to be a solitary offense, committed by people otherwise uninvolved in criminal activity who have no one on whom to snitch. The rates are: 4.4% in 1996, 3.4% in 1997, 3.4% in 1998, 4.5% in 1999 and 5.5% in 2000. 1996 Fed. Sent. SrCbK., supra note 134, at tbl.11; 2000 Fed. Sent. SrCbK., supra note 22, at tbl.27. For a discussion of the narcotics sentencing data, see infra Part VI.
145. In 1996, there were substantial assistance departures in 14.2% of the robbery cases. In 1997 the rate was 13.5%, in 1998 it was 16.7%, in 1999 it was 15.3%, and in 2000 it was 13.1%. 1996 Fed. Sent. SrCbK., supra note 134, at tbl.27; 2000 Fed. Sent. SrCbK., supra note 22, at tbl.27.
In these cases, stable average sentence length, as well as consistent plea and substantial assistance departure rates flow from a balance of forces. The key to this particular equilibrium is that prosecutors and judges can offer apparent mitigation attractive enough to induce pleas while keeping overall sentence length, substantial assistance rates, and plea rates stable.\textsuperscript{149} The resulting stability in sentence length promotes equity over time. Consistency in sentencing practices also encourages rational decision making. From the defendant’s point of view, predictability is desirable because so much of the decision about resolving a case turns on predictions about the probable outcome of contingent events. Prosecutors benefit from predictability because they can better predict how cases will be resolved and can make allocative decisions according to the way in which a case will be resolved. Stability in these indicators also suggests a systemic balance of forces that improves the odds of efficient law enforcement. Stable sentence length and sentencing practices actually may represent one correct price of the 1980 reforms for both defendants and prosecutors in non-narcotics cases.

\textit{B. Stable Sentencing Strategies for Apparent Mitigation}

One might, however, take the efficiency hypothesis too far. That the current sentencing scheme with respect to non-narcotics cases produces some efficient outcomes does not mean that the scheme is the best of all possible sentencing schemes; rather, it is just one scheme that has reached a certain equilibrium and is predictable over time. Furthermore, the pattern of sentencing practices that characterize the equilibrium responds to systemic pressure. The pattern results not from the aggregate wisdom of thousands of decisions based on the merits of each individual case but from consistent responses by judges and prosecutors to their needs to move their dockets, to process cases, and to individualize sentencing in the interest of fairness.

These pressures cause virtually all the observable sentencing discretion in the federal system to be exercised to mitigate, and not to enhance, sentences. Out of 18,405 case involving departures from the Guidelines in 2000, 358, or 0.7\%, increased the sentence and 99.3\% of all departures decreased the sentence.\textsuperscript{150} 

\begin{footnote}

note 22, at tbl.11. This also reflects the precise dynamics of these cases, which are particularly hard to defend at trial.

149. The impact of a downward departure cannot be assessed without evaluating the charging decision. For example, a defendant charged with an offense that carries a Guideline range of sixty-three to seventy-eight months who receives a downward departure of two levels and a sentence of fifty-one months has apparently received mitigation. That case will appear in the downward departure statistics. That defendant, however, has still received a longer sentence than if he had been charged with an offense that carries a range of forty-one to fifty-one months, received no apparent mitigation at sentencing, and was sentenced at the low end of that range.

150. In 1999, there were downward departures in 34.5\% of the cases. 0.6\% of the cases involved an upward departure. \textit{United States Sentencing Comm'n, 1999 Sourcebook of Federal Sentencing Statistics}, at tbl.27, fig.G (2001) [hereinafter "1999 Fed. Sent. Sourcek."]. Just under 0.2\% of all departures were upward departures, or 313 out of 18,405. \textit{Id.} Although this very low number is consistent with an overall view that
\end{footnote}
About half of those departures were controlled by prosecutors and about half were judicial departures.\textsuperscript{151} Other prosecutorial mechanisms for mitigation, such as lenient charging decisions and sentence-fact bargaining, are not visible in the sentencing statistics but also are a significant factor. We have seen already that the system is structured to encourage systemic prosecutorial harshness in the early phases of a case, followed by very moderate offers of mitigation to induce pleas. In almost all cases, prosecutors correctly bring charges that carry sufficiently long penalties without sentencing phase enhancement. Only lack of prosecutorial foresight would require consistent enhancement of punishment at the sentencing phase.

Judges also understand that sufficient prosecutorial charging discretion should result in a lack of need for judges to take upward departures. Thus, while judges have the power to impose sentencing enhancements through upward departure, they rarely exercise that power. From the judicial point of view, excessive prosecutorial mitigation in the form of unwarranted undercharging or excessively favorable sentence-fact bargaining (the kind that would have to be corrected by sentencing enhancement through departure), is not the systemic danger judges must protect against. Given that prosecutors have many different tools to achieve very high sentences, and some incentive to seek such sentences, judges rarely have an incentive to correct prosecutorial lack of enthusiasm by using their discretion to increase punishment. Our system of overcriminalization and prosecutorial charging discretion has little use for sentencing phase enhancement.

A second area of strategic sentencing patterns reinforces the conclusion that mitigation relates to systemic reliance on overcriminalization and prosecutorial discretion and not solely on the merits of the case or attitudes about sentence severity. Almost two thirds of all sentences within the Guidelines range are imposed at the Guidelines minimum.\textsuperscript{152} More than sixty-one percent of all defendants who did not receive departures were sentenced at the low end of the Guidelines range in 1999.\textsuperscript{153}

\textsuperscript{151} There were a total of 19,371 departures in 2000. 9743 were substantial assistance departures, controlled by prosecutors, and 9628 were other downward departures and upward departures, all controlled by judges. 2000 FED. SENT. SRCBK., supra note 22, at tbl.27.

\textsuperscript{152} The sentencing judge has discretion to determine where in the applicable range of sentences the actual sentence will fall. This is the remaining bit of the once almost limitless, unreviewable sentencing discretion of district court judges. See 18 U.S.C. §§ 3742(a)(3) & (b)(3) (2002) (authorizing appeal of any sentence above or below range, but not within a lawfully determined sentencing range).

\textsuperscript{153} Id.
Why do almost two thirds of the non-departure cases fall at the Guidelines minimum? In these cases, placement in the range is the only sentencing discretion judges can use to influence the course of the case. Judges are sensitive to caseload pressures. The difference between resolving a case by trial and by plea has very significant docket management consequences. Research indicates that the variation between sentences after trial and sentences after plea has increased under the Guidelines as has the variation among judges. This research suggests that many judges use the discretion they retain to promote pleas and that parties still bargain in the shadow of the judge. As relatively little is at stake (the difference between the top and bottom of the range is the greater of six months or twenty-five percent of the bottom of the range, with the ranges overlapping their neighbors on either side), it makes a great deal of sense for judges to seek and maintain a reputation for typically sentencing at the bottom of the range after a plea. Coupling that reputation with one for sentencing harshly after trial is the only incentive for pleas that judges can control with their sentencing, whatever their attitudes about sentence severity in general.

Consistent with the view that placement in the range is motivated by systemic concerns is the fact that the second largest group of within-range sentences, 14.3%, are imposed at the Guidelines maximum. This includes the very small group of violent offenses in the federal system. The extreme conduct involved justifies the highest sentence available. The conduct also includes some cases in which the prosecutor already had exercised so much leniency in charging, plea-, and sentence-fact bargaining that the judge felt it appropriate to impose a sentence at the high end of a range that was very low for the particular case. Overall, more than seventy-five percent of all within-range sentences were imposed at the extremes, suggesting strategic use of sentencing rather than an effort to distribute the cases fairly along a continuum.

---


156. If the cases were arrayed within each range in a distribution correlating with their relative merits, they would be more evenly distributed along the range and would not cluster with two thirds at the bottom of the range. The representativeness bias, see infra Part V.B, also skews these sentences.

157. In 1999, 14.3% of all within range sentences were at the maximum. 1999 FED. SENT. SRCBK., supra note 150, at tbl.29.

158. Of the ninety murder and manslaughter cases, forty-two percent, or thirty-eight, were at the guideline maximum. Similar high percentages at the maximum are found in arson cases, 44.2% of the forty-three cases and racketeering, at 31.7%. Id.

159. This interpretation is supported by the fact that the highest percentage of maximum sentences, 68.3%, is found in the 218 cases involving charges of using a communications facility in narcotics dealing. 21 U.S.C. § 843(b) (2002). This charge, which carries a four-year maximum, is often used in cases in which a very favorable plea is offered to resolve a more serious drug charge.

160. 1999 FED. SENT. SRCBK., supra note 150, at tbl.29.
A third area that fits the model of harsh charging and apparent mitigation is the slow but steady increase in rates of judicial downward departures.\textsuperscript{161} Even as average sentences in non-narcotics cases do not decrease,\textsuperscript{162} judges have granted downward departures at an increasing rate from 1996 to 2000. The independence of departure rates and sentence length is the clearest evidence that prosecutors and judges can control the final sentences, even in the face of rising sentence phase mitigation. In examining downward departure rates, one can see the intersection of the common law focus on individual cases, the differing incentives of prosecutors, judges and defense lawyers, and the interest in sentencing stability. These interests converge to make the slowly increasing judicial mitigation and stable sentences a hallmark of a reasonably presumptive Guidelines system.\textsuperscript{163} The system can permit the resulting increasing rates of downward departure because it can render them meaningless in many cases through charging and other sentencing decisions that enhance sentences in anticipation of later mitigation. Such a pattern also increases individualization and is a slow and subtle way to regain some judicial control over sentencing. Downward departures become more frequent because judges tend to see the cases before them as more atypical than they really are, and the law enshrines those views in new departure grounds through a lenient standard of appellate review.

When the Guidelines first came into force, no one ever had requested or sought a departure. There was much speculation about how the courts would implement departures. The more adventurous and creative defense lawyers advanced new and untested legal theories. Creative arguments were well received in the trial courts for at least two reasons. First, many judges were resistant to losing their sentencing authority, which disposed them to be receptive in that initial period.\textsuperscript{164} More importantly, however, the departure argument, by its nature, has a special intuitive appeal. The key question in judicial departure cases is whether the case at hand is typical or atypical among the cases that fall under a given guideline. Studies show that decision makers are not skilled at determining how usual or unusual an event

\textsuperscript{161} In fraud cases, there were non-substantial assistance downward departures in 8.5% of the cases in 1996, 7.5% in 1997, 8.3% in 1998, 9.7% in 1999, and 11.6% in 2000. In forgery/counterfeiting cases the rates are: 4.4% in 1996, 4.4% in 1997, 7.1% in 1998, 6.4% in 1999, and 7.8% in 2000. Embezzlement cases, which have among the lowest average sentence lengths at around 6 months, and among the highest rates of non-incarceratory sentences, have had rates of: 13.5% in 1996, 12.3% in 1997, 11.8% in 1998, 11.1% in 1999, and 13.0% in 2000. 1996 \textit{Fed. Sent. SRCbk.}, \textit{supra} note 134, at tbl.27; 1999 \textit{Fed. Sent. SRCbk.}, \textit{supra} note 150, at tbl.27; 2000 \textit{Fed. Sent. SRCbk.}, \textit{supra} note 22, at tbl.27.

\textsuperscript{162} See \textit{supra} Part V (arguing that average sentence length is stable in non-narcotics cases).

\textsuperscript{163} As noted above, state systems also offer examples of this pattern. See \textit{supra} note 39.

\textsuperscript{164} Judicial downward departures started to rise in 1992 and have risen steadily since then. They fell during the first three years of Guidelines sentencing, from 1989 to 1991. 1996 \textit{Fed. Sent. SRCbk.}, \textit{supra} note 134, at fig.G. The early fall in these rates suggests that when the sentencing laws changed, judges who had been accustomed to exercising a great deal of authority used the departure mechanism to continue to exert power over sentencing. As judges grew accustomed to the Guidelines, departures fell for a brief period until the dynamics I have described took hold.
is in the absence of valid and well-presented data about the relevant base rate. Known as the representativeness heuristic or bias, it suggests that everyone, including a judge,\textsuperscript{165} tends to favor highly salient or available information. In other words, the case at hand, being quite familiar, seems special compared to the many unknown and only dimly imagined cases on other judges’ dockets. Trial judges are not in the position to determine how typical a particular circumstance is among the many cases across the country. In the Guidelines cases, the representativeness illusion causes judges to conclude much too quickly that the case at hand is atypical and merits the special treatment of a departure.\textsuperscript{166}

A problem related to the lack of complete information is the egocentric bias. We all have data about our own experiences and see the world from that perspective. We overestimate the importance of our own experiences, in part because we do not have data against which to compare it. All the players at the trial court level devote a great deal of time to each of the cases, as they should; but that very narrow focus only intensifies the biases, all of which cause the players to evaluate unrealistically how special is the case at the bar. These two pervasive cognitive biases may help explain why there are so many “special” cases under the Guidelines.

The tendency to see each case as special was given the force of law by judges as they decided each case. Case law began to develop and a whole array of departure arguments became available through case research, treatises and other materials.\textsuperscript{167} Every year, new departure grounds were created and older grounds continue to spread and develop. Although some grounds were narrowed or even reversed by subsequent cases, the asymmetrical but complimentary incentives of defense lawyers and prosecutors, coupled with the general reluctance of courts to reverse earlier cases, tend to increase the number and range of available downward departure arguments over time.

The dynamics of appellate practice in the area of Guidelines departures reflects these forces. First, defendants file the vast bulk of all sentencing appeals. Defendants filed 6749 sentencing appeals in 1999.\textsuperscript{168} Prosecutors filed

\textsuperscript{165. See generally Chris Guthrie et al., Inside the Judicial Mind, 86 CORNELL L. REV. 777 (2001) (reporting results of a study of United States Magistrates showing they are subject to the representativeness bias).}

\textsuperscript{166. Judicially controlled downward departures are reserved for cases that present unusual circumstances which the Commission did not “adequately” consider. In determining whether a given drug defendant’s family situation is so extraordinary that it merits a departure, judges do not have access to valid data about how that defendant’s situation compares to those of all other narcotics defendants. Even if they did, judges, like all decision makers, would still overemphasize the ailing parent in the case at the bar and ignore or underplay the question of how many other defendants have similarly ailing parents. After all, the case at bar is a good representative of the class of cases that involve extraordinary family situations in that it looks a great deal like such a case, so we too easily conclude that it is such a case. Each district judge sees only a very thin slice of the whole set of cases. Yet they are asked to judge what is and is not typical. Although they get some help from probation officers, lawyers, and the reported cases, all these sources suffer the same cognitive biases.}

\textsuperscript{167. See generally BAMBERGER & GOTTLEIB, supra note 16 (discussing departures).}

\textsuperscript{168. 1999 FED. SENT. SRCBK., supra note 150, at tbl.57.}
Although one might think that prosecutors simply win most of the time in the district courts and defendants have a strong incentive to appeal even very weak issues because defendants have very little adverse risk, these numbers must be viewed against the steady increase in downward departures. The government appeals very few of the cases in which downward departures are granted. This datum is consistent with the notion that prosecutors do not seek the highest possible sentencing in every case. It is difficult to generalize about the outcomes of the cases that are appealed. Since Koon v. United States, some of the Circuit Courts have been receptive to the departure cases they have reviewed but others, notably the Fourth Circuit, have not supported the exercise of discretion by district judges. Although the Supreme Court reaffirmed the importance of judicial sentencing discretion in Koon, the greatest supporters of the judicial departure power seem to be the prosecutors who do not challenge it.

The rising rate of downward departure introduces an element of complexity into the simple story of stable average sentences. The mechanisms by which sentences remain stable are not clear and vary from district to district, but there is strong evidence that prosecutors use their charging discretion to influence sentences. The evidence includes empirical work and case studies. For example, prosecutors changed their charging behavior in response to the Guidelines. Case studies report that prosecutors consider sentencing when they make charging decisions and

169. About ninety-two percent of the time, judges' rulings in defendant-initiated appeals are affirmed. About 540 rulings are reversed upon defendant appeals. In 1999, the affirmance rate for government-initiated appeals was 68.5%, so only seventeen cases were reversed on government appeal in that year. There has been a decrease in the government's success rate since, when the government filed 176 appeals in Guidelines cases and the district court was affirmed in 38.1% of the cases. In that year the government succeeded in getting reversals in 109 cases. In 1997, the government filed 113 and prevailed in about sixty percent of the cases, and in 1998 the government filed 122 Guidelines appeals and succeeded in about fifty percent of the cases. In that period of time, the government's success rate in downward departure cases dropped even more dramatically, with the district courts being affirmed in only fifteen percent of those cases in 1996, but courts of appeals affirmed seventy-six percent of the twenty-one downward departure cases filed in 1999. These are, of course, very small numbers given the total number of cases, but judges and lawyers pay attention to the courts of appeals. Id. at tbls. 56-58.
170. Prosecutors only filed 176 sentencing appeals in 1999, while there were over 8300 cases in which there non-substantial assistance downward departures. Id. at tbl.24.
171. Koon v. United States, 518 U.S. 81, 96-97 (1996) (holding that abuse of discretion is the proper standard of review of district court decisions to depart downward from the Guidelines, and applying that standard to reverse some departure grounds and affirm others in the case of the two police officers convicted in the beating of Rodney King).
172. My review of all the downward departure cases heard in the Circuit Courts from the time Koon was decided through the end of 1998 revealed that seven circuits had reversed every downward departure reviewed, one had affirmed every departure and four issued a mixed set of decisions. Weinstein, supra note 25 (analyzing appellate review of downward departure cases); see also Bowman & Heise, supra note 92, at 1116 (reviewing the impact of appellate cases on departures).
engage in sentence-fact and factor bargaining.\textsuperscript{174}

Other evidence suggests that although prosecutors play the largest role, their decisions influence and are influenced by judicial discretion. For example, the rate of substantial assistance departures, which are controlled by prosecutors, tends to be counterbalanced by judicially controlled departures. As one goes up, the other tends to go down.\textsuperscript{175} In this way, the two kinds of departures interact to keep sentences stable. Judges may also grant more, but smaller, downward departures. It is harder to observe the details of charging discretion.\textsuperscript{176} In the best case, prosecutors would charge ever more serious cases over time, justifying the longer sentences that counterbalance the slow increase in downward departures. The more likely scenario, given that prosecutors already try to prosecute the worst conduct by the most evil people, is that prosecutors will charge the same cases a bit more harshly and engage in less sentence-fact and factor bargaining over time to balance increases in mitigation.

Some might see slowly increasing rates of mitigation and stable average sentence length as yet another kind of disparity among defendants. Those defendants who do not receive mitigation may receive longer sentences to counterbalance the increase in mitigation. Of course, the same average sentence will result if there are more, but smaller departures and the non-departure sentences remain unchanged. It seems more likely, however, that harsher prosecutorial charging and sentence bargaining would increase distinctions among defendants as measured by sentence length and thus enable the system to better differentiate among defendants. A wider range of sentences can better account for differing levels of culpability. If prosecutorial and judicial discretion are functioning as they should among the cases not affected by the mandatory minimum statutes, stable overall sentence length and slowly rising downward departure provide a mechanism to better individualize justice while minimizing unwarranted disparity over time.

Distinguishing warranted and unwarranted disparity is at the heart of the debate


\textsuperscript{175} Weinstein, \textit{supra} note 25, at 609 (reporting negative correlation between rates of judicial and substantial assistance departures).

\textsuperscript{176} There are many examples of the variation in practices among districts. Weinstein, \textit{supra} note 25, at 608 (discussing variations); Bowman & Heise, \textit{supra} note 44, at 531-58 (reporting results of regression analyses that support the conclusion that variations in sentencing practice cannot be completely explained by local variations in cases, defendants or other factors). In almost all districts, the local equilibrium appears quite stable. To offer one example, I have calculated that when the ninety-four districts are ranked by rates of substantial assistance and those rankings are compared for 1996 and 1999, only twenty of the ninety-four districts moved more than twenty places and only eight moved more than thirty places. In other words, fewer than one in ten changed position by more than a third and only two moved between the top and bottom quarters. Data on file with author.
about our current sentencing regime. But whatever view one takes on that question, addressing sentencing issues through sentence stage reform only has limited impact. As this Article demonstrates, sentencing is the final stage of a process that begins with the allocation of investigative resources and a process in which charging decisions and negotiated settlements often determine an offender’s punishment. Sentencing is often a ceremonial act that only formalizes the decisions made earlier in the case. Politicians who see subterfuge and circumvention by judges or prosecutors miss the point and vastly underestimate the difficulties judges and prosecutors face. Those who imagine that society can sanction every offender with the harshest penalty a legislature can pronounce have not thought through the magnitude of the task, the many different reasons society authorizes stiff sentences, and the demands of justice.

Our criminal law system relies on the feedback of negotiation and shared authority to find a reasonable level of punishment within the very broad statutory parameters. Stable average sentence length is an indicator of one kind of desirable balance and a good one in its own right. Stable sentences suggest that prosecutors find the level of punishment adequate to meet their instrumental goal of effective law enforcement and their view of the normative goal of appropriate punishment. Of course these judgments should not be given controlling weight, but they are entitled to consideration. We should not dismiss lightly a system that displays stability over time, especially when an obvious alternative, legislatively determined mandatory sentences, has much more obvious flaws.

VI. Narcotics Sentencing: An Unreasonable and Unstable System

It is very difficult to find praise for federal narcotics sentencing.\textsuperscript{177} The currently pervasive mandatory minimum statutes make narcotics sentences indefensibly rigid, often unfair and unjustly harsh. The excessive rigidity mutes the systemic feedback that would encourage efficiency. The emphasis on drug quantity instead of culpability causes great unfairness by inequitably sanctioning offenders. Most importantly, we are imprisoning people for many years and destroying lives and families while gaining nothing.\textsuperscript{178}

The excessive harshness and rigidity of narcotics sentences has resulted in a long-term trend of declining average sentences, very high but falling rates of substantial assistance departures, increasing rates of judicial departures, and increasingly rare trials. The long-term decrease in average sentence lengths has received attention. As Hofer and Semisch first noted, narcotics sentences peaked in

\textsuperscript{177} See comments by Justice Breyer, supra note 19 and accompanying text. Among the others who have voiced opposition to the use of mandatory minimum sentences are the Judicial Conference of the United States, the Federal Courts Study Committee, and the Commission. See Beale, supra note 69, at 27.

\textsuperscript{178} See Beale, supra note 69, at 25 (noting that overwhelming majority of criminal justice experts agree that harsher sentences fail to either greatly increase public safety or reduce crime).
1992 and have steadily declined since then, from a high of 77.4 months in 1992 to 68.5 months in 1998, an 11.5% decline.\textsuperscript{179} The average has continued to decline and narcotics sentences are between fifteen percent and twenty percent lower than they were in 1992.\textsuperscript{180} This steady and significant decrease distinguishes this very large group of sentences from others imposed in the federal courts.

The most thoughtful commentators in this area, Bowman and Heise, have analyzed a variety of factors that could explain the decrease in sentence length, including changes in the law,\textsuperscript{181} changes in case type,\textsuperscript{182} types of defendants prosecuted,\textsuperscript{183} prosecutorial caseload, and other local factors.\textsuperscript{184} Bowman and Heise's analysis suggests that these changes are responsible for a portion of the decrease in sentence length.\textsuperscript{185} They conclude, however, that some of the decrease stems from prosecutors' and judges' use of discretion to reduce sentences.\textsuperscript{186} Bowman and Heise must be right, as prosecutors clearly can maintain stable sentence length over time even in the face of increasing mitigation rates. The suggestion that something else must explain the continued decline in narcotics sentences is strengthened by the observation that average sentence length appears to continue to decrease even as cooperation departures have fallen and judicial departures have risen.

The very high rate of substantial assistance departures has become a signal feature of narcotics sentencing. In recent years, more than one third of all narcotics defendants received substantial assistance departures and more than one out of four receives this benefit today.\textsuperscript{187} The trend, however, is downward. Judicial

\begin{footnotes}
\item[179] Hofer & Semisch, supra note 47, at 17.
\item[180] For a thorough discussion of the data and varying interpretations of the data on the magnitude of the decrease, see Bowman & Heise, supra note 44, at 479-80. The bottom line, however, is that average narcotics sentences have fallen since 1992.
\item[181] Id. at 487-97 (discussing impact of safety valve, super acceptance of responsibility and other changes in the law).
\item[182] Id. at 501-05 (discussing impact of mix of cases involving crack, powder cocaine and other drugs).
\item[183] Id. at 505-12 (discussing impact of changes in offenders' role in offense, criminal history and other factors). One other factor that may play a role is the shift in the racial makeup of sentenced offenders. As one careful researcher suggests, African-American defendants receive the harshest sentences in federal court, followed by Latinos and then whites. See David B. Mustard, Racial, Ethnic and Gender Disparities in Sentencing: Evidence from the US Federal Courts, 44 J.L. & ECON. 285, 285 (2001) (finding that "blacks, males, and offenders with low levels of education and income receive substantially longer sentences"). The relative decrease in black defendants and the rise in the percentage of Latino defendants would tend to lower overall sentences. Although eyeballing the tables is not a substitute for regression analysis, whites were about the same proportion of the group of narcotics trafficking defendants in 1996 as in 1999, while blacks decreased from 35.2% to thirty-one percent and Hispanics increased from 37.8% to 42.9%. 1996 FED. SENT. SRCBK., supra note 134, at tbl.7; 1999 FED. SENT. SRCBK., supra note 150, at tbl.7.
\item[184] Bowman & Heise, supra note 44, at 535-58 (reporting results of regression analysis on district level data and finding that caseload, drug type, and other factors did not account for all decrease in average sentence length).
\item[185] See Bowman & Heise, supra note 44, at 525-58.
\item[186] Id. at 528-29.
\item[187] In 1996 thirty-five percent of narcotics trafficking sentences involved a substantial assistance downward departure. In 1997 it was 31.4%; 1998, 31.1%; 1999, 29.3%; and 2000, 27.8%. 1996 FED. SENT. SRCBK., supra
\end{footnotes}
departures have increased in narcotics cases, from 10.2% in 1996 to 15.3% in 2000. The last indicator of the problems in narcotics cases is that the number of guilty pleas has risen at a higher rate than in other cases, despite falling sentences. In 1996, 89.7% of all narcotics trafficking cases ended in a plea. In 2000 it was 94.9%. Over five years the rate of narcotics trials has dropped by almost fifty percent.

The overall picture in the realm of narcotics cases is one of instability. Unlike non-narcotics sentencing, prosecutors and judges involved in narcotics sentencing have not maintained sentence length or trial rates for these cases. Sentences fall despite fewer cooperation departures and fewer trials, suggesting that prosecutors do not need higher sentences to induce pleas, nor do they recognize the current level of cooperation as optimal. It appears likely that prosecutors are using charge bargaining to continue to let sentences fall.

These patterns of falling sentences, falling trial rates, and high cooperation rates are troubling. Falling sentences lead to disparity over time and suggest sentences are longer than required to induce pleas, a phenomenon that is particularly true when coupled with falling rates of trials. I offer no suggestion about the optimal trial rates and sentence lengths from the standpoint of efficiency. Whatever rates might be best, however, the long-term trend of decreasing trials is a detrimental result of the shift in power to federal prosecutors. The realistic threat of going to trial is one of the few sources of bargaining leverage for defendants and their lawyers. As the possibility of trial grows even dimmer, evidence is not revealed or tested and prosecutorial charging decisions are subjected to less scrutiny. We note 134, at tbl.27; 1999 FED. SENT. SCRIBK., supra note 150, at tbl.27; 2000 FED. SENT. SCRIBK., supra note 22, at tbl.27. Some data suggest that two thirds of all narcotics defendants attempt to cooperate. A sentencing commission study, using methodology that likely undercounted, reported that 67.5% of all the defendants provided some kind of assistance to the government. See MAXFIELD & KRAMER, supra note 61, at 26, exh. 5 n.21.

188. 1996 FED. SENT. SCRIBK., supra note 134, at tbl.27; 1999 FED. SENT. SCRIBK., supra note 150, at tbl.27; 2000 FED. SENT. SCRIBK., supra note 22, at tbl.27.

189. Trial rates have fallen for most offenses, although the decrease is greater for narcotics cases. 2000 FED. SENT. SCRIBK., supra note 22, at fig.C. See also supra note 141. Although the general pattern supports my argument, the fall in plea rates across almost all cases indicates that other forces, beyond the mandatory minimums, are driving this trend.

190. In 1996, about one in ten narcotics trafficking cases went to trial. In 2000, it was about one in twenty, a decrease by half. Among the much smaller group of fraud cases, about one in fourteen went to trial in 1996, and about one in twenty-one went to trial in 2000, a decrease of about one-third. Among embezzlement cases, the rate has remained steady at about one in thirty. 1996 FED. SENT. SCRIBK., supra note 134, at tbl.11.

191. 2000 FED. SENT. SCRIBK., supra note 22, at tbl.11.

192. For examples of the ways prosecutors charge bargain under the Guidelines, see Bowman & Heise, supra note 92, at 1119-22.

193. See generally Amy Farmer & Dek Terrell, Crime Versus Justice: Is There a Trade-off?, 44 J.L. & ECON. 345 (2001) (arguing that there is a trade-off between crime control and justice and offering a model that quantifies, but does not advocate for, the crime reduction that could be gained by increasing unjust convictions).

194. The current regime puts great pressure on defendants to attempt to cooperate very early in the life of the case. Defendants are often committed to pleading guilty long before discovery is made. The Guidelines have also introduced an incentive for defendants not to seek full discovery. In many cases, continued investigation by the
have lost a vital check on the prosecutorial discretion on which we rely but about which some skepticism would be healthy.\textsuperscript{195}

High cooperation rates are another problem area. Cooperation departures play a particularly significant role in narcotics cases for two reasons. The first reason is that narcotics cases are very difficult to prove without insider testimony. Unlike violent or theft offenses, narcotics cases generally do not involve victims who come forward to complain and to testify. The government must penetrate drug conspiracies through undercover agents, informants, or cooperators. Thus, higher rates of cooperation should be expected in narcotics cases, relative to other case types. The current high rates, however, are far above the historic use of cooperation and the huge rise in cooperation in the early years of the Guidelines system points to another explanation for the current rates.\textsuperscript{196}

Thus, the second reason is that cooperation is firmly entrenched as the most important device for prosecutorial control of sentence length in narcotics cases.\textsuperscript{197} Unfortunately, of all the ways prosecutors can control sentences, the mandatory

\textsuperscript{195} Given data showing that ten percent of all death row inmates should not have been convicted, see \textit{United States v. Quinones}, 205 F. Supp. 2d 256 (S.D.N.Y. 2002) (holding federal death penalty unconstitutional and marshaling evidence for the high rate of wrongful convictions in death penalty cases), we might reasonably worry about innocent defendants pleading guilty. There are many important differences between those cases and federal narcotics cases, not the least of which is that death cases are subject to intense local pressures and given little in the way of defense resources. Most of narcotics convictions result from pleas, which may be less subject to this problem. \textit{But see} Stuntz, supra note 66, at 1946-66 (discussing problem of innocent defendants taking advantageous pleas and suggesting reforms to improve plea bargaining). The extensive use of cooperators in narcotics cases is another reason to worry about this problem in narcotics cases. \textit{See generally} Ellen Yaroshefsky, \textit{Cooperation with Federal Prosecutors: Experiences of Truth Telling and Embellishment}, 68 FORDHAM L. REV. 917 (1999). Given the rates of error in death cases and other violent crimes in which DNA has now made more reliable fact finding possible, concern about the problem of the conviction of innocent defendants is warranted. Given the number of cases, even a rate of innocence one-tenth the rate in the death penalty cases would result in more than 220 innocent people being sentenced in narcotics cases every year.

\textsuperscript{196} I am not aware of any data on cooperation rates before the Sentencing Commission began collecting and analyzing data pursuant to its statutory directive, see 28 U.S.C. § 995(12)-(16) (1988), but the relatively sparing use of cooperation in the first years of the Guidelines is consistent with the sense of experienced practitioners that cooperation was not as widespread before the Guidelines as it is now. In 1989, the first year data was available, only 3.5\% of the cases involved substantial assistance departures. In 1990 the number climbed to 7.5\% and continued to move up to 11.9\% in 1991, 15.1\% in 1992, and 16.9\% in 1993. In 1994 the figure reached 19.5\% and has remained at about that level (19.7\% in 1995 and 19.2\% in 1996). 1996 \textit{FED. SENT. SRCBK.}, supra note 134, at fig.G. This information is also available at http://www.ussc.gov and is on file with the author. \textit{See also} John Gleeson, \textit{Supervising Criminal Investigations: The Proper Scope of the Supervisory Power of Federal Judges}, 5 J.L. & POL’Y 423, 424 n.9 (1997):


\textsuperscript{197} \textit{See} Weinstein, supra note 25, at 611-17 (marshaling evidence that high rate of substantial assistance departures in narcotics cases reflects an effort to control sentences, not maximize efficient use of resources).
minimum statutes force them to use the least efficient mechanism.\textsuperscript{198} Cooperation requires many prosecutorial and investigative resources. Unlike dropping a count in the indictment, cooperation requires meetings with both prosecutors and agents. If cooperation is successful, the information gained often gives rise to a new investigation. While it is desirable to prosecute new cases, it is not at all clear that the best way to allocate investigative resources is to order them according to how sentences should be imposed in already-prosecuted cases.

Narcotics sentencing has not reached a stable equilibrium because neither prosecutors nor judges have found the sentence level that meets their systemic needs. The continued decline of narcotics sentences suggests that the sentences are longer than necessary to encourage rapid pleas, to obtain adequate cooperation, and to allocate law enforcement resources efficiently. An optimally harsh sentence will encourage rapid pleas at the minimum investment of investigative and prosecutorial resources. An overly harsh sentence will also encourage the same plea but requires more resources as additional investigation and legal advocacy is required. If the same law enforcement result can be accomplished with a lower investment in each individual case, longer sentences requiring higher investments need another justification.

Beyond efficiency, systemic balance, and incentives, the fundamental problem in federal narcotics sentencing is rampant injustice. Narcotics sentences are severe in both the absolute and in comparison to pre-1980s sentences. From 1991 to 1996, the average sentence for a narcotics trafficking case was about six and two-thirds years and the median sentence was about five years.\textsuperscript{199} The average includes sentences that ranged from several months to life. Given that half of the sentences were above sixty months and the average is eighty months, sentences of fifteen and twenty years were not rare. Sentences have moderated, but they remain quite harsh.\textsuperscript{200} By any yardstick, whether or not one believes the sentences are just or fair,\textsuperscript{201} narcotics sentences are long. For many defendants, the prospect of a sentence of two or even three years is very daunting, but many can imagine returning to their families and lives. Defendants facing sentences of six or ten years often despair at the thought of maintaining family relationships and friendships and have difficulty imagining their lives after prison. Defendants facing fifteen years face emotional pressures of even greater magnitude. The current allocation of authority prevents discussion of the severity of these sentences. After all, when mitigation is largely in the control of the prosecutor, it makes little sense to mount

\textsuperscript{198} For the argument that substantial assistance is the right tool to address disparity, see Simons, \textit{supra} note 76 (arguing for more substantial assistance departures to address sentencing inequities).

\textsuperscript{199} 1996 FED. SENT. SRCK., \textit{supra} note 134, at fig.G.

\textsuperscript{200} In 2000, the average narcotics sentence was about six years and two months and the median was fifty-one months. 2000 FED. SENT. SRCK., \textit{supra} note 22, at fig.E. Very long sentences are still not uncommon.

\textsuperscript{201} It is beyond argument that these long sentences are not "winning the war on drugs." The reality is that however many couriers and other low-level people we lock up, others will come forward to take their place.
a frontal challenge to the harshness of the sentence and there is really nowhere to take the complaint. The current regime of overcriminalization, increased prosecutorial discretion, and very harsh sentencing reinforces itself. The only judgment that counts comes from an advocate.

CONCLUSION

This Article highlights two kinds of problems, one practical and one normative. The practical problem is that we have hobbled federal prosecutors' ability to exercise their central function in the enforcement of our narcotics laws. The normative problem is the injustice of these very harsh sentences.

Mandatory minimum sentences do not work in a system of overcriminalization and prosecutorial discretion. Although we may roll back some of the overcriminalization of the last fifteen years, it would be more consistent with our practices to restore checks and balances in the enforcement of federal criminal law. Like others, this Article urges a return to a greater judicial role in sentencing. The most important reforms would eliminate the mandatory minimum statutes, which would bring the federal criminal law system back to one in which prosecutors had greater incentives to ensure efficient law enforcement and could respond better to the information that flows from a more robust plea bargaining system. Judges would once again play a significant role in the normative judgments about blameworthiness and just deserts in all narcotics cases. Prosecutors, for all their virtues, should not have a monopoly on judgments about justice.

But once we move past the monopolistic effects of mandatory minimum statutes, finding the ideal balance between judicial and prosecutorial power is much more complicated. As this Article contends, sentences are the result of a long and complex process. Any system that has enough flexibility to achieve individualization will be subject to strategic behavior that will be labeled circumvention or subterfuge by some. The goal cannot be to eliminate that behavior, but to channel it as best we can. We should not idealize sentencing before the Guidelines. Unfettered judicial discretion did not promote justice in every case. There was great disparity among judges and jurisdictions and it was difficult to make informed judgments about sentencing. At least today we have benchmarks and data, even if it can be challenging to make sense of what we see.

Several perceptive commentators have discussed why our current politics make it so difficult to roll back mandatory minimum sentences and to restore some judicial authority in narcotics sentencing. Legislators can see no political advantage in changes that likely would make them appear soft on crime and that

202. For the most thoughtful proponents of this position, see generally KATE STITH & JOSE A. CABRANES, FEAR OF JUDGING: SENTENCING GUIDELINES IN THE FEDERAL COURTS (1998).

203. See Stuntz, supra note 66 (analyzing political pressures for harsher, more rigid sentencing); Richman, supra note 52 (same); Beale, supra note 69 (same).
would deny them the opportunity to shift blame for crime onto the prosecutors whom they have empowered to solve all the nation's ills. Although this analysis has great force, the politics of crime has tides like all other issues. Over the past 150 years, harsh policies have given way to reforms in several great waves. The nineteenth century stepped away from the routine use of the death penalty, transportation, and other cruel sanctions. Although now discredited in popular discourse, the sentencing reforms of the 1960s, including indefinite sentences and an elimination of an earlier generation of mandatory minimum sentences, likely will come back in vogue. The historical conditions that fuel reform movements are larger than the trends in an area as insular as federal sentencing reforms.

When those winds shift, perhaps popular opinion will see the normative question in the same light as those involved in the daily administration of these laws. Although I do not think prosecutors and judges are in rebellion against these laws, their behavior makes it plain that many have no great enthusiasm for them either. The systemic pressures this Article has described require these actors to exercise some moral judgment as they exercise their discretion. When those whose professional identity is bound up in enforcement of the criminal law against wrongdoers do not, and cannot, use the full weight of these harsh sanctions, we should consider that very carefully. Someday, many more Americans will look back and ask how we could have locked up so many people for so long for such crimes. So long as the law compels prosecutors to select some couriers, lookouts, and drivers to face life sentences because they work for an ambitious boss, narcotics sentences will pose a serious challenge to our long-cherished goal of doing justice through the law.

---

204. Bowman and Heise note:

Our system invests federal judges, federal prosecutors, the federal criminal defense bar, and federal probation officers with immense collective power based in large measure on the faith that these are people of intelligence and good judgment. The emergence of a widely-shared judgment among these trustees of the federal criminal justice system is a fact of which Congress and the Commission must take account in making drug sentencing policy.

Bowman & Heise, supra note 44, at 558-59.