

# Chapter Three: Presentencing, Inter-Judge, and Regional Disparity

## A. Introduction

Eliminating unwarranted sentencing disparity was the primary goal of the Sentencing Reform Act [SRA]. The legislative history of the SRA, reviewed in Chapter One, devotes more space documenting the “shameful disparity” constituting a “major flaw in the existing criminal justice system” than on any other aspect of pre-guidelines sentencing practices.<sup>104</sup> The SRA directs the United States Sentencing Commission to establish policies and practices that will “provide certainty and fairness in meeting the purposes of sentencing, avoiding unwarranted sentencing disparities among defendants with similar records who have been found guilty of similar criminal conduct while maintaining sufficient flexibility to permit individualized sentences when warranted by mitigating or aggravating factors not taken into account in the establishment of general sentencing practices.”<sup>105</sup> The Commission is also directed to “assure that the guidelines and policy statements are entirely neutral as to the race, sex, national origin, creed, and socioeconomic status of offenders.”<sup>106</sup>

The following two chapters review the empirical evidence concerning whether these central goals of sentencing reform have been achieved. This chapter focuses on inter-judge and regional differences and the impact of presentencing stages on sentencing uniformity. The next chapter investigates potential sources of racial, ethnic, and gender disparity.

### 1. *Definitions of Disparity*

***Form and substance.*** While there is widespread agreement that unwarranted disparity should be eliminated, there is less agreement on how to define it. Similar treatment for similar offenders and different treatment for different offenders is the hallmark of fair sentencing. But this formal definition is incomplete because it does not tell us how to classify offenders as similar or different (Cole, 1997). We need to identify which characteristics of offenses and offenders are relevant to our sentencing goals to know how to classify offenders.

In the federal system, sentencing goals are supplied by the statutory purposes of sentencing found at 18 U.S.C. § 3553(a)(2), which were prioritized by the federal sentencing guidelines (Hofer

---

<sup>104</sup> SENATE REPORT, *supra* note 1.

<sup>105</sup> 28 U.S.C. § 991(b)(1)(B).

<sup>106</sup> 28 U.S.C. § 994(d).

& Allenbaugh, 2003). The guidelines place primary importance on proportionate punishment—fitting the severity of punishment to the seriousness of the offense. Offense characteristics bearing on the *harms* caused by the offense and the offender’s *culpability* for those harms are especially relevant to assessing offense seriousness. The need to protect the public from additional crimes by the offender makes the offender’s *risk of recidivism*, as measured by their criminal history, also highly relevant. Unwarranted disparity is eliminated when sentencing decisions are based only on offense and offender characteristics related to the seriousness of the offense, the offender’s risk of recidivism, or some other legitimate purpose of sentencing.

***Common sources of unwarranted disparity.*** In the debates leading to passage of the SRA Congress identified differences among judges in sentencing philosophy and, to a lesser extent, differences among regions in sentencing practices as common sources of unwarranted disparity. Research evidence demonstrated that philosophical differences among judges affected the sentences they imposed, and that sentences varied significantly depending on the judge to whom an offender was assigned. The Federal Judicial Center’s Second Circuit Study—cited in the legislative history of the SRA—found dramatic differences among judges in the sentences they imposed on an identical set of hypothetical offenders (Eldridge & Partridge, 1974). Judges were sent presentence reports based largely on 20 actual federal cases representing a range of typical offenses and were asked what sentences they would impose. Differences of several years were common; in one case more than 17 years separated the most severe from the least severe sentence. The data showed that handfuls of judges were consistently more severe or more lenient than their colleagues. More important, judges varied in their approaches to particular crime types. Some judges treated drug traffickers relatively leniently while sentencing white collar offenders more harshly; other judges displayed the reverse pattern. Forst and Wellford (1981) also found significant inter-judge disparity and analyzed the role played by each judges’ sentencing philosophy in greater detail. Using a sample of 264 federal judges sentencing a different series of hypothetical cases, they found that judges who were oriented towards utilitarian goals (incapacitation and deterrence) gave sentences at least ten months longer on average than judges who emphasized other goals.

In addition to differences in philosophy among individual judges, several studies of pre-guidelines sentencing found geographical variations in sentencing patterns, suggesting that different political climates or court cultures can affect sentences. Research sponsored by the Department of Justice in the 1970s showed that judges placed differing importance on various factors depending on the region in which they sat (Sutton, 1978; Rhodes & Conly, 1981). Regional differences arise not just from the exercise of judicial discretion, but also from differences in policies among U. S. attorneys and in the practices of individual prosecutors.

## **2. *Increased Transparency of the Sentencing Decision***

Sentencing may now be the most transparent part of the criminal justice system. Not only is sentencing done publically in open court, with factual findings and determinations of law made on the record, but a detailed database of offense and offender characteristics and the judge’s decisions are compiled by the Sentencing Commission. With the exception of confidential

information contained in the presentence report (such as an offender’s medical history) most of this information has been made available to researchers and to the public.

We know more about the federal sentencing process today than ever before. One measure of this growth in understanding is what researchers call the “percentage of variance accounted for” by statistical models of the sentencing decision. Variance is a statistical term for the total of all variations in a group of sentences from the average sentence. Prior to the guidelines, researchers typically could account for 30 to 40 percent of the variance in sentences. Today that percentage has risen to over 80 percent, primarily because the factors that determine sentences are in large part identified in the guidelines themselves and in the data the Commission collects. Sentencing is the most-studied stage of the criminal justice process, and investigations of sentencing disparity are the most common subject of empirical inquiry in part because of this transparency.

## **B. Disparity Arising at Presentencing Stages**

The SRA focused primarily on sentencing, but Congress and the Commission recognized from the beginning that sentencing could not be considered in isolation. Decisions regarding how investigations should be conducted, what charges to decline or dismiss, what plea agreements to reach, and other decisions made prior to conviction and sentencing can all affect the fairness and uniformity of sentencing.

As described in Chapter One, the SRA directed the Commission to develop several mechanisms to monitor, and if necessary, control some of the effects of presentencing stages. The Department of Justice and the Judicial Conference of the United States also developed policies and procedures designed to ensure that the guidelines were not undermined by charging or plea bargaining variations. In 2003, the Department reaffirmed its belief that “[j]ust as the sentence a defendant receives should not depend upon which judge presides over the case, so too the charges a defendant faces should not depend upon the particular prosecutor assigned to handle the case.”<sup>107</sup>

The detailed federal sentencing guidelines take into account a large number of aggravating and mitigating factors in order to precisely tailor the severity of punishment to the seriousness of the crime and the dangerousness of the offender. Presentencing decisions can result in punishment that is either more severe or more lenient than the guideline sentence that would otherwise apply to the case. Uneven charging or plea agreements that fail to fully account for offenders’ criminal conduct can result in sentences that are both disproportionate to the seriousness of the crime and disparate among offenders who engaged in similar conduct. As reviewed below, several mechanisms leading to this type of disproportionality and disparity have been identified by researchers and by participants in the sentencing process.

---

<sup>107</sup> Ashcroft Charging Memo, *supra* note 24, September 22, 2003, at 2.

Disparity arising at presentencing stages has been described as often occurring “below the radar screen” (Saris, 1997). Plea and sentencing procedures require that plea agreements be disclosed to the court, but not all the reasons for presentencing decisions are necessarily stated on the record, and fewer public documents and statistical data exist to investigate the reasons for, and effects on sentences of, decisions to decline or bring charges, or to stipulate to particular sentencing facts or guideline provisions. Some commentators have called circumvention of the guidelines that would properly apply to a case a form of “hidden departure” (Wolf & Broderick, 1991; Schulhofer & Nagel, 1997; Hofer et al., 1999; Berman, 2000). Avoiding potentially applicable penalty statutes or circumventing applicable guidelines may result in sentences, in some cases, that are better suited to achieve the purposes of sentencing than the sentence that would result from strict adherence to every applicable law. But unlike judicial departures—which require reasons stated on the record and reviewable on appeal—presentencing decisions may open a gulf between sentencing “by the book” and sentencing “by the bargain,” which can undo the transparency and uniformity intended by the SRA (Freed, 1992).

The extent to which concerns about the effects of presentencing decisions have proven valid is an important issue for federal sentencing today and in the years ahead. The data available to assess these effects are not as detailed and complete as data on the sentencing decision itself. The data that are available, however, suggest that presentencing stages remain important in achieving sentence uniformity, and that some of the components of guidelines implementation that were designed to ensure uniformity have proven inadequate to the task or have not worked as intended.

## ***1. Presentencing Stages That Can Affect Sentencing Uniformity***

Commentators have identified several stages and decision points prior to sentencing that can affect the uniformity of sentencing decisions.

***Investigation techniques.*** Under the guidelines, sentences are based on detailed facts concerning offenders’ criminal conduct. The links between sentences and, for example, the quantity of drugs or money involved in the offense, the presence of a gun, or the location of the crime, are widely known to prosecutors and police. This enables them to predict the likely sentence based on the facts developed and brought forward at sentencing. Berlin (1993) argues that manipulation of defendant’s sentencing exposure during the investigation phase is a significant source of continuing disparity in the federal system. For example, rather than arrest a drug seller when his crime first becomes known, police could choose to make additional purchases until the quantity of drugs involved reaches the amount needed to trigger the sentence the police believe appropriate (Zlotnick, 2004). Judges have recognized a ground for departure from the guidelines for some types of police conduct—“sentencing entrapment”—when, for example, police induce a defendant to cook powder cocaine into crack cocaine in order to qualify the defendant for a harsher penalty (Fisher, 1996).

***Charging decisions.*** Decisions about which charges to decline or bring against a defendant have binding consequences for the final sentence. The guidelines were designed to minimize the effects of uneven charging decisions in many circumstances, as described in Chapter One. But in

some cases the statutory minimum and maximum penalties provided for the counts of conviction constrain the judge's discretion and trump guideline range that would otherwise apply to the offenders' conduct (Nagel & Schulhofer, 1989, 1992).

Mandatory minimum statutory penalties sometimes require a sentence *above* the otherwise applicable guideline range. For example, many offenders who are convicted of trafficking drug amounts just above the five- and ten-year thresholds cannot receive the full benefit of the guidelines' adjustments for acceptance of responsibility or mitigating role in the offense, because they do not qualify for a waiver of the mandatory minimum penalty under USSG §5K1.1 or the "safety valve." Charging an offender with possession of a firearm in furtherance of a drug trafficking crime under 18 U.S.C. § 924(c) usually, although not always, results in a sentence above the guideline range that would apply if the drug trafficking guideline's enhancement for possession of a firearm were applied instead (Hofer, 2000). Charging several counts of 18 U.S.C. § 924(c) in the same indictment—so-called "count stacking"—can result in sentences dramatically higher than the otherwise applicable guideline range, because mandatory prison terms are increased from at least five to at least twenty-five years for each subsequent conviction, even if sentenced at the same time, and must be imposed to run consecutively (Etienne, 2003). Including a money laundering count in an indictment once substantially increased the guideline range for conduct that might otherwise have been appropriately sentenced under a different guideline (USSC, 1997). (A 2001 amendment to the guidelines reduced, but did not eliminate, this effect of charging decisions.)<sup>108</sup>

In other cases, use of charges that understate the true offense conduct—for example, charging drug possession or use of a communications facility instead of drug trafficking—caps the statutory sentencing range *below* the level required by the guidelines for offenders' real offense conduct (USSC, 1995). The Commission has more than quintupled the number of cross-references to the guidelines through the years based on research findings that some offenders guilty of serious crimes, such as aggravated sexual abuse, were being charged and sentenced for less serious crimes like statutory rape or abusive sexual contact (USSC, 1992). Despite these attempts to undo the effects of undercharging, some evidence suggests that cross references are viewed as optional by some prosecutors and courts and are not always used as intended (USSC, 1996, 1997).

***Plea bargaining.*** Agreements not to charge or to dismiss charges are often made as part of plea negotiations between the parties. Plea agreements can be reached before indictment or between the time of indictment and sentencing. Plea agreements also may contain stipulations that a particular sentencing factor or provision of the sentencing guidelines does or does not apply to the case, or recommendations that a specific sentence is appropriate.<sup>109</sup> These stipulations and recommendations may be either non-binding or binding. Courts may reject plea agreements, but once an agreement with a binding recommendation is accepted, the court is obligated to sentence in accordance with the agreement. If an agreement is rejected, the court must personally advise the

---

<sup>108</sup> USSG, App. C, Amend. 634 (Nov. 1, 2001).

<sup>109</sup> Fed. R. Crim. P. 11(c)(1)(B) and (C).

defendant that the sentence may be less favorable than the agreed-upon sentence and give the defendant the opportunity to withdraw his or her guilty plea.<sup>110</sup>

Early in the guidelines era, Commission-sponsored research suggested that plea bargaining was leading to circumvention of the guidelines in a significant number of cases (Schulhofer & Nagel, 1989, 1992, 1997). In addition to charge bargaining, researchers reported that new types of plea agreements had developed under the guidelines, including fact bargaining and date bargaining (Alschuler & Schulhofer, 1989). Fact bargains include agreements on sentence-relevant considerations, such as the amount of drugs involved in an offense or the presence of a firearm. Date bargaining occurs when the parties negotiate over when an offense occurred, in order to use a particular edition of the *Guidelines Manual* that is more favorable to a defendant than a later edition.

While charging and plea bargaining are officially regulated by nationwide DOJ policies, researchers reported that in practice these policies were less determinative of prosecutorial conduct than internal U.S. Attorney's office policies. Judicial scrutiny of plea agreements, supported by probation officers' independent presentence investigations, were often inadequate to control plea bargaining because both judges and probation officers were heavily dependent on the information provided by the prosecutor in a given case. In addition, resource limitations and a reluctance to reject agreements, for a variety of reasons discussed further below, made judicial rejection of plea agreements that undermined the guidelines relatively rare.

***The presentence investigation.*** Probation officers are responsible for conducting the presentence investigation and informing the court in the presentence report of all the facts relevant to guidelines application. The probation officers' review of the offense conduct and explicit analysis of the impact of the plea agreement on the sentence helps inform judges' review of plea agreements. This investigative function that helps ensure that the guidelines are faithfully applied has led probation officers to be called the "guardians of the guidelines" (Bunzel, 1995).

Surveys of probation officers have consistently found variations in how offense conduct is investigated (FJC, 1990; USSC, 1991; Bowman, 1996). In some districts, probation officers conduct independent investigations, even interviewing prosecution witnesses and examining taped conversations or laboratory reports. In other districts, probation officers rely on the prosecution's version of the offense, even incorporating the government's written version of the offense directly into the "Offense Conduct" section of the presentence report. The amount of relevant conduct outside the counts of conviction, or potential grounds for upward or downward departure, uncovered by the probation officer depends on the intensity of the presentence investigation (Zlotnick, 2004). If these investigations are waived or abbreviated, relevant differences among offenders may go undiscovered and dissimilar offenders may be treated similarly.

***Filings of motions and notices.*** Statutory provisions give prosecutors sole discretion to seek certain increases or reductions of sentences. For example, 21 U.S.C. § 841 provides lengthier

---

<sup>110</sup> Fed. R. Crim. P. 11(c)(5).

mandatory minimum prison terms for offenders who commit certain drug trafficking offenses subsequent to a prior conviction for a drug felony in either state or federal court. To obtain the increased penalties, 21 U.S.C. § 851 provides that prosecutors must file notice of their intention to seek the enhancement prior to the sentencing hearing. Similarly, departures from the guidelines for a defendant's substantial assistance in the prosecution of other persons under USSG §5K1.1, 18 U.S.C. § 3553(e) and Federal Rules of Criminal Procedure, Rule 35(b) may be granted only upon motion of the government. (These motions are the only means by which a mandatory minimum penalty can be waived, other than the "safety-valve" exception at 18 U.S.C. § 3553(f) for certain first-time nonviolent drug offenders.) More recently, the PROTECT Act gave prosecutors sole authority to move for an additional reduction of one offense level for offenders who accept responsibility for their crime in a timely manner, thereby saving the government the cost of preparing for trial—a decision that had previously been left to the court. In addition, the departure of up to four levels below the otherwise applicable offense level authorized by the PROTECT Act for offenders who participate in an authorized early disposition program may be granted only upon motion of the prosecutor.<sup>111</sup>

## ***2. Surveys Suggest Sentencing Disparity Results from Presentencing Stages***

Preliminary evidence suggesting that sentencing disparity results from presentencing decisions comes from surveys of court practitioners. Several times in the life of the guidelines, researchers have asked practitioners to complete detailed questionnaires on the practical operation of the sentencing guidelines. While the results of these surveys are not strictly comparable due to differences in the wording of questions, the surveys reveal the perceptions and concerns of court practitioners and how those concerns have evolved over the guidelines era.

The earliest warning that charge and sentence bargaining were persisting into the guidelines era came from a survey of judges in one circuit (Alschuler & Schulhofer, 1989). Judges reported that the frequency of charge and sentence bargaining was roughly the same during the guidelines era as before, and that new forms of plea bargains were being developed. Just over two years later, as part of its Four-Year Evaluation, the Commission conducted its own survey in which judges were asked the effects of presentencing bargaining on sentencing disparity (USSC, 1991). However, the majority of judges reported that pre-indictment or post-indictment plea agreements were a source of unwarranted disparity in only some or a few cases. Only a minority of judges supported additional regulation of plea agreements beyond the policy statements contained in Chapter Six of the *Guidelines Manual*. Many other members of the court community, however, including probation officers and defense attorneys, identified presentencing decisions as a source of unwarranted sentencing disparity.

Two more surveys of court practitioners were conducted in the mid-1990s. The Commission's Probation Officers Advisory Group conducted a nationwide survey of federal

---

<sup>111</sup> USSG App. C, Amends. 649 (April 30, 2003) and 651 (October 27, 2003); Pub. L. No. 108-21, §§ 401(g), 401(m)(2)(B), 117 Stat. 650 (April 30, 2003).

probation offices and received responses from eighty-five districts (Bowman, 1996). Although not without methodological limitations (Berman, 1996), the survey reported some troubling findings. Just over half of the districts reported that “when guideline calculations are set forth in a plea agreement, they are supported by offense facts that accurately and completely reflect all aspects of the case.” But 43 percent of the districts reported that this was true just half the time or less. Probation officers reported preparing presentence reports that described the real offense conduct in almost all cases, but officers relied to a great extent on information supplied by prosecutors. In some districts and cases, respondents indicated that prosecutors tried to limit or manipulate information used in applying the guidelines. In a significant number of districts, probation officers reported that the court would usually or nearly always defer to the plea agreement when it conflicted with information in the presentence report.

The Federal Judicial Center conducted a survey of chief probation officers as well as Article III judges in 1996. The findings showed that “respondents believe much of the discretion that resided with judges before the guidelines has been shifted to prosecutors” (FJC, 1997, p. 6). About three-quarters of district judges and over half of chief probation officers reported that prosecutors had more influence on the final sentence than did judges. The vast majority of respondents reported that plea agreements in their district contained stipulated facts. More than a quarter of the judges reported that plea stipulations understated the offense conduct somewhat frequently or very frequently, while another 12 percent said they did so about half the time. Judges reported that they did sometimes “go behind” the plea agreements to examine underlying conduct, but they reported doing so “infrequently.” In contrast to the 1991 survey, 73 percent of judges felt that plea agreements were a hidden source of unwarranted disparity.

### ***3. Field Studies Suggest Sentencing Disparity Results from Presentencing Stages***

More evidence that disparity arises at presentencing stages comes from field studies conducted in several federal districts. This research suggests that different districts have evolved different “adaptations” to the guidelines system and to caseload pressures and other local conditions (Braniff, 1993; Bersin & Feigin, 1998). These various adaptations may be more or less formalized and regularized within a given district, and may be developed by U.S. Attorneys in each district with or without coordination with local judges and probation officers. The various types of “fast-track” programs that were developed in several districts beginning in the late 1990s are an example of a relatively formalized adaptation. The provisions of the PROTECT Act and recent initiatives of the Department dealing with early disposition programs are an attempt to centralize and regulate these mechanisms (DOJ, 2003).

Some districts control their workload with strict intake and charge declination policies, declining to prosecute, for example, marijuana cases that involve less than a ton of drugs—an amount that would be a major federal case in another district (Gleeson, 2003). Still other districts utilize post-indictment charge bargaining, fact bargaining, or other plea agreements to move their cases and obtain defendant cooperation, either as part of a systematic program or on a more ad hoc



basis. One district may employ liberal use of section 5K1.1 motions and departures, while a neighboring district achieves a similar overall departure rate through use of departures for mitigating circumstances other than substantial assistance (Farabee, 1998). Comparing data from four different federal courts, Storto (2002) suggested that different paths—for example, systematic charge bargaining in immigration cases in one district and departure bargaining in another—can sometimes lead to similar results.

One of the earliest and most comprehensive series of field studies was the work of Nagel and Schulhofer mentioned above (Schulhofer & Nagel, 1989, 1992, 1997). They concluded that circumvention of the guideline sentence was common, but that such circumvention was not necessarily “wrong” but “a covert vehicle for downward departure.” These hidden departures were motivated by a variety of reasons, including efforts to save time and resources and to provide incentives for defendant cooperation in addition to the incentives already included within the guidelines. In addition, several areas where the guidelines lacked flexibility were identified, which caused prosecutors, defense attorneys, and judges to search for ways to circumvent the guidelines’ strict requirements. These areas included an overemphasis on harm- and quantity-driven offense characteristics, a relative neglect of offender characteristics, and overall severity levels required by statutory minimum penalties and the guidelines pegged to them that were regarded by a significant number of judges and prosecutors as unnecessarily harsh in some cases, particularly for drug trafficking offenses. The authors concluded that prosecutorial discretion “if unchecked, has the potential to recreate the very disparities that the Sentencing Reform Act was intended to alleviate” and they warned that the system for regulating plea bargaining—relying on 1) probation officers’ investigations, 2) judicial review of plea agreements, and 3) Department of Justice charging and plea policies—might prove ineffectual.

More recently, Marks (2002) studied the effects of prosecutorial decisions in one district court, focusing on these same three mechanisms designed to help control disparity, as well as the relevant conduct rule. Interviews revealed that key participants in the sentencing process were generally unfamiliar with the contents of the policy statements in Chapter Six governing judicial review of plea agreements. “Informational asymmetry” between the government and the court made it unrealistic for probation officer investigation to fully inform the court about offenders’ real offense conduct. Implementation of the relevant conduct provision was further hindered by ambiguity in the language of the rule, discomfort with the role of law enforcement in establishing relevant conduct, and discomfort with the severity of sentences that often result from inclusion of all relevant conduct in guidelines determinations. Department of Justice policies then in place were also viewed as ineffective at achieving charging and plea bargaining uniformity.

A former prosecutor and current federal judge has argued that regional disparities in prosecutorial conduct are endemic and may be impossible to eliminate (Gleeson, 2003). Using drug couriers as an example, the judge demonstrated how two similar couriers arriving into the country in two different districts are subject to penalties over fifty percent higher in one than the other, due to the existence in one district of an agreed-upon program of maximum reductions for role in the offense for drug couriers.

The successful effort to restrain judicial discretion...has not produced a system in which similarly situated offenders are treated alike. Prosecutors have always been vested with ample discretion, and the Guidelines' diminution of the power of judges further enhanced the power of federal prosecutors. Differences in the exercise of that discretion among U.S. Attorneys, and by individual U.S. Attorneys in specific cases, have resulted in the differential treatment of similar cases, and account for the lion's share of the remaining disparities in federal sentencing (*id.* at 1701).

Judge Gleeson advised the Commission not to concern itself with presentencing disparity, however, since regional differences in public attitudes toward different types of crime and other local conditions make disparate practices inevitable. He concluded this does "not mean that the Guidelines have failed to achieve their essential goal" of controlling judicial discretion (*id.* at 1711).

#### **4. *Quantifying the Extent of Disparity Arising at Presentencing Stages***

**Limited data.** While surveys and field research suggest that unwarranted disparity arises at presentencing stages, such evidence can be challenged as anecdotal and impressionistic. Quantifying the effects of presentencing decisions is hampered by a lack of systematic data on police and prosecutorial practices. Only a few numerical estimates have been attempted. Based on their field studies in ten federal districts early in the guidelines era, Schulhofer and Nagel (1997, p. 1284) estimated that circumvention of the guidelines occurred in 20 to 35 percent of cases. The only other attempt to quantify the exact impact of plea bargaining through statistical analysis was conducted by the Commission in its Four-Year Evaluation. The Commission reported that in 14 percent of all guilty plea cases sentenced in 1989, the plea agreement resulted in a sentence below the minimum of the original guideline range.

These early estimates are unreliable bases for quantifying the precise impact of presentencing stages on sentencing today and more research is sorely needed. Recent revisions to Department of Justice policies, which reiterate that plea agreements are to be placed on the record<sup>112</sup> and forwarded to the Commission,<sup>113</sup> may help facilitate additional research in the coming years. Yet the existence of pre-indictment bargaining, limitations in the ability of probation officers to investigate and report offenders' real offense conduct, and judicial inability or unwillingness to review and reject plea agreements that understate the real offense will continue to hamper research.

**Research based on case documentation submitted to the Commission.** The Commission has periodically used the case documentation it receives on the vast majority of cases sentenced under the guideline to shed light on the sentencing effects of charging and plea bargaining decisions. Original and superceding indictments, plea agreements, and information provided by probation

---

<sup>112</sup> Ashcroft Charging Memo, *supra* note 24, September 22, 2003, at 5.

<sup>113</sup> The PROTECT Act amended 28 U.S.C. § 994(w) to require the Chief Judge of each district to submit a report of each sentence to the Commission, including any plea agreement.

officers in presentence reports are studied to reconstruct, as best as possible, offenders' real offense conduct and to compare it with the conduct for which they were ultimately held accountable. This has proven some of the most difficult research undertaken at the Commission. First in 1990, as part of the Four-Year Evaluation, and again in 1995 and 2000, the Commission collected an Intensive Study Sample (ISS)—a representative, random sample of cases sentenced in each year. The ISS is discussed in greater detail in Technical Appendix D.

Research based on information provided in presentence reports has been challenged as not accurately reflecting, for example, evidentiary problems that may attend proof of criminal conduct in some cases. Thus, estimates of undercharging or fact bargaining may be overstated if they are based on criminal conduct described by probation officers for use at the sentencing hearing, where the rules of evidence and standard of proof are more lax, which could not have been readily proven at trial. On the other hand, probation officers report that information on potentially applicable charges is sometimes not provided to them or to the court, or is excised from the presentence report if it is not used to determine the final sentence. This would cause comparisons of the conduct described in the presentence report with the conduct used for sentencing to *understate* the extent of undercharging or fact bargaining. On balance, although imperfect, data on undercharging or fact bargaining derived from presentence reports are the most reasonable and best available to quantify how presentencing stages affect the uniformity of sentencing.

***Uneven use of statutory penalty enhancements based on prior record.*** Research over the past fifteen years has consistently found that mandatory penalty statutes are used inconsistently in cases in which they appear to apply. Early in the guidelines era, the Commission reported that, among all offenders who engaged in conduct that qualified them for a mandatory minimum sentence, only 74 percent were initially charged with a count carrying the highest mandatory penalty applicable to their conduct (USSC, 1991b, pp. 56-58). Only sixty percent were ultimately convicted and sentenced at this penalty level or above.

Perhaps the firmest evidence of uneven use of statutory penalties concerns 21 U.S.C. § 841, which doubles the minimum statutory penalty for drug trafficking offenders who have a previous conviction for a felony drug offense, as long as the government files notice of its intention to seek the enhancement. Because criminal records are relatively straightforward compared to evidence concerning drug amounts or other factors, evidentiary problems are unlikely to prevent prosecutors from seeking this enhancement in a large number of cases. Yet the enhancement is more often avoided than sought. In 1991, the Commission reported that the enhancement was applied in a minority of qualified cases (USSC, 1991(b), p. 57). Analysis of data from both the 1995 and 2000 ISS samples found that the proportion of offenders with prior felony drug convictions who received the enhancement was under seven percent (6.5% and 6.9%, respectively).

Department of Justice policies explicitly permit prosecutors to forego the enhancement “after giving particular consideration to the nature, dates, and circumstances of the prior convictions, and

the extent to which they are probative of criminal propensity.”<sup>114</sup> This policy, however sound in theory, vests in prosecutors discretion to make sentencing judgments that were traditionally vested in judges, and that the Commission was designed to make with the benefit of research and study. There is reason to believe that the criminal history guidelines, which were developed based on empirical evidence on the links between prior convictions and the likelihood of recidivism, are better able to identify high-risk offenders than prosecutors deciding whether to pursue mandatory penalty enhancements available in the statutes (Krauss, 2003; USSC, 2004).

***Uneven use of firearms enhancements.*** Research on sentencing for possession or use of a firearm during a drug trafficking or violent offense has also consistently found uneven use of the statutory enhancement found at 18 U.S.C. § 924(c), as well as the firearm offense level adjustments contained in the guidelines. In 1991, the Commission reported that among drug offenders, only about 45 percent who qualified for a mandatory penalty enhancement under 18 U.S.C. § 924(c) were initially charged under the statute. This firearms count was later dismissed for 26 percent of the offenders initially charged. Analysis of 1995 ISS data found that only 34 percent of offenders who qualified for the statutory enhancement based on use of a firearm received the enhancement. Thirty percent received the guideline SOC instead, while 35 percent received no weapon increase of any kind (Hofer, 2000). Offenders who had merely carried or possessed a firearm, as opposed to using it, were even less likely to receive the statutory enhancement. Notably, Blacks accounted for 48 percent of the offenders who appeared to qualify for a charge under 18 U.S.C. § 924(c) but represented 56 percent of those who were charged under the statute and 64 percent of those convicted under it.

Analyses conducted for this report found that in 2000, just 20 percent of offenders who used a firearm received the statutory enhancement, 35 percent received the SOC, while 49 percent received neither. (Percentages add to more than 100 because a small number of offenders received both the statutory enhancement and the SOC. These estimates have a margin of error of about plus or minus ten percent because they are based on a random sample of cases.) As in 1995, offenders who merely carried or possessed a firearm were even less likely to receive the statutory enhancement than those who used it. Data from 2000 also showed the same pattern of disproportionate overrepresentation of Blacks among qualified offenders who actually received the statutory enhancement.

There is little empirical research exploring why enhanced penalties are sought in some cases and not in others, or whether their use reflects legally relevant factors, extra-legal factors, or arbitrary variation. A re-analysis of the Commission’s 1991 data found that racial disparity in use of mandatory penalties disappeared after controlling for additional factors, including whether the offender had pled guilty (Langan, 1992). Field research has reported that defense counsel believe the existence of penalty enhancements that are applicable at the sole discretion of the government gives prosecutors tremendous bargaining power to encourage defendant cooperation and discourage zealous defense advocacy (Etienne, 2003). Without more complete data on the legitimate

---

<sup>114</sup> Ashcroft Charging Memo, *supra* note 24, September 22, 2003, at sec. I.B.5.

considerations that affect charging decisions, it is not possible to evaluate the reasons for relatively rare use of these enhancements or the disparities observed.

Department policy was recently clarified to give prosecutors additional direction regarding use of the statutory firearm enhancement. The policy directs charging one count of 18 U.S.C. § 924(c) in every case in which it is applicable. If an offender has three or more possible counts and the predicate offenses are crimes of violence, prosecutors are directed to charge and pursue the first two such counts. The policy is silent on what should happen to offenders with two possible counts, or where the predicate offenses are drug convictions. It also permits exceptions to these rules if an office is “particularly overburdened” and in other circumstances.<sup>115</sup>

***Statutory floors and statutory caps.*** *Not* seeking statutory minimum penalties can lead to *more* proportionate sentencing, because statutory penalties would often trump the otherwise applicable guideline range and prevent mitigating adjustments contained in the guidelines from being taken into account. From this perspective, high rates of circumvention of potentially applicable mandatory penalties may be desirable. Many offenders do not benefit from avoidance of the penalties in these circumstances, however. In 2002, ten percent of federal offenders (over 6,000) received sentences above the top of the guideline range that would otherwise have applied to their case because of a trumping statutory minimum penalty. For another five percent, a statutory penalty restricted the judge’s discretion above the minimum of the guideline range that would otherwise have applied. Hispanic offenders, who were forty percent of all offenders, were forty-nine percent of those whose guideline range was completely exceeded by a statutory minimum penalty. Drug trafficking and firearm mandatory minimum penalties are the primary cause of trumping. In a small number of cases, 80 defendants in 2001, stacking of firearm counts resulted in statutory penalties that far exceeded the otherwise applicable guideline range.

On the other hand, charging decisions sometimes limit offenders’ exposure to punishment *below* the guideline range that would otherwise apply to their offense. In 2002, 1,379 offenders were convicted of charges carrying a statutory *maximum* sentence that was below the bottom of the guideline range that applied to their offense. This was most often due to conviction of a less serious immigration offense than they actually committed, or conviction for use of a communication facility to commit a drug trafficking offense instead of drug trafficking itself.

Sentences that result from avoidance of applicable penalties may seem to those most familiar with a particular case sufficient to meet the purposes of sentencing and more appropriate than the penalty required by strict application of the statutes and guidelines. Present practices, however, which lead to strict application in some cases and avoidance in others, result in disparity that cannot be accounted for by existing data and may be unwarranted. The fact that charging decisions disproportionately disadvantage minority offenders is further reason for additional research.

---

<sup>115</sup> *Id.* at sec. I.B.6.

***Departures pursuant to plea agreements.*** Finally, evidence that plea bargaining has resulted in unwarranted disparity is found in data that “pursuant to a plea agreement” has been the most or second most-frequently cited ground for downward departures in recent years (USSC, *Sourcebooks*, 1997-2002, Tbl. 25). While the policy statements in Chapter Six have always attempted to prevent judicial acceptance of plea agreements that undermine the guidelines, these policy statements were amended in 2003 to reiterate that the fact of an agreement alone is not sufficient to justify downward departure absent other mitigating circumstances.<sup>116</sup> Whether this change will be sufficient to reinvigorate the standards for acceptance of plea agreements, which field research suggests are largely unknown and widely disregarded, is an important question for the future.

One danger is that restriction of explicit downward departures will lead to an increase in “hidden departures” achieved through fact bargaining or other methods that fall “below the radar screen.” Quantifying the extent of fact bargaining is among the most difficult research issues because the effects of the bargain are built into the offense level reported to the Commission. Only by inclusion of all real offense conduct in the presentence report can the extent of fact bargaining be detectable to researchers.

## ***5. Presentencing Stages, Disparity, and the Mechanisms Designed to Control It***

Although a lack of data raises a serious obstacle to quantitative research, a variety of evidence suggests that disparate treatment of similar offenders is common at presentencing stages. Disparate effects of charging and plea bargaining are a special concern in a tightly structured sentencing system like the federal sentencing guidelines, because the ability of judges to compensate for disparities in presentence decisions is reduced. While the guidelines contain some mechanisms to ameliorate the effects of disparate charging and plea bargaining practices—such as the relevant conduct and multiple count rules, and judicial review of plea agreements—some of these mechanisms are not working as intended. By their nature, some of these mechanisms tend to work in one direction. The relevant conduct rule, for example, can increase sentences to account for criminal conduct that was not charged or that was dismissed prior to sentencing. But there is no guidelines mechanism to decrease sentences for an offender who, for example, is convicted of several counts of 18 U.S.C. § 924 (c) and is therefore subject to multiple consecutive mandatory penalty enhancements. If some offenders are charged in this manner while other similar offenders are not, there is little a judge can do to compensate for the resulting sentencing disparity.

The remainder of this chapter is focused exclusively on sentencing. Uniformity is defined as similar treatment of offenders who appear to be similarly based on the charges of conviction and the facts established at the sentencing hearing. Achievement of the more ambitious goal of similar treatment of offenders who engage in similar real offense conduct will also depend on uniform treatment at presentencing stages.

---

<sup>116</sup> USSG, App. C, Amend. 651 (Oct. 27, 2003).

## C. Inter-judge and Regional Sentencing Disparity

In the legislative history of the SRA, Congress identified unwarranted sentencing disparity among judges and, to a lesser extent, disparity among regions, as particularly disturbing national problems.

Sentencing disparities that are not justified by differences among offenses or offenders are unfair both to offenders and to the public. A sentence that is unjustifiably high is clearly unfair to the offender; a sentence that is unjustifiably low is just as plainly unfair to the public. Such sentences are unfair in more subtle ways as well. Sentences that are disproportionate to the seriousness of the offense create a disrespect for the law. Sentences that are too severe create unnecessary tensions among inmates and add to disciplinary problems in the prisons.<sup>117</sup>

With over fifteen years of experience under the guidelines, it is fitting to evaluate the success of the guidelines system at achieving this goal and to identify any problem areas that may remain.

Analyzing sources of inter-judge and regional disparity is complicated because the potential sources are so many, varied, and interacting. Differences among judges in sentencing philosophy has long been identified as an important source of variation in sentencing (Hogarth, 1971; Carroll, 1987). Research sponsored by the Department in the 1970s showed that judges differed in the importance they placed on various factors depending on the region in which they sat (Sutton, 1978; Rhodes & Conly, 1981). Sentencing can be influenced by differences among the districts and circuits in their sentencing case law and “personas” (Demlietner, 1994). These, in turn, are influenced by the political climates of different regions of the country. A great deal of research has established the importance of the local norms of different district courts—what some researchers have called court communities (Eisenstein & Jacobs, 1977; Ulmer & Kramer, 1996; Ulmer, 1997). The norms of different courts are also influenced by practical constraints, such as court workload and the availability of different types of sentencing options.

The use of sentencing guidelines was intended to control the effects of philosophical differences among judges and varying local conditions.<sup>118</sup> But even under a detailed and binding system like the federal sentencing guidelines, differences might arise among judges in how they use the guideline range, the available sentencing options, or the departure power, all of which could result in disparity. This chapter begins by examining whether implementation of the guidelines reduced inter-judge and regional sentencing disparity. It then turns to an examination of the various sources of inter-judge and regional disparity that may remain in federal sentencing today.

---

<sup>117</sup> SENATE REP. NO. 225, 98th Cong., 1st Sess. 45-46 (1983).

<sup>118</sup> The legislative history of the SRA states “[f]or the first time, Federal law will assure that the Federal criminal justice system will adhere to a consistent sentencing philosophy.” *Id.* at 59.

## 1. *The Effects of the Guidelines on Inter-judge and Regional Disparity*

The “central question” about the success of sentencing reform is whether implementation of the guidelines reduced unwarranted sentencing disparity. In its 1992 evaluation of the guidelines system, the General Accounting Office declared that this question “remained unanswered” (GAO, 1992). Today, better data and methodological innovations permit a more complete answer. The results of the latest analyses indicate that the guidelines have significantly reduced inter-judge disparity compared to the preguidelines era. Although some inter-judge disparity remains, the influence of judges’ personal philosophy on sentencing decisions has been reduced.

Regarding regional disparity, however, the available data and the methods for analysis are less robust and the conclusions are less reassuring. The available evidence suggests that regional disparity remains under the guidelines, and some evidence suggests it may have even increased among drug trafficking offenses.

***The best and most recent statistical analyses indicate that the guidelines have significantly reduced inter-judge disparity compared to the preguidelines era.***

Rates of use of guidelines mechanisms for sentencing outside the presumptive guidelines range, such as downward departures for substantial assistance to the government or departures for other mitigating circumstances vary dramatically among the circuits and the districts. In addition, with passage of the PROTECT Act<sup>119</sup> Congress re-opened the question of what types of regional disparities are to be considered unwarranted by creating a new mechanism for regional variation, “early disposition programs.”<sup>120</sup>

## 2. *Evidence of Inter-judge and Regional Variation in the Preguidelines Era*

***Uncontrolled studies.*** Data showing wide variations in the percentage of offenders sent to prison by different judges or in different regions, or in the average length of prison sentences imposed, were well known in the years preceding guidelines implementation. Congress cited some of these data (Sutton, 1978) in the legislative history of the SRA.<sup>121</sup> However, as discussed further below, simple tabulations of variations in sentences do not demonstrate *unwarranted* disparity because different judges and different regions have different types of cases, with differing offense seriousness and offender criminal histories. Some variation in average sentences is fully warranted. Only by controlling for case differences can we determine how much, if any, of the variation was unwarranted.

---

<sup>119</sup> Pub. L. No. 108-21, § 401(m)(2)(B), 117 Stat. 650 (April 30, 2003).

<sup>120</sup> 28 U.S.C. § 991(b)(1)(B).

<sup>121</sup> SENATE REPORT, *supra* note 1, at 41.



***Primary judge effects and interaction effects.*** Congress was also presented data from experimental research, which controlled for case differences and isolated the disparity attributable to judges. The Federal Judicial Center’s Second Circuit Study (Partridge & Eldridge, 1974) found dramatic differences among judges in the sentences imposed on hypothetical offenders. Judges were sent presentence reports representing a range of typical offenses, based largely on actual cases. Differences of several years were common. In one case more than seventeen years separated the most severe from the least severe sentence. The data showed that differences in the *average* sentences—what researchers call the “primary judge effect”—was fairly small for the majority of judges, even though a handful of judges were consistently more severe or lenient than their colleagues. This average similarity masked more substantial underlying disparities, however. Judges varied significantly in their approach to different types of cases—what researchers call “interaction effects.” Some judges treated white collar offenders more harshly than their peers, but drug offenders less harshly, while other judges’ treatment was the opposite. A second study quantified these two different types of disagreement among judges (Forst & Wellford, 1981). Again using hypothetical cases, the researchers identified how much of the variance in sentences was due to offense and offender characteristics, and how much was attributable to judges. Twenty-one percent of the variance was attributable to the primary judge effect, while thirty-four percent was attributable to interaction effects. The researchers also demonstrated that differences among judges in sentencing philosophy helped explain their differences in sentencing decisions.

***Limitations of research using hypothetical cases.*** Research using hypothetical cases demonstrated that some disparity in sentences can be attributed to the judges to whom cases were assigned. However, critics have questioned whether these findings can be generalized to the real world (Stith & Cabranes, 1998; *see also* discussion of the limitations of using hypothetical cases to evaluate the guidelines in Hofer et al., 1999). Waldfoegel (1997) used data on real cases from one federal district to evaluate the extent of inter-judge disparity in the preguidelines era. Like previous researchers, he found that the primary judge effect was less important than the interaction effects, but these accounted for only 2.3 percent and 9 percent of the variance in sentences, respectively. Research using hypothetical cases may exaggerate the extent of inter-judge disparity in the actual caseload.

### ***3. Research Concerning the Guidelines’ Effects on Disparity***

The federal system has been evaluated perhaps more thoroughly than any other, by both the U.S. Sentencing Commission itself and outside researchers. However, this increased scrutiny did not initially result in consensus about whether disparity had been reduced. Early research, using a variety of research methods and assumptions, resulted in a spectrum of opinions that varied from those who believed disparity was reduced (USSC, 1991; Karle & Sager, 1991) to those who could not tell whether there had been significant change (GAO, 1991) to those who believed disparity had actually gotten worse under the guidelines (Heaney, 1991).

***Survey results indicate growing judicial support for sentencing guidelines.*** When the federal guidelines were adopted, many judges doubted that the guidelines would be effective in

reducing disparity (Alschuler & Schulhofer, 1989). As time passed and experience with the guidelines increased, judges began to have more favorable views on the guidelines system. A 1991 survey conducted by the Commission found that judges were evenly split between those that thought the guidelines increased disparity (31.8 percent), those that thought the guidelines decreased disparity (36.2 percent), and those that thought the guidelines had no impact on disparity (32.0 percent) (USSC, 1991). In later surveys, more respondents recognized that sentencing guidelines could be an effective tool to reduce unwarranted disparity. By 2001, more than a third (36.9 percent) of federal district judges indicated that the guidelines “almost always” avoided unwarranted sentencing disparity for similar offenders convicted of similar conduct. A similar proportion (32.1 percent) thought that the guidelines often avoided this form of disparity. Just about a quarter (25.4 percent) reported that the guidelines only sometimes avoided this disparity, and only a handful of judges (5.6 percent) reported that the guidelines rarely avoided disparity (USSC, 2002).

***Early empirical evaluations of the guidelines.*** While surveys provide insights into judges’ impressions of the effects of the guidelines, empirical research that examines data on changes in actual sentencing practices is necessary to assess if the guidelines have been successful. Even with such data, however, it is difficult to isolate the effects of the guidelines from the effects of other changes that occurred at the same time as guidelines implementation. Several early evaluations illustrate problems in isolating the effects of the guidelines from shifts in the types of cases sentenced in the preguidelines and guidelines eras. (Heaney, 1989; Karle & Sager 1991; GAO, 1992. *See* Hofer et al., 1999, for a review and critique of these early studies.)

The Commission’s previous attempt to evaluate the guidelines’ success at reducing unwarranted disparity was included in the Four-Year Evaluation (USSC, 1991, pp. 279-299). The report featured a “Distributional Analysis” that compared bank robbery, cocaine distribution, heroin distribution, and bank embezzlement cases sentenced during fiscal year 1985 with similar cases sentenced in the first months of full guidelines implementation. The Commission matched offenders from the two time periods on factors deemed relevant to sentencing, *e.g.*, the approximate amount of drugs, any injury caused to any victims, the defendant’s role in the offense and criminal record, and whether the defendant pled guilty or went to trial. Variations among the matched groups at each time period were compared in terms of the sentence imposed and the expected time the defendant would actually serve. Because only offenders who met the strict matching criteria could be included in the study, the number of defendants in each group was relatively small. However, the analysis showed that the distribution of sentences for each group under the guidelines was narrower than in the preguidelines era. The Commission concluded from these results that unwarranted disparity was reduced by the guidelines.

Several reviewers criticized the Commission’s methods, however, and questioned whether the study so clearly demonstrated success. (Tonry, 1997; McDonald & Carlson, 1993; Rhodes, 1992; Weisburd, 1992. For more thorough analysis of these criticisms, *see* Hofer et al., 1999.) At the request of Congress, the GAO re-analyzed the data on several offenses using somewhat different techniques. Their analyses replicated and confirmed the Commission’s basic findings, but the GAO concluded there was insufficient evidence to establish clearly that the guidelines had reduced

disparity (GAO, 1991). One of the problems identified by critics was the small and unrepresentative sample of cases used in the study, due to the use of matched groups to achieve comparability of cases. What was needed was a different method that could examine judge-created disparity in the caseload as a whole.

#### **4. *Recent Research Concerning the Guidelines' Effects on Disparity***

***The “natural experiment” method.*** In 1991, the economist Joel Waldfogel introduced a method that has since been used by researchers both inside and outside the Commission to evaluate the effect of the guidelines on inter-judge disparity (Waldfogel, 1991). This “natural experiment” method exploits a common court procedure: *random assignment* of cases to judges. By measuring variation in average sentences before and after guidelines implementation among judges in the same random assignment pool, the extent to which judges influence sentences can be quantified. Waldfogel’s initial study reported no decrease in disparity under the guidelines (Waldfogel, 1991). In 1997, Payne replicated Waldfogel’s approach in more districts, with mixed results (Payne, 1997). Limitations in these early studies, and the great promise of the natural experiment method, led researchers to extend the work using more recent data and more robust statistical models. (See Hofer et al., 1997, for further discussion of all studies using this methodology.)

***Later research using the “natural experiment” method.*** Commission staff published results of its research in the *Journal of Criminal Law and Criminology* (Hofer et al., 1999) and details of the analysis can be found in the Technical Appendix accompanying that article. The statistical model permitted quantification of inter-judge disparity across all the different cities included in the study, while comparing each judge only to other judges in the same city who were part of the same random assignment pool. Both the primary judge effect and interaction effects between judges and seven different offense types were studied. In addition, measures of the amount of variation among different regions were calculated. The magnitude of the influence of each of the explanatory factors during the two time periods was measured with the R-squared—the percentage of variance in sentences accounted for by each factor.

The study compared sentencing in fiscal years 1984-85, immediately before implementation of the guidelines, with 1994-95, after the guidelines were fully implemented. Two sets of analyses were conducted. To control for changes in the composition of the bench, the first analysis involved only judges who sentenced during both time periods. This limited the analysis to just nine cities that had at least three judges meeting this criteria, however. The primary judge effect accounted for 2.32 percent of the variation in sentences in the preguidelines era. Under the guidelines, this dropped to 1.24 percent, a reduction of almost half. The effect of judges was statistically significant at both time periods, but was substantially reduced under the guidelines.

The second analysis included all judges who were part of a random assignment pool involving at least three judges, regardless of whether they sentenced during both time periods. This expansion allowed the analysis to include 41 cities. The overall pattern of results was similar to the nine-city analysis. Again, the primary judge effect was significant at both time periods, but was

reduced under the guidelines—2.40 percent in the preguidelines era and 1.64 under the guidelines, a reduction of one-third. In this 41-city analysis, the effect of the guidelines appeared more limited and differences among cities appeared more important.

As in studies conducted in the preguidelines era, the interaction effects between different offense types and judges were substantially larger than the primary judge effect. Judges disagree about the appropriate sentence for specific cases to a greater extent than is revealed by differences in the primary judge effect, which measures only the “tip of the iceberg” of sentencing disparity.

To examine whether the effects of the guidelines varied for different offenses and in different cities, results were calculated for seven offense types separately. The results suggest that the effect of the guidelines has not been uniform. For most offenses, the judge effect decreased under the guidelines, but for robbery and immigration offenses the influence of judges increased.

Most troubling were changes in the influence of cities on sentences, which actually increased in the guidelines era. Almost all of the increase was found in drug trafficking offenses, where the city effect increased from 6.2 before guidelines implementation to 12.7 percent under the guidelines. This suggests more regional disparity in the sentencing of drug cases under the guidelines. Interpreting the city effect is difficult, however, because cases are not randomly assigned to cities. In addition, policy changes between 1984-1985 and 1994-95 could exacerbate the effects of the guidelines. For example, the greater emphasis on drug quantity in sentencing following enactment of the Anti-Drug Abuse Act of 1986 could create greater disparity between small cities and large cities, which are distribution centers for larger quantities. Note also that fiscal years 1994-1995 were prior to adoption of early disposition programs in some districts, which may have contributed to growing regional disparity in other types of cases, such as immigration.

***Research outside the Commission.*** A reduction in inter-judge disparity under the guidelines was found in a study by researchers outside the Commission, who also used the natural experiment methodology and a highly sophisticated statistical model. Anderson, Kling & Stith (1999) studied sentencing patterns in 26 cities over 12.5 years among judges who sentenced in both the preguidelines and guidelines eras. They conducted rigorous tests to confirm the randomness of assignment to judges and constructed a statistical model that allowed them to test the significance of changes in inter-judge disparity over the entire period of their study. These improvements in the model allowed them to detect “changes . . . more pronounced than the mixed results of previous studies” (p. 294).

In the preguidelines era, the expected difference in sentence lengths between two typical judges was about 17 percent of the average sentence length. Under the guidelines, this difference fell to 11 percent. Because average sentences are lengthier in the guidelines era, a given percentage of disparity among judges results in a larger absolute difference in months of imprisonment. Taking into account these changes, the authors report: “For 1986-87 when the mean sentence length was 29, the expected inter-judge difference was 4.9 months, which fell to 3.9 months in 1988-93 when the mean sentence length was 35.” Further tests indicated that the switch to guidelines, and not

changes in the composition of the bench or in the types of cases sentenced over the period of their study, accounted for the decrease in inter-judge disparity.

These authors suggest that the enactment of mandatory minimum penalties for drug offenses in 1986 may have contributed to the decrease in disparity for drug offenses. They also note that the effects of discretionary decisions by law enforcement officers and prosecutors at presentencing stages have greater influence under the guidelines system, since later actors have fewer opportunities to ameliorate any effect of disparate treatment prior to sentencing. Some of the inter-judge disparity in preguidelines sentencing was likely ameliorated by the parole guidelines, which affected the prison time actually served by offenders prior to implementation of “truth-in-sentencing.”

*The sentencing guidelines have reduced inter-judge disparity.* At this time, findings from research using the natural experiment method have not been challenged and it appears unlikely that a more powerful method for studying the effects of the guidelines on inter-judge disparity will be found. The convergence of findings by researchers both inside and outside the Commission lends additional credibility to the results. The conclusion is clear: the federal sentencing guidelines have made significant progress toward reducing disparity caused by judicial discretion.

*The sentencing guidelines have made significant progress toward reducing disparity caused by judicial discretion.*

## D. Continuing Disparity Under the Guidelines

Though clearly reduced by the guidelines, inter-judge sentencing variations that cannot be explained by differences in the caseload remain statistically significant today. Regional disparity also appears to remain and may have increased for some types of cases. How can disparity continue in a system of detailed and binding sentencing rules? The remainder of this chapter reviews the evidence that inter-judge and regional disparity continue to exist in federal sentencing and explores how it occurs.

### 1. Continuing Regional and Inter-judge Disparities

The natural experiment method is best for establishing whether inter-judge disparity has been reduced by implementation of the guidelines, but it is not as precise as other methods for measuring the amount of inter-judge and regional disparity today and comparing the effects of judges and regions with the warranted effects of legally relevant considerations, such as the seriousness of the crime and the criminal history of the offender.

*Uncontrolled comparisons.* The Commission’s annual *Sourcebook of Federal Sentencing Statistics* and additional information found at the Commission’s website (<http://www.ussc.gov/LINKTOJP.HTM>) contain a wealth of information on departure rates and

other sentencing variations that some commentators have used to question the guidelines' success (Miller, 2002; Mercer, 2003). These data compare the sentencing practices of each federal district with other districts and with national rates and averages. The data show that the types of sentences imposed and the average sentence lengths for offenders convicted of various types of crime vary among the districts. The rates of departure for substantial assistance or other mitigating or aggravating conditions also vary substantially among the districts and circuits.

These regional variations do not necessarily indicate *unwarranted* disparity, however, because different districts sentence different types of crimes within the general offense categories found in the reports. The types of fraud sentenced in the Southern District of New York (average fraud sentence 23.5 months) are different than the frauds sentenced in the District of North Dakota (average sentence 11.4 months). The guidelines themselves require different sentences for frauds involving different amounts of monetary loss, different numbers of victims, and many other specific offense characteristics.<sup>122</sup> Similarly, variations in the rates of a particular type of departure among different districts must be evaluated within a larger context of each district's distinctive adaptation to the guidelines system. Inferring unwarranted disparity from uncontrolled comparisons of average sentences or rates of departure may be erroneous.

***Multiple regression studies.*** One source of variation in sentences that is clearly warranted is differences in the types of cases sentenced by each judge and in each district. Researchers have sought to control for legally relevant differences among cases using the statistical technique of multiple regression. Many studies of racial and ethnic disparity have also included measures of the district and circuit in which each case was sentenced. These studies suggest that differences in offense seriousness, defendant criminal history, or other legally relevant factors account for the largest share of variation among cases, but that some statistically significant variation among regions remains unexplained.

Albonetti (1997) reported that the probability and length of imprisonment for drug offenses sentenced in the early years of the guidelines was affected by region in about half of the circuits, after controlling for offense level, criminal history points, and a number of other legally relevant factors. Everett and Wojtkiewicz (2002) grouped the circuits into five regions and reported harsher sentencing in the southern circuits and more lenient sentencing in the northeastern and western circuits. Kautt and Spohn (2002) reported a statistically significant effect in drug cases sentenced in 1997-1998 in a minority of circuits. These studies did not use the presumptive sentence method, discussed in greater detail in Chapter Four and in Technical Appendix D, to control for legally relevant differences among cases. This introduces some avoidable errors in the results, because the effects of trumping mandatory minimum statutes, mandatory firearm sentencing enhancements, and other legally relevant considerations binding on the judge were not properly specified.

---

<sup>122</sup> See USSG §2B1.1.

***A new approach: hierarchical modeling.*** Recent developments in the field of statistics (*see e.g., Raudenbush & Bryk, 2002*) have introduced researchers to a new method that may be superior to multiple regression for studying inter-judge and regional differences today. Hierarchical models avoid some types of statistical bias that can arise in multiple regression models and also permit researchers to explore the *direct effects* of factors on sentences and the *conditioning effects* of higher-level factors on lower-level factors. For example, the effect of district workload on the extent of departure in a district might be studied using a hierarchical model. Several studies have appeared in the last two years using these new methods to investigate regional variation in sentencing under the federal guidelines (Ulmer et al., 2001; Albonetti, 2003; Spohn, 2003). One such study has been published at the time of this writing (Kautt, 2002).

In a study of federal drug trafficking sentences, Kautt found variations between jurisdictions that could not be accounted for by the legally relevant differences included in her model. Districts and circuits both affected sentences, with the influence of districts more important than the influence of circuits. Districts also appeared to differ in the way that legally relevant factors influenced sentences. From these results Kautt concluded that “despite the federal system’s congressionally mandated return to determinate sentencing, extra-legal factors (specifically jurisdictional effects) continue to influence the federal sentencing system and its outcomes directly and indirectly . . . these findings indicate a far greater concern: that the mechanisms of federal structured sentencing may *foster* certain forms of extralegal sentence disparity” (Kautt, 2002, p. 659).

Hierarchical models are a powerful new tool for studying regional variation, but they are “delicate and complex” (Kautt, 2002, p. 645). Mis-specification of the relationship among explanatory factors might distort results in unpredictable ways. For this reason, using the presumptive sentence method to represent legally relevant considerations seems desirable. (Kautt considered but rejected the use of a presumptive sentence model for reasons that are unclear, *see id.* at 649, n. 15.) Additional research with hierarchical models using the presumptive sentence to control for legally relevant differences among cases could prove very useful.

***A new hierarchical analysis of inter-judge and regional disparity.*** For this report, a hierarchical model of the determinants of sentence length was developed, using the presumptive sentence to control for differences among cases in the legally relevant factors taken into account by the guidelines and mandatory minimum statutes. Both judges and districts were included as levels of analysis. All cases sentenced in 2001 with full information were analyzed, with the exception of cases handled by visiting judges. Appendix D contains additional details of this analysis.

The analysis showed that legally relevant differences among cases explain the vast majority of variation among judges and regions in sentence length. Fully 73 percent of sentence variation is accounted for by the guidelines and statutes. The amount of variation that is associated with judges or

***Research using the most current statistical models continues to show relatively minor inter-judge and regional disparity not explained by case differences.***

districts is relatively small. Of the 27 percent of variance that is not accounted for by the presumptive sentence, just 2.9 percent is associated with judges within each judicial district, and 2.8 percent is among judicial districts. The remaining is unexplained case variation. The judge and district levels provide relatively minor variation of sentences compared to case differences.

## **2. *Mechanisms for Disparity Within the Guidelines System***

Congressional attention has recently focused on potential disparity arising from varying downward departure rates for mitigating circumstances identified by judges (Mercer, 2003; PROTECT Act, 2003). Less attention has been paid to other potential sources of disparity. These include variation in rates of other types of departure, such as rates of departures for substantial assistance to the government or the extent of such departures for different forms of assistance (Saris, 1997; Maxfield & Kramer, 1998; Farabee, 1998). The PROTECT Act authorized a new ground for departure for defendants who participate in qualified early disposition programs. In addition, the guidelines give judges discretion over placement of the sentence within the guideline range, including, in some cases, whether to use a sentencing option such as probation.

***Relative contribution of different mechanisms to sentence variation.*** To assess the influence of each of these mechanisms on sentencing disparity, a multiple regression analysis of sentences imposed in 2001 was undertaken. Details of this analysis are provided in Technical Appendix D. In addition to the presumptive sentence, the analysis included variables indicating whether the case 1) received a sentence within the guideline range, but above the minimum of the range, 2) received an upward departure, 3) received a downward departure for a mitigating circumstance identified by the judge, or 4) received a downward departure for substantial assistance.

As in other analyses using the presumptive sentence, the guideline and statutory factors represented by the presumptive sentence accounted for the vast majority of variation in sentences in 2001—75 percent. (Slight differences in the amount of variation accounted for by the presumptive sentence are expected due to the different populations of cases included in each analysis.) Among the mechanisms,

***Among discretionary mechanisms within the guidelines system, substantial assistance departures contribute the greatest amount to variation in sentences, while judges' use of the guidelines range contributes the least.***

substantial assistance departures accounted for the greatest amount of the remaining variation in sentence length—4.4 percent. Other downward departures contributed 2.2 percent, while upward departures contributed just 0.29 percent. Only 0.07 percent of the variation was explained by use of the guideline range above the guideline minimum.

It may be surprising that substantial assistance departures account for so much more variability in sentence length than other types of downward departures, because the *rate* of the two types of departure are similar—17.1 percent and 18.3 percent, respectively (USSC, 2001, Tbl. 26).



However, the *extent* of departure for substantial assistance is on average far greater. The mean departure length for substantial assistance was 43 months (or a mean 56 percent decrease from the guideline minimum), while the mean departure length for other downward departures was just 20 months (or a mean 47 percent decrease from the lower average guideline minimum found among offenders receiving these departures).

The relative unimportance of placement within the guideline range above the minimum may also be surprising. Almost 38 percent of offenders received a sentence above the guideline minimum in 2001, and 13.8 percent were sentenced at the guideline *maximum* (USSC, 2001, Tbl. 29). But the average difference between the guideline minimum and the sentence imposed in these cases was just 6.8 months. It should be noted, however, that at the lower end of the sentencing table a six month difference may be the difference between a sentence of simple probation and six months in prison—a distinction of considerable importance to the offenders involved.

A total of just over 82 percent of the variation in individual sentence lengths could be explained by the model. This does not mean that 18 percent of the variation in sentences is unwarranted. The model only identified the relative importance of different mechanisms—how frequently the mechanism is used and how far from the presumptive sentence offenders affected by the mechanism are sentenced. It did not attempt to determine if the mechanisms were being used appropriately and uniformly. For this we must turn to research specific to each one.

### ***3. Departures upon Motion of the Government***

Several types of sentence reductions can be made only upon motion of the government. Departures from the guidelines and guideline adjustments for various forms of defendant cooperation—such as “substantial assistance in the prosecution of other persons” under USSG §5K1.1, 18 U.S.C. § 3553(e) and Federal Rules of Criminal Procedure, Rule 35(b); participation in an “early disposition program” under USSG §5K3.1 p.s.; and timely “acceptance of responsibility” under USSG §3E1.1(b)—may all be granted only upon motion of the government. Research on several of these mechanisms has revealed considerable regional variation, suggesting that uniform practices have not been in place. Policies put in place subsequent to the PROTECT Act are too new to be evaluated with the data available for this report.

***Downward departures for substantial assistance.*** Downward departures for substantial assistance to the government in the prosecution of other persons are made pursuant to USSG §5K1.1. The policy statement permits such a departure only upon the motion of the government, but it does not require that the judge depart whenever the government so moves. Research has shown, however, that judges almost always grant these departures when a motion is made (Maxfield & Kramer, 1997). The rates of substantial assistance motions vary among the districts. In 2002, the national rate was 17.4 percent of all offenders. Five districts had rates twice

***Rates of substantial assistance departures vary widely among the districts.***

as high as the national rate, with three districts having rates over 40 percent (USSC, 2002, Table 26). On the other hand, 12 districts had substantial assistance departure rates less than half that of the national rate, while three had rates of less than five percent.

Policy statement, section 5K1.1 sets out a non-exhaustive list of reasons for judges to consider when determining the appropriate extent of a reduction. However, no detailed nationwide policies governing how substantial assistance motions should be used, nor how the extent of the departure should be determined, have been promulgated by either the Department of Justice or the Sentencing Commission (Lee, 1997; American College of Trial Lawyers, 2001). Case law has established some principles for determining the extent of departure in some circuits.<sup>123</sup> The Commission has received limited reports of standardized discounts in some districts, although in other districts, once a motion is made the determination of the sentence is left entirely to the discretion of the sentencing judge. The majority of sentencing judges reported cases where a defendant had substantially assisted the government, but had not received a motion for a departure on that ground (FJC, 1997).

Given the wide variety of behaviors that can qualify a defendant for a substantial assistance departure, some commentators have suggested that courts are given insufficient guidance regarding the appropriate extent of departure (Berman, 2002; Bowman, 1999). Some have argued that substantial assistance departures are a source of continuing unwarranted disparities (Tease, 1992; Marcus, 1993; Gyurici, 1994; Lee, 1994, 1997), although others have cautioned that differences in rates of departure do not necessarily result in sentencing disparities (Weinstein, 1998; Storto, 2002).

Empirical research on substantial assistance departures is extremely difficult because detailed information on the most important legally relevant consideration—the nature of the defendant’s assistance to the government—is available only to the prosecutors familiar with the case. While U.S. attorneys offices are required to document the reasons underlying every substantial assistance motion (DOJ, 1992), these records are not collected into a comprehensive database that can be used for empirical analysis. However, what research has been done indicates that substantial assistance departures may be a source of continuing sentencing disparity.

A comprehensive study of substantial assistance departures was undertaken by the Sentencing Commission in the mid-1990s (Maxfield & Kramer, 1998). It included a survey of U.S. attorney offices’ policies on substantial assistance, site visits to eight districts, and an examination of the types of cooperation given by a random sample of defendants receiving substantial assistance departures, as determined by analysis of the presentence reports prepared in the case. The research uncovered irregular and inconsistent policies and practices among the various districts.

---

<sup>123</sup> See Federal Judicial Center, *Guideline Sentencing: An Outline of Appellate Case Law on Selected Issues*, § VI.F.2 (2002).

- (1) While every U.S. attorney office reported some review process for the approval of substantial assistance motions, and four out of five offices had written policies regarding the use of substantial assistance motions, review of a sample of cases showed that in practice districts frequently diverged from their stated policies.
- (2) Different U.S. attorneys offices were consistent in authorizing motions for offenders who provided information against other persons, or participated in the investigation of other persons, or testified against them. But different offices varied in whether and how they considered information offenders provided concerning their own criminal conduct.
- (3) Six out of ten offenders who provided some assistance did *not* receive a section 5K1.1 motion, suggesting that prosecutors generally require that the assistance be substantial.
- (4) Offenders at higher levels of a criminal conspiracy are *not* more likely to benefit from a departure for substantial assistance than are lower-level offenders. Although occurring in some specific cases, the so-called “cooperation paradox” in which more culpable offenders receive shorter sentences than less culpable offenders was not common.
- (5) Offenders providing similar types of assistance received varying magnitudes of departure. Offenders with longer presumptive guideline sentences tended to receive greater reductions.

A lower annual rate of substantial assistance departures for Blacks has been a consistent finding in the guidelines era. Minority defendants may, in fact, be less trusting of government officials, less willing to become “snitches” due to pride or fear of reprisal, or less well-positioned to provide useful information than White offenders. Maxfield and Kramer found that Whites and women were more likely to receive a motion after controlling for offense type, guideline range, weapon involvement, and a host of factors relevant to sentencing. However, a re-analysis by Langan (1996, 2001) found that part of the difference between the races was due to their different rates of pleading guilty, and that the statistical significance of the remaining difference was questionable. Neither of these studies could adequately evaluate whether there might be legitimate reasons for differences in substantial assistance departure rates among different groups due to lack of data on the nature of the assistance various offenders provide. Beginning in 1992, Department policy required prosecutors to “maintain documentation of the facts behind and justification for each substantial assistance pleading.”<sup>124</sup> No standards for how this information is to be recorded appear to have been promulgated and the data have not been compiled or made available to outside researchers.

---

<sup>124</sup> Terwilliger Bluesheet, *supra* note 43.

***Rule 35 sentence reductions.*** In addition to substantial assistance departures, U.S. attorneys offices can file motions with the court to have previously imposed sentences reduced based upon post-sentencing cooperation by an offender. These are known as Rule 35(b) motions. Data on the use of Rule 35(b) motions or on factors that might account for their use have been very difficult to obtain (Marcus, 1985). The Commission does not reliably receive reports on sentence reductions following the sentencing hearing, so analyses using the Commission’s monitoring databases cannot detect their effects on sentences. However, a recent paper using data from the Federal Bureau of Prisons presented the first empirical look at these reductions (Adams, 2003). It suggests that regional variations in practice found with §5K1.1 motions may be present with Rule 35(b) motions as well.

The number of offenders receiving Rule 35(b) sentence reductions has increased dramatically over the guidelines era, from 30 offenders in 1988 to 1,453 offenders in 2000, the last year for which data are available. The average extent of the reduction has remained fairly stable through the years, varying between 30 and 42 percent of the originally-imposed sentence, although the extent varies by district. Drug offenders are by far the most frequent beneficiaries of Rule 35(b), accounting for 80 percent of the reductions. The use of Rule 35(b) varies significantly among the districts. Most districts account for less than one percent of offenders receiving Rule 35(b) reductions in any given year, but in one recent year ten districts accounted for over two percent of offenders receiving the reduction while one district accounted for over 15 percent.

***Downward departures for participation in early disposition programs.*** The Department recently informed the Commission that prosecutors in certain districts have developed early disposition, or “fast-track” programs, that grant participating offenders sentencing concessions.<sup>125</sup> How many districts have employed these programs, and for how long, is not known. These programs offer defendants a sentence reduction, in the form of a downward departure, charge dismissal, or some other benefit, in return for the defendant’s waiver of certain procedural rights. These rights might include a prompt guilty plea, a waiver of appeal rights, and in cases involving non-citizens, the defendant’s agreement to immediate deportation. Practitioners and commentators have expressed concern that the presence of these programs in some districts, and their absence from neighboring districts, could lead to disparate sentencing outcomes for offenders convicted of similar conduct (USSC Hearing, 2003). The absence of reliable information on the types of cases which are, and which are not, sentenced pursuant to early disposition procedures prohibits analysis of the impact of these programs on sentencing disparity. But as discussed below, the presence of fact track programs in some districts explains a great deal of regional variation in downward departure rates.

The PROTECT Act sought to formalize and standardize these practices. Per the act, the Sentencing Commission authorized a downward departure from the guidelines of “not more than

---

<sup>125</sup> Letter from Eric H. Jaso, Counselor to the Assistant Attorney General, DOJ, to Hon. Diana E. Murphy, Chair, USSC, regarding “Fast-Track Program,” August 12, 2003.

four levels” for offenders who participate in these programs.<sup>126</sup> The Department has outlined criteria to be used to authorize early disposition programs in some districts.<sup>127</sup> Whether these developments will ensure uniformity of sentencing under these programs cannot yet be determined. Evaluation of these programs will be possible only if the Department and relevant U.S. attorney offices provide data which reveal the workings of the fast track process.

#### 4. *Variability of Within-Guidelines Sentences*

Congress recognized in the SRA that no set of rules could anticipate every circumstance relevant to the sentencing decision. In addition to authorizing departures in exceptional circumstances, Congress permitted the Commission to design guidelines that provide a limited range of prison time for each category of offender.<sup>128</sup> The Commission determined that in the lower zones of the sentencing table, judges should have discretion to sentence offenders to prison terms or to choose from a variety of sentencing options.<sup>129</sup> At the highest offense levels the guideline range is over six years and judges may impose sentences anywhere within it. Discretion within the guideline range permits consideration of subtle differences among offenses and offenders that are not considered by the guidelines, but that do not meet the exceptional standards for departure. Guidelines commentary encourages use of the range to take account of differences in offense seriousness in some circumstances.<sup>130</sup>

***Use of sentencing options.*** The Commission’s annual *Sourcebook* contains information on the use of the sentencing range and sentencing options for various types of offenses. Figure F from the 2002 *Sourcebook*, reproduced on the following page, shows the imprisonment rates of offenders who are eligible for a non-prison sentencing option for nine offense categories. Many offenders who could receive a sentence of probation under the guidelines are imprisoned instead. Imprisonment rates of probation-eligible offenders range from over 80 percent for immigration offenders (reflecting their frequent lack of a United States residence and imminent deportation) to about 20 percent for larceny offenders. All other offense types vary between 20 and 50 percent of probation-eligible offenders receiving imprisonment instead. These findings have remained fairly stable from year-to-year.

---

<sup>126</sup> USSG §5K3.1 (policy statement).

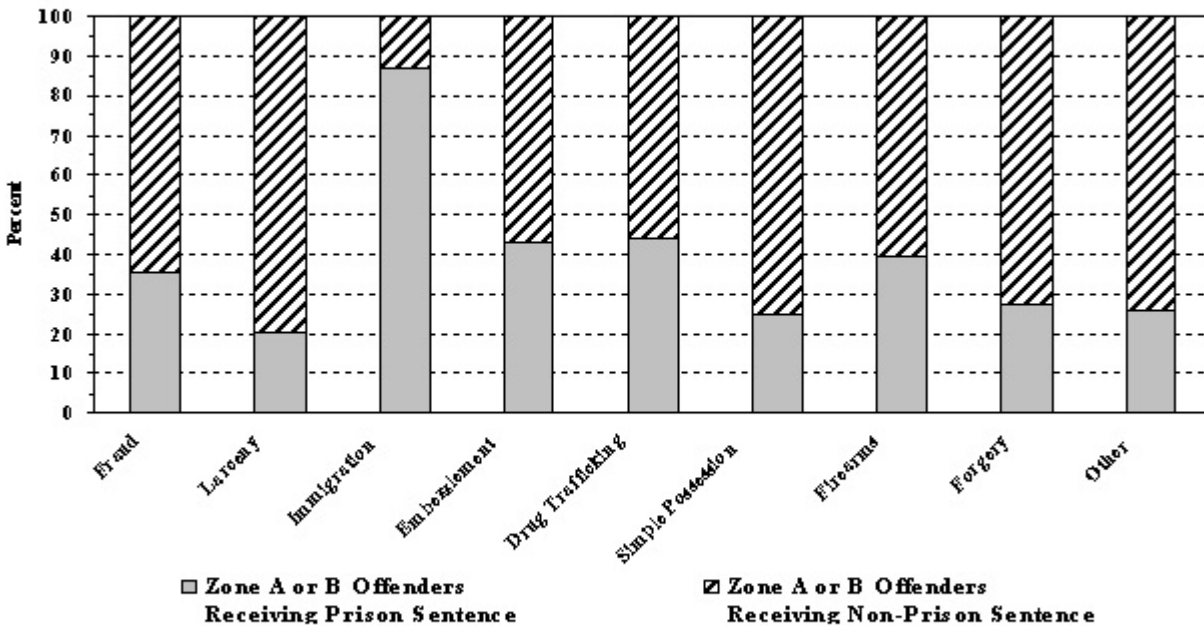
<sup>127</sup> Ashcroft Fast-Track Memo, *supra* note 46, Sept. 22, 2003.

<sup>128</sup> 28 U.S.C. § 994(b)(2).

<sup>129</sup> USSG §5C1.1.

<sup>130</sup> *See e.g.*, USSG §2C1.2, comment. n. 3.

**Figure F**  
**IMPRISONMENT RATES OF OFFENDERS ELIGIBLE FOR**  
**NON-PRISON SENTENCES IN SELECTED OFFENSE TYPES<sup>1</sup>**  
**Fiscal Year 2002**



<sup>1</sup> Of the 54,355 cases, the Commission received 40,416 guidelines applicable information for 52,348 cases. Of those 52,348 cases, 8,114 are ineligible for any prison sentence because the guidelines are less restrictive than the state law. Prisoners are also ineligible for prison sentences if they are on probation or parole. Full results are available at [www.uscourts.gov](http://www.uscourts.gov). The prison rates are based on the total population with any eligible offense. Some jurisdictions have a higher rate of imprisonment. Reported in R.

© SOURCE: U.S. Sentencing Commission, 2002 Data File, USSC FY02.

The Commission’s website contains this information for each district and each circuit. (See webpage for each jurisdiction at <http://www.uscourts.gov/JUDPACK/JP2002.htm>, Tbl. 6). Nationally, 45.9 percent of offenders for whom probation is an option receive imprisonment instead, but this rate varies significantly by district. In 2001, the rates in each district varied from a low of 9.3 percent to a high of 78.1 percent. (Districts that sentenced fewer than 30 probation-eligible offenders were excluded from these analyses because their rates can be dramatically affected by a small number of offenders.) The incarceration of probation-eligible fraud offenders, for example, varied from 17 percent to 38 percent between New Jersey and Pennsylvania, two contiguous districts.

Some of this regional variation can be accounted for by differences in the specific types of offenses and offenders sentenced in each region. A 1996 Commission report examined factors associated with judges’ use of sentencing options (USSC, 1996). Using a multiple regression model, it was found that criminal history, employment status, role in the offense, citizenship, and mode of conviction accounted for much, but not all, of the variation in the use of sentencing options. Judges are more willing to consider community placement for offenders who are employed, who plead

guilty, and who played a lesser role in their offense. The guidelines themselves discourage judges from using a probationary sentence for offenders with a criminal history category of III or above.<sup>131</sup>

**Placement within the guideline range.** For offenders who do not receive a sentencing option or a departure, judges must decide on a term of imprisonment within the prescribed guideline range. The *Sourcebook of Federal Sentencing Statistics* provides the percentage of offenders who are sentenced at the bottom, lower half, middle, upper half, or top of the guideline range for 32 different offense categories (USSC, 2002, Tbl. 29). Overall, in 2002, 59.8 percent of offenders were sentenced at the bottom of the ranges, 14.8 in the lower half, 8.9 percent at the middle, 6.4 percent in the upper half, and 10.1 percent at the top of the range. This distribution is slightly skewed to the bottom of the range compared to state guidelines systems on which data are available. For example, the Virginia Criminal Sentencing Commission reports that in fiscal year 2003, 65 percent of offenders were sentenced below the midpoint, 16 percent at the midpoint, and 19 percent above the midpoint (Virginia Criminal Sentencing Commission, 2003, p. 18).

Judges in different guidelines systems, and different judges in the federal system, vary in how they approach the guideline range. In some state guidelines systems, the presumptive sentence is in the middle of the guideline range. (See e.g., Kansas Sentencing Commission, 2002, p. 42;

***Different guidelines systems, and to a minor extent different federal judges, use the guideline range differently.***

Pennsylvania Commission on Sentencing, 2000.) Judges use the lower and upper ends of the range for mitigated and aggravated sentences that do not rise to the level of a departure. In other systems, including the federal system, the bottom of the range is most typically used.<sup>132</sup>

There appears to be general consensus among federal judges about how to approach the guideline range. In addition, plea agreements often specify where in the guideline range the parties agree the sentence should fall. Only a few judges use another part of the guideline range more frequently than the bottom of the range. Among the 911 federal judges who sentenced at least ten cases between 1999 and 2001, the bottom of the range was the most typical sentence for 880 of them. Twenty-four judges, however, most typically sentenced between the bottom and midpoint of the range, while two most typically sentenced between the midpoint and the top. Just one judge used the midpoint of the range most frequently, while four judges sentenced at the top of the range most frequently. It seems likely that judicial sentencing philosophy, rather than differences among the types of cases sentenced, account for these different approaches to the guideline range. While

---

<sup>131</sup> USSG §5C1.1, comment. n. 7.

<sup>132</sup> 18 U.S.C. § 3553(a) requires judges to “impose a sentence sufficient, but not greater than necessary, to comply with the purposes” of sentencing. When determining “the particular sentence to be imposed” the court shall consider “the kinds of sentence and the sentencing range established” by the guidelines.

generating a form of inter-judge disparity, these atypical practices are not widespread and fall within the range of discretion reserved for judges by the SRA.

## 5. *Departures for Exceptional Circumstances*

Commentators (Berman, 2000) and empirical analyses have suggested that departures for aggravating and mitigating circumstances articulated by the judge could be a continuing source of unwarranted sentencing disparity in the guidelines system (Gelacak et al., 1996; Farabee, 1998; Adams, 1998). Congress enacted the PROTECT Act to further limit the circumstances when downward departures are authorized. The Act became effective on April 30, 2003, and the Commission's guidelines amendments pursuant to it became effective October 27, 2003, after the data for this report were collected. Thus, the effects of the Act are not addressed in this report. Research on sentencing practices prior to the Act suggest that downward departures may well be contributing to inter-judge and regional disparities, but that the reasons for variations in downward departure rates have been poorly understood.

***The Commission's report on departures.*** As part of its Fifteen-Year Evaluation of the guidelines system, the Commission published a special report, *Downward Departures from the Federal Sentencing Guidelines* (USSC, 2003), which discussed the PROTECT Act and some of the concerns that motivated it. Empirical analyses presented in the report demonstrate that the number of cases sentenced within the guideline range decreased from 1991-2001. Until 1994 this decrease was attributable largely to an increase in departures for substantial assistance to the government, but after 1995 these departures declined slightly and other downward departures for mitigating circumstances began to increase.

Rates of departure vary by offense type, with the rate of departure for immigration offenses increasing substantially over the same years that the number of immigration offenses increased substantially. Rates of departure also vary

***Several possible factors have contributed to the increases and variations in downward departure rates.***

dramatically from district to district. While a clear majority of districts in 2001 had mitigating circumstance departure rates of less than ten percent, a quarter had rates of 10-20 percent and the remainder had rates even higher. Three districts had rates over 50 percent.

The Commission's departure report discussed several possible reasons for the increasing departure rate, as well as the concerns raised during debates preceding passage of the PROTECT Act. The Supreme Court's decision in *Koon v. United States*,<sup>133</sup> which held that an abuse of discretion standard applied to appellate review of departures, was discussed in Congress as a cause of increased departures. However, the Commission's report cited evidence suggesting that the impact of *Koon* was negligible. The report also showed that appeals of downward departures by the

---

<sup>133</sup> 518 U.S. 81 (1996).



government were rare both before and after the *Koon* decision, only ranging between 25 and 43 cases per year in five recent years. The data suggested that in 2001, the government initiated at least 40 percent of all downward departures for mitigating circumstances, often as part of an early disposition program or other guilty plea arrangement. The rate of downward departures for reasons other than substantial assistance that were *not* initiated by the government appeared to be approximately 10.9 percent in 2001.

The causes of variation in the rates of departure, and their potential effect on unwarranted sentencing disparity, is a complicated issue that cannot be resolved through simple examination of the reported rates. Problems with document submission (Mercer, 2003) and data accuracy (GAO, 2003) also complicate careful analysis. The strengthened reporting requirements put in place by the PROTECT Act and data collection improvements undertaken by the Commission are expected to address some of these concerns. When assessing the role of departures in creating unwarranted sentencing disparity during the first fifteen years of guidelines sentencing, however, caution is advisable and caveats are unavoidable.

***District factors influence the rate of departure more than circuit factors.*** For this report, a new analysis using a hierarchical model compared the amount of variation in departure rates associated with circuits with the amount associated with districts. (Details can be found in Technical Appendix D.) The case law governing various grounds for departure varies somewhat from circuit to circuit (Nagel & Galacek 1996; Lee 1997; Johnson 1998) and different circuits have been recognized as having different climates or “personas” regarding their amenability to departure (Demlietner, 1994). But results from the hierarchical analysis suggest that differences in circuit case law or climate, while exerting some significant influence over departure rates, are less important than differences among the districts. Only about one quarter of the variation in downward departure rates is attributable to the circuits, while three quarters is attributable to districts.

***The GAO’s exploration of regional variations.*** Recent research by the GAO investigated how much regional variation in departure rates can be accounted for by differences in offense and offender characteristics (GAO, 2003). The GAO found “major variation among certain judicial circuits and districts” (*id.* at 3-4) in the likelihood of departure in drug trafficking cases, even after controlling for a variety of offense and offender characteristics, including the type of drug involved, the presence of a weapon, the severity of the offense, whether the defendant pled guilty, and whether the offense was eligible for a mandatory minimum penalty or the safety valve. Differences among circuits and districts in the likelihood of departure were usually reduced after controlling for these characteristics, indicating that some of the regional variation is due to the different types of cases and offenders in the various regions. Significant regional variation remained, however. For example, downward departure remained 6.78 times more likely in the Ninth Circuit than in the Eighth, even after controlling for offense and offender characteristics.

Because “empirical data on all factors that could influence sentencing were not available” the GAO noted that the remaining differences “may not, in and of themselves, indicate unwarranted sentencing departures or misapplication of the guidelines.” As the Commission noted in its response

to the draft GAO report, several factors that might help account for regional variation in departure rates were not included in the GAO's analysis. Most noteworthy, the GAO did not take into account the existence in several districts of formal, government-created "fast track" programs that offer departures as part of a plea agreement as an incentive for quick waiver of certain defendant rights.

***USSC replication and extension of the GAO analysis.*** To estimate the impact of "fast track" on regional variation in departure rates, the GAO's analysis was replicated including a variable indicating whether a "fast track" program involving departures was in place in a particular district. A letter from the Department of Justice to the Commission on August 12, 2003 was used to identify those districts having such programs during the years of our analysis. The results show that regional variation in downward departure rates is greatly reduced when the presence of "fast track" programs in some districts is taken into account. In particular, the increased odds of departure in the Fifth, Ninth, and Tenth Circuits (when compared to the Eighth Circuit, the same circuit used for comparison in the GAO report) are reduced by more than two-thirds. The variation in departure rates accounted for by "fast track" programs is much greater than the variation accounted for by all of the offense and offender characteristics included in the GAO's analysis combined.

However, while the highest departure rates are clearly due to the presence in some districts of "fast track" programs, significant variation remains after controlling for these programs. The odds of receiving a downward departure for mitigating circumstances remain over three times higher in the Ninth Circuit than in the Eighth, almost three times higher in the Second, and two times higher in the DC circuit. In the Fifth Circuit, on the other hand, the odds of departure are just 17 percent that of the Eighth Circuit. As noted above, however, district practices are more important than circuit factors in determining the departure rate. Within the Ninth Circuit, the odds of departure vary from Montana, with odds less than half those found in the District of Minnesota (again, the comparison districts used by the GAO) to almost twice the odds in the Northern District of California. (Arizona and Southern California were excluded due to the unusual case types and workload found in these border districts.) Similarly, while most of the districts in the Fourth Circuit have lower odds of downward departure than Minnesota, the District of Maryland has slightly higher odds. Clearly, practices particular to each district have a substantial impact on the departure rates in those districts.

***Continuing debate over which regional variations are warranted.*** Identifying the reasons for regional variation in departure rates will not settle the policy question of whether the variation is warranted or unwarranted. Numerous commentators have argued that some regional variation is warranted by local conditions. In addition to different workload pressures (Braniff, 1993) commentators have suggested that different crimes generate different levels of public concern in different regions, which should be reflected in the sentences imposed (Broderick, 1993; Ragee 1993, Sifton, 1997). It has also been argued that departure can be used to ameliorate the unwarranted disparity that can arise when some offenders are prosecuted in federal court while others are prosecuted in state court where sentences are more lenient (O'Hear, 2002). Regional variation in sentencing has been, and will likely continue to be, a lively area of research and debate.