United States General Accounting Office

GAO

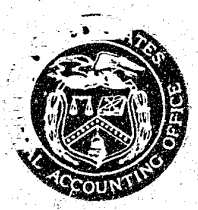
[redacted]

[redacted]

[redacted]

[redacted]

[redacted]



152288

GAO

United States
General Accounting Office
Washington, D.C. 20548

General Government Division

B-249290

August 14, 1992

The Honorable Joseph Biden, Chairman
Committee on the Judiciary
United States Senate

The Honorable Strom Thurmond
Ranking Minority Member
Committee on the Judiciary
United States Senate

The Honorable Jack Brooks, Chairman
Committee on the Judiciary
House of Representatives

The Honorable Hamilton Fish, Jr.
Ranking Minority Member
Committee on the Judiciary
House of Representatives

Throughout most of this century, an important goal of the federal government's sentencing system was the rehabilitation of the convicted offender. Despite some successes, a growing number of studies in the 1970s and 1980s showed that rehabilitation was not working. Rates of recidivism among released inmates remained essentially unchanged.

Under this system, Congress established broad sentencing ranges—1 to 10 years in prison, for example—for specific crimes. Judges could choose a sentence anywhere within this range and were not required to articulate the bases for their sentences. Thus, judges enjoyed wide discretion to sentence in accordance with their own theories of justice and rehabilitation. Similarly, the Parole Commission exercised broad discretion in determining the amount of time a prisoner actually served; a prisoner could be released after serving as little as one-third of his or her sentence.

A number of studies showed considerable sentencing disparity under this system—that is, wide variations in the sentences imposed on similar offenders for similar criminal behavior. Some studies indicated that, at least for some offenses, the identity of the sentencing judge was a better predictor of incarceration than the defendant and the crime. Other studies showed that factors such as race, gender, economic status, marital status, employment status, and education affected sentences and created wide

disparity in sentences imposed. Moreover, given the Parole Commission's discretion to determine release dates, the actual sentence the judge imposed was not necessarily a reliable indicator of the actual prison term a convicted offender might serve. A recent study summarized the situation as follows:

"The view that sentencing was more a product of a lottery than a rational punishment scheme undermined public confidence in, and respect for, the criminal justice system. Additionally, the public became increasingly dissatisfied with the perceived leniency of some criminal sentences and with the dishonesty of judicial sentencing practices that exaggerated the length of the ultimate prison term."[1]

Faced with such evidence, and a perception that crime was growing, Congress enacted the Sentencing Reform Act of 1984.[2] The act made fundamental changes to federal sentencing policy in an attempt to bring more certainty to the sentences received by persons convicted of violating federal laws and to reduce sentencing disparity. The act created the United States Sentencing Commission and required the Commission to develop a system of sentencing guidelines. The act eliminated parole, and Congress instructed that sentencing under the guidelines should be neutral as to the race, gender, creed, national origin, and socio-economic profile of offenders, while taking into account the nature and circumstances of the offense and the criminal history and characteristics of the offender. The act also identified rehabilitation as an inappropriate goal of imprisonment, but retained it as a general goal of lesser sanctions, such as probation.

The act instructed the Commission and GAO to evaluate the impact of the guidelines. Specifically, it required us to evaluate the impact of the guidelines and compare the operation of the new system with the old.

The Commission's Report

In January 1992, the Commission issued its report evaluating the federal sentencing guidelines, as required by the act.[3] The Commission reported that implementation of the guidelines was moving steadily forward but remained in transition. For the limited number of preguidelines and guidelines cases of bank robbery, cocaine distribution, heroin distribution,

[1]Theresa Walker Karle and Thomas Sager, "Are the Federal Sentencing Guidelines Meeting Congressional Goals? An Empirical and Case Law Analysis," Emory Law Journal, Vol. 40, No. 2 (Spring 1991), p. 396.

[2]Public Law 98-473.

[3]The Federal Sentencing Guidelines: A Report on the Operation of the Guidelines System and Short-Term Impacts on Disparity in Sentencing, Use of Incarceration, and Prosecutorial Discretion and Plea Bargaining, issued in January 1992.

and bank embezzlement cases it was able to compare, the Commission found that the amount of sentencing disparity had been reduced.

In a separate analysis of guidelines cases only, the Commission evaluated whether disparity resulting from offender characteristics that should normally be irrelevant to determining a sentence was present under the guidelines. The Commission found cases where the sentence imposed under the guidelines was related to the offender's race, gender, and employment status, but not to age, marital status, or education.

Finally, the Commission analyzed how plea and charging agreements between prosecutors and defense attorneys affected sentencing under the guidelines but was unable to determine the effect of these decisions on sentencing disparity. The Commission found that prosecutorial charge reductions and other bargaining appeared to have had an impact on the sentencing process in 17 percent of the cases it examined.

Results in Brief

At this time, neither we nor the Commission can definitively answer the central question posed by Congress regarding how effective the sentencing guidelines have been in reducing sentencing disparity. Where there were sufficient data for analysis of comparable preguidelines and guidelines cases, our findings concurred with those of the Commission in that we also concluded that the amount of sentencing disparity had been reduced. However, data limitations precluded generalizing either our or the Commission's results from the specific cases analyzed to all sentencing under the guidelines. Similarly, a lack of empirical data on workload prevented a reliable quantification of the impact of the guidelines on the operations of the criminal justice system.

Although the available data were not sufficient to permit overall generalization, the results of our analysis of whether unwarranted disparity continued under the guidelines were consistent with the Commission's. Unwarranted disparity is disparity related to offender characteristics which should normally be irrelevant in determining a sentence. Using the Commission's data and a different approach, our analyses showed that of the six variables tested by the Commission, only education appeared to have had no relationship to the sentences imposed. Though no clear or consistent pattern emerged, both our analysis and the Commission's showed that the guidelines had not removed all unwarranted disparity within a sentencing range.

An area which may lead to unwarranted sentencing disparity under the guidelines involves presentencing decisions by prosecutors, defense attorneys, and others. These decisions include the use of discretion in the choice of charges filed and plea bargaining. The exercise of discretion at each presentencing stage can potentially affect the final sentence imposed. For example, we found that for heroin distribution, both whites and Hispanics were more likely than blacks to have counts reduced or dismissed prior to going to trial and consequently may have received shorter sentences. These results raise the issue of whether, and if so to what extent, the presentencing decisions result in unwarranted disparity in sentencing. However, as data are not routinely collected on such decisions and the data available were limited, neither we nor the Commission were able to address this issue.

As required by the act, we also examined the impact of the guidelines on the operations of the criminal justice system. Based on the limited statistical data available and our interviews with criminal justice system personnel in four districts, it appeared that the guidelines increased to some degree the workload of judges, prosecutors, probation officers, and defense attorneys. Interviewees said that the guidelines introduced new elements to their jobs and caused some existing elements to take longer. While the available data precluded us from quantifying the amounts of change, the greatest increase appeared to be for probation officers. Future evaluation efforts should include provisions for developing the data and analyses necessary to address these impacts and their magnitude.

Delays caused by the decision to phase in implementation of the guidelines by making them applicable only to offenses committed on or after November 1, 1987, and subsequent litigation has limited the amount of guidelines experience and, consequently, the data available for review. Also, limited comparability between the data available on preguidelines offenders and guidelines offenders restricted the number of cases suitable for comparison.

As a practical matter, it is not possible to rectify the shortcomings in preguidelines data and develop a more meaningful baseline for comparing sentencing outcomes before and after the guidelines. Consequently, future analytical resources could best be used to evaluate current disparity under the guidelines rather than continuing to focus on comparisons between the preguidelines and current systems. Future efforts should also be designed to ensure that the guidelines' impact on system operations can be better determined.

Background

Congress fundamentally altered national sentencing goals and practices with the passage of the Sentencing Reform Act of 1984. The act created the Commission as an independent entity in the judicial branch and authorized it to develop and implement guidelines for federal judges to use in sentencing offenders. Before the guidelines were established, federal judges had broad sentencing discretion for federal offenses. A bank robber, for example, could receive a sentence ranging from probation to 20 years in prison. The judge could choose any sentence within this range. However, after serving one-third of the sentence, the offender could be eligible for parole, with the Parole Commission determining the actual date of release. As a result, the time actually served could vary considerably. Because rehabilitation was one of the goals of prison sentences, inmate behavior and self improvement efforts were key determinants of when, within the range of time established by the judge's sentence, the Parole Commission would approve release.

Wide Sentencing Discretion Generated Concern

The 1970s was a period of growing public disenchantment with wide sentencing discretion and its results in encouraging the rehabilitation of offenders. During this time, support for sentencing reforms, including limits on sentencing discretion, escalated. Federal legislation calling for the establishment of sentencing guidelines was first introduced in 1975 and was introduced in each successive Congress until the act was passed. According to the legislative history and other literature on federal sentencing reform, three major factors led to reform:

- public frustration over increases in crime rates and a perception that the criminal justice system was too soft on crime;
- studies reporting that the rehabilitative model of sentencing was not working; and
- a belief that indeterminate sentencing caused unwarranted disparity in sentences given by different judges and courts.

The Senate Judiciary Committee Report, in citing reasons for enacting sweeping changes to the system for sentencing federal offenders, stated:

"Almost everyone involved in the criminal justice system now doubts that rehabilitation can be induced reliably in a prison setting, and it is now quite certain that no one can really detect whether or when a prisoner is rehabilitated. Since the sentencing laws have not been revised to take this into account, each judge is left to apply his own notions of the purposes of sentencing. As a result, every day federal judges mete out an unjustifiably wide range of sentences to offenders with similar histories, convicted of similar crimes, committed under

similar circumstances. . . . These disparities, whether they occur at the time of the initial sentencing or at the parole stage, can be traced directly to the unfettered discretion the law confers on those judges and parole authorities responsible for imposing and implementing the sentence."[4]

In the act, Congress delegated to the Commission the responsibility for developing mandatory guidelines to be used by judges in determining appropriate sentences. The act stipulated that under the guidelines offenders would serve the full prison sentence imposed by the court less any credit of up to 54 days a year for satisfactory behavior. The act also eliminated provisions for parole and set a schedule for phasing the Parole Commission out of existence.

The act directed the Commission to ensure that sentencing under the guidelines was neutral in terms of an offender's race, gender, national origin, creed, and socioeconomic status and to take care that the guidelines reflected the general inappropriateness of considering an offender's education and vocational skills, employment record, family ties and responsibilities, and community ties in sentencing.

The first set of guidelines went into effect on November 1, 1987, for offenders sentenced for offenses occurring on or after that date. The Commission has described the first 15 months after the guidelines became effective as a period of "extraordinary litigation." Before the Supreme Court upheld the guidelines' constitutionality in January 1989,[5] more than 200 district court judges had invalidated them in whole or in part, while about 120 district judges had upheld their constitutionality. The Supreme Court decision cleared the way for nationwide application. The Commission assumes that virtually all sentencing is now taking place under the guidelines.

Guidelines Use Sentencing Grid

An offender's sentence under the guidelines is to be determined using a sentencing table or grid (see fig. 1). The grid reflects the act's guidance to the Commission that the guidelines consider both an offender's criminal history and the offense for which the offender was being sentenced. Consequently, the left side of the grid consists of 43 offense levels with the least serious crimes falling within the lower offense levels and the most serious crimes at the high end. For example, failing to register for military

[4]Report on S.1762, The Comprehensive Crime Control Act of 1983, Committee on the Judiciary, United States Senate (Sept. 14, 1983), p. 38.

[5]Mistretta v. United States, 488 U.S. 361 (1989).

service is an offense level 6, while transmitting top secret national defense information to a foreign government is an offense level 42.

Figure 1: Sentencing Guidelines Table (Sentencing Ranges in Months)

Offense Level	Criminal History Category					
	I	II	III	IV	V	VI
1	0-6	0-6	0-6	0-6	0-6	0-6
2	0-6	0-6	0-6	0-6	0-6	1-7
3	0-6	0-6	0-6	0-6	2-8	3-9
4	0-6	0-6	0-6	2-8	4-10	6-12
5	0-6	0-6	1-7	4-10	6-12	9-15
6	0-6	1-7	2-8	6-12	9-15	12-18
7	1-7	2-8	4-10	8-14	12-18	15-21
8	2-8	4-10	6-12	10-16	15-21	18-24
9	4-10	6-12	8-14	12-18	18-24	21-27
10	6-12	8-14	10-16	15-21	21-27	24-30
11	8-14	10-16	12-18	18-24	24-30	27-33
12	10-16	12-18	15-21	21-27	27-33	30-37
13	12-18	15-21	18-24	24-30	30-37	33-41
14	15-21	18-24	21-27	27-33	33-41	37-46
15	18-24	21-27	24-30	30-37	37-46	41-51
16	21-27	24-30	27-33	33-41	41-51	46-57
17	24-30	27-33	30-37	37-46	46-57	51-63
18	27-33	30-37	33-41	41-51	51-63	57-71
19	30-37	33-41	37-46	46-57	57-71	63-78
20	33-41	37-46	41-51	51-63	63-78	70-87
21	37-46	41-51	46-57	57-71	70-87	77-96
22	41-51	46-57	51-63	63-78	77-96	84-105
23	46-57	51-63	57-71	70-87	84-105	92-115
24	51-63	57-71	63-78	77-96	92-115	100-125
25	57-71	63-78	70-87	84-105	100-125	110-137
26	63-78	70-87	78-97	92-115	110-137	120-150
27	70-87	78-97	87-108	100-125	120-150	130-162
28	78-97	87-108	97-121	110-137	130-162	140-175
29	87-108	97-121	108-135	121-151	140-175	151-188
30	97-121	108-135	121-151	135-168	151-188	168-210
31	108-135	121-151	135-168	151-188	168-210	188-235
32	121-151	135-168	151-188	168-210	188-235	210-262
33	135-168	151-188	168-210	188-235	210-262	235-293
34	151-188	168-210	188-235	210-262	235-293	262-327
35	168-210	188-235	210-262	235-293	262-327	292-365
36	188-235	210-262	235-293	262-327	292-365	324-405
37	210-262	235-293	262-327	292-365	324-405	360-life
38	235-293	262-327	292-365	324-405	360-life	360-life
39	262-327	292-365	324-405	360-life	360-life	360-life
40	292-365	324-405	360-life	360-life	360-life	360-life
41	324-405	360-life	360-life	360-life	360-life	360-life
42	360-life	360-life	360-life	360-life	360-life	360-life
43	life	life	life	life	life	life

The top of the grid consists of six criminal history categories, with category I being the least severe and category VI being most severe. Placement in one of these categories is to be determined by such factors as number of prior convictions. For example, if an offender is sentenced for crimes of violence or drug offenses and has had two or more prior felony convictions of this nature, the offender is to be considered a career offender and placed in category VI, the highest criminal history category.

The intersection of an offender's final offense level (after all adjustments) and criminal history category determines his or her placement within a "cell" or sentencing category on the grid, and, thus, the potential sentence.[6] For example, an offender convicted of robbery (base offense level 20), with no adjustments, who has a criminal history score that equates to category II could receive a prison sentence ranging from 37 to 46 months. The statute directed that the maximum sentence a defendant could serve for a particular offense should generally not exceed the minimum by 25 percent or 6 months, whichever was greater. A judge is to determine the exact number of months within the range to which the offender is sentenced.

By directing judges to sentence according to this matrix, the guidelines have substantially narrowed judges' sentencing discretion. However, the statute permitted judges to depart from the guidelines when they found special circumstances that the Commission did not adequately consider. But the act also purposefully limited the use of judicial departures to these special circumstances.

Objectives, Scope, and Methodology

The act required the Commission and us to evaluate and report on the impact of the guidelines. It also required us specifically to compare the operation of the old system with the new. Within this broad statutory mandate, our objectives were to determine the following:

- whether or not the guidelines reduced the variation in sentences imposed and time served by groups of offenders who committed similar crimes and who had similar criminal histories, and whether the average time served for such similar groups of offenders increased or decreased;

[6]Adjustments are to be made to the initial, or "base," offense level to account for aggravating or mitigating circumstances in the crime and other factors. For example, a robbery offense with a base offense level of 20 can be increased for such aggravating factors as injury to a victim, use of a firearm, or losses of more than $10,000. The offense level can be decreased if, for example, the offender accepts responsibility for the crime and/or had a minimal or minor role in the activity.

- how the guidelines sentences related to offense characteristics, such as severity of the offense, and to offender characteristics, such as race;
- the perceptions of judges, prosecutors, defense attorneys, and probation officers regarding the benefits, problems, and long-range effects of the sentencing guidelines;
- the impact of the guidelines on the operations of the federal criminal justice system. Our work focused on how the guidelines affected the workload and budget and case processing times of the courts and investigative agencies.

We did our work between March 1991 and May 1992 in accordance with generally accepted government auditing standards. All of the sentencing data used in our analyses were provided by the Commission. Discussions of the scope of our work and the specific methodologies used to respond to the above objectives are provided in appendixes I through V.

Overall Effectiveness of Guidelines in Reducing Preguidelines Disparity Could Not Be Determined

One of the primary objectives of the sentencing guidelines was to reduce sentencing disparity, so that offenders with similar criminal histories who are convicted of similarly serious criminal offenses would receive similar sentences. Congress was particularly interested in reducing or eliminating disparity caused by demographic factors such as an offender's race, gender, and education. However, limitations and inconsistencies in the data available for preguidelines and guidelines offenders made it impossible to determine how effective the sentencing guidelines have been in reducing overall sentencing disparity.

Significant differences in much of the offender data available made it difficult to reliably match and compare groups of preguidelines and guidelines offenders. Preguidelines offender data focused on personal information, such as socioeconomic status and family and community ties, that was supposed to be irrelevant under the guidelines in all or most cases. Conversely, most of the detailed data available on guidelines offenders, such as role in the offense, were not available for preguidelines cases.

Delays in the nationwide implementation of the guidelines restricted both the Commission's evaluation and ours to about 2-1/2 years of limited experience under the guidelines rather than the 4 years Congress envisioned when setting the statutory dates for the evaluations. Two factors contributed significantly to the limited amount of data available for evaluation: the constitutional challenge ultimately resolved by the

Supreme Court and the decision to apply the guidelines only to offenses committed on or after November 1, 1987. According to the Commission's data, by June 1991, about 74 percent of federal criminal offenders were sentenced under the guidelines.

As a result of these events, the number of guidelines cases available to the Commission and us for analyses was small. Consequently, neither our conclusions nor those of the Commission could be generalized to the guidelines as a whole.

Preguidelines Disparity for Selected Groups Was Reduced

While data limitations made it impossible to determine if the guidelines had been effective in reducing overall sentencing disparity, it was possible to make a determination for selected groups.

On the basis of a comparison of a sample of preguidelines and guidelines cases for bank robbery, cocaine distribution, heroin distribution, and bank embezzlement, the Commission concluded that the overall disparity in sentences of similar offenders in these cases had declined under the guidelines.[7] We applied a different approach to the Commission's data on cocaine and heroin distribution offenders and reached the same conclusion.

To obtain a more general comparison of sentencing disparity before and after the guidelines, we analyzed sentencing dispersion using larger samples of offenders sentenced under the preguidelines system and those sentenced under the guidelines system.[8] We analyzed 68 groups of offenders.[9] Of these 68 groups, 57 groups had less dispersion under the guidelines and 11 had higher. (See app. I for additional details.)

We recognized that the dispersion increases we found might have been the result of factors other than the guidelines, such as mandatory minimum[10]

[7]The Commission's sample contained 479 of the 25,940 offenders sentenced under the guidelines in fiscal year 1990. The small number resulted from the Commission's effort to match several specific offense and offender characteristics for offenders in each of the four crime categories examined.

[8]Our analysis was able to include a larger sample of offenders because we used a more limited set of criteria to match preguidelines and guidelines offenders than the Commission did. In addition, our analysis focused on sentence dispersion as an indicator for disparity. Dispersion is defined as the spread of sentences around the average sentence. See appendix I for additional details.

[9]A group may be a single cell on the sentencing grid, or a group of cells with the same sentencing range. See appendix I for additional explanation.

[10]Mandatory minimum sentence refers to a statutory provision requiring the imposition of at least a specified minimum sentence when the statutorily specified criteria have been met.

and consecutive sentences,[11] which have increased sentence length. For example, an additional 5-year sentence must be imposed when a defendant is convicted of using a firearm during a "crime of violence" or "drug trafficking crime." When we controlled for the impact of significant increases in sentence lengths for those sentenced under the guidelines, dispersion in all cells decreased by varying degrees.

A detailed discussion of how we selected our samples for the preguidelines and guidelines groups for comparison and other methodological issues is provided in appendix I.

Unwarranted Disparity Continued

We and the Commission used the limited data available at the time of our study to conduct additional disparity analyses on samples of offenders sentenced under the guidelines between April 1, 1990, and September 30, 1990.[12] These analyses evaluated the relationship between an offender's race, gender, marital status, employment status, education, and age and the sentence imposed. As previously discussed, under the guidelines, none of these variables should normally be relevant in determining an offender's sentence. The Commission found cases where the sentence imposed under the guidelines was related to the offender's race, gender, and employment status, but not to age, marital status, or education.

Using the same data, we extended the Commission's analysis to control separately for offense severity level, criminal history category, offense type, and mode of disposition (whether by plea or by trial). Our analysis confirmed the existence of disparity and showed that, of the variables listed above, only education had no statistically significant relationship to the sentence imposed.

A different Commission analysis and other available evidence suggest that the perceived effects of these factors may change when the sentences imposed for specific offenses are analyzed. Nevertheless, both our analyses and those of the Commission suggest that under the guidelines

[11]A mandatory consecutive sentence refers to a statutory "enhancement" or "add-on" sentence that must be served in addition to any other sentence imposed for the offense for which the consecutive sentence is statutorily required.

[12]Sentencing disparity exists when offenders with similar criminal records found guilty of similar criminal conduct receive dissimilar sentences. Unwarranted sentencing disparity exists when those dissimilar sentences result in some way from offender characteristics—race, gender, creed, national origin, or socioeconomic status—that, under the guidelines, are to be neutral in their effect on sentences imposed. Unwarranted disparity may also exist when dissimilar sentences result from offender characteristics that should not generally be relevant in determining the sentence imposed, such as education and family or community ties.

not all unwarranted disparity in sentencing has been eliminated. Further, our analyses indicated that for some variables, such as race, the effects were not consistent. For example, in some cases sentencing outcomes appeared to disfavor blacks; in others, they appeared to favor blacks. Additional research as more data become available is necessary to determine under what conditions disparity is occurring and whether the amount of disparity is increasing, decreasing, or remaining stable. Details on these analyses can be found in appendix II.[13]

While both our analysis and that of the Commission found that the sentence imposed may be related to variables that should not normally be relevant, we disagree on how these results should be interpreted. The basis of our disagreement is in what constitutes similar sentences under the guidelines.

The Commission maintained that as long as the sentences imposed are within the guidelines range, they are by definition similar. Consequently, according to the Commission, unwarranted disparity cannot exist if the sentences imposed fall within the guidelines range. We disagree with the Commission's position that disparity can only exist when the sentence imposed is outside the guidelines range.

We have taken a broader view than the Commission as to what constitutes dissimilar sentences and, accordingly, unwarranted disparity. Sentencing variations within a given guidelines range can result in quite different sentencing outcomes. For example, at the lower offense levels one can be sentenced to prison for 6 months or to probation. At the highest offense levels the difference between sentences can be 7 or more years. The Commission's rationale would obviate any finding of unwarranted disparity despite the different sentences permitted within a given guidelines range.

Both we and the Commission found that some offender characteristics that are supposed to be neutral, such as gender, were related to the within range sentences imposed. We believe that such statistically significant differences in the within range sentences should be considered an indication of unwarranted disparity.

[13]An analysis of racial differences in sentencing under the guidelines for bank robbers is contained in appendix III.

Presentencing Decisions Also Affected by the Guidelines

A number of judges and others argued that focusing solely on what happens after conviction when looking at the impact of the guidelines misses potential sources of sentencing disparity that occur earlier in the process. The act anticipated the possibility of such disparity and directed the Commission to evaluate "the impact of the sentencing guidelines on prosecutorial discretion [and] plea bargaining."

A variety of discretionary decisions by prosecutors, defense attorneys, judges, and probation officers can introduce unwarranted disparity into sentencing outcomes. These decisions include the choice of criminal charges filed, plea bargaining policy, and the extent of the presentence investigation. The guidelines increased the importance of these decisions to ultimate sentencing outcomes and their potential impact on unwarranted disparity.

For example, for bank robbery and larceny, we found that Hispanics were more likely than either whites or blacks to have counts reduced or dismissed and, consequently, to receive a lower sentence. For heroin distribution both whites and Hispanics were more likely than blacks to have counts reduced or dismissed; there were no significant differences for embezzlement.

These results raised the issue of whether different guilty plea rates between blacks and whites might have been associated with the sentencing gap between blacks and whites. Available data showed that persons convicted by plea tended to receive shorter sentences than persons who were convicted at trial and that blacks were less likely to be convicted by plea than whites. The limited data on why blacks were convicted by guilty plea less often made it difficult to determine the extent to which disparity in presentencing decisions, such as plea negotiations, affected the ultimate disparity in sentences imposed. Because of the absence of data, it's not possible to determine the differential rates of conviction by plea and the extent to which any racial disparity in sentencing outcomes might be due to that factor.

The Commission's evaluation study acknowledged the potential importance of charging and plea practices under the guidelines:

"Under the current determinate guidelines system, the potential exists for closer association between charging and plea practices and sentence severity than in the former indeterminate system. This is particularly true for offenses such as bank robbery that are treated by the guidelines as separate and distinct instances of criminal conduct. Although

the relevant conduct guideline takes into account criminal behavior beyond the elements of the offense of conviction for all offenses, plea negotiation practices and, in the case of separate and distinct offenses, charging practices by the prosecutor have the potential to influence the applicable guideline range."[14]

Our interviews, the Commission's own studies, and other studies indicated that there was considerable variation in how preconviction decisions were made under the guidelines. For example, in the districts we and the Commission visited, some judges said they would accept plea agreements with certain provisions that would affect the sentencing outcome, while others said they would not. Interviewees in our study and others questioned whether prosecutors were consistent in reflecting the most serious readily provable charges in their plea negotiations.

At the time of our study, little data existed on presentencing decisions, but existing studies using that data suggested that these practices do affect sentencing outcome and disparity. However, the Commission's report succinctly stated the problem involved in evaluating this issue:

"Primarily, much of the plea negotiation process involves 'behind the scenes' discussions between prosecutors and defense attorneys that generally are not memorialized. Evidentiary problems and defendant cooperation may affect the outcome of a plea negotiation (i.e., the sentence), but often there is little record of how this outcome evolved. Without data on specific decision points in this plea process, quantitative analysis cannot be performed."[15]

These same limitations also affect case file analyses of the type the Commission undertook for its evaluation. However, given the existing data limitations, such analyses, supplemented by interviews with participants, were the only ones possible and do offer useful insights into the actual and potential impact of presentencing negotiations on sentencing outcomes and disparity. The Commission's evaluation, using case file analyses, found that prosecutorial charge reductions and other bargaining appear to have had an impact on the sentencing process in about 17 percent of the 1,212 cases it examined and in 26 percent of the 428 drug trafficking cases it examined (this was also the largest category examined, more than 3 times as large as the next largest category, fraud). These findings are consistent with those of another recent study the preliminary results of which were reported at a March 1992 sentencing institute for the Second and Eighth

[14]Sentencing Guidelines, Executive Summary, p. 24.

[15]Sentencing Guidelines, Executive Summary, p. 65.

Judicial Circuits.[16] This study found that prosecutors, through charging and plea decisions, affected sentences in 20 to 35 percent of the cases reviewed. According to the study, prosecutors did so largely to "circumvent" the guidelines for sympathetic offenders, especially to avoid mandatory minimum sentences.

The Commission has promulgated policy statements for judicial review of plea agreements. However, the Commission's interviews with judges revealed that nearly three-fourths of the judges interviewed said they accepted plea agreements when presented, rather than waiting for the presentence report, and about two-thirds said they felt bound to honor the terms of plea agreements. This suggests that there is potential for the charging and plea negotiation process to have a significant impact on sentences imposed under the guidelines.

We agree with the Commission's conclusion, from its interviews, that the "experimentation period" of the guidelines was not yet complete. With regard to the impact of charge and plea practices, the Commission concluded:

"On the basis of these data, it is difficult to determine to what extent reductions occur due to plea agreements that, for example, involve dismissal of charges, or occur due to a combination of prosecutorial behavior circumventing the guidelines and judicial acquiescence in the face of such agreements."[17]

In our view, such factors, along with other aspects of presentencing practices, need to be analyzed if the impact of the guidelines in reducing unwarranted disparity in sentencing is to be fully addressed.

Guidelines Debate Has Continued

From their inception the sentencing guidelines have been the focus of widespread debate within the criminal justice system. Both the Commission and we interviewed a cross section of people involved in the criminal justice system, including judges, prosecutors, defense attorneys, and probation officers, to determine their current views. In both our and the Commission's interviews, supporters of the guidelines believed that the new system made sentencing more consistent and predictable.

[16]Ilene H. Nagle, and Stephen J. Schulhofer, "A Tale of Three Cities: An Empirical Study of Charging and Bargaining Practices under the Federal Sentencing Guidelines," Southern California Law Review (forthcoming—Nov. 1992).

[17]Sentencing Guidelines, Executive Summary, p. 81.

Critics, on the other hand, argued that the guidelines are too harsh and rigid and have gone too far in reducing judicial discretion by, for example, not allowing consideration in sentencing of relevant personal characteristics of offenders, such as age or family situation. Some judges and defense attorneys, in our interviews, also said that the guidelines largely ignore the hidden disparity resulting from the largely undocumented discretion exercised by prosecutors in charging decisions and plea bargaining.

We asked 10 district court judges, 12 federal prosecutors, 19 public and private defense attorneys, and 12 probation officers in 4 court districts what benefits, problems, and long-range effects they believe have resulted and will result from the guidelines.[18] Prosecutors generally believed that the guidelines improved the sentencing system, while most judges, defense attorneys, and probation officers on balance did not believe the guidelines were an improvement over the prior system.

The most frequently mentioned benefits were less disparity or more uniformity in sentencing and more certainty of the prison time offenders would actually serve. The most frequently cited problems were that the guidelines were too harsh (primarily for drug offenders and first-time offenders), too inflexible, and "dehumanizing" because they (1) reduced multifaceted human behaviors to a set of numbers; (2) were too complex; and (3) by limiting judges' discretion, gave prosecutors too much control over sentencing outcomes based on how they charged offenders and what pleas they accepted. Most anticipated long-term effects focused on the impact of the guidelines on prisons, including increased costs and populations, discipline, and the readjustment of offenders back to society after long prison terms. Tables 1 through 3 show the number of times common benefits, problems, and anticipated long-term effects were cited by each practitioner group.

Two policy decisions were frequently questioned by interviewees. A majority of the personnel we interviewed thought that judges should have more ability to depart, or sentence outside of the guidelines, than the act permits. The act provides for departures only in those instances where judges find special circumstances not considered by the Commission in developing the guidelines. However, many of our interviewees believed that factors the act generally found inappropriate, such as family responsibilities, should be considered.

[18]For a discussion on how the interviewees were selected see appendix IV.

Interviewees had mixed views of the Commission's policy that only the prosecutor should be allowed to submit a motion to have the normal guidelines sentence reduced because the offender provided substantial assistance in the investigation or prosecution of another person who has committed a crime. All 12 of the prosecutors we interviewed thought the Commission should allow this authority to remain exclusively theirs because they are in a unique position to determine the value of the assistance provided.

Conversely, 6 of the 10 judges we interviewed and all 19 defense attorneys believed this authority should be extended to judges and/or defense attorneys to reduce opportunities for abuse by prosecutors. Many of the opinions expressed in our interviews are consistent with what has been reported in other studies of the guidelines' impact and at a judicial sentencing institute for the Second and Eighth Circuits in March 1992.

While our interviewees did not believe that unwarranted sentencing disparity occurred often, they were able to provide some examples of situations where they thought it had and could occur. These included (1) drug cases where the guidelines reflect statutory provisions which require similar sentences for offenders with different levels of culpability, (2) the exercise of prosecutorial discretion in negotiating charges and plea agreements, and (3) the judicial use of departures and awards of acceptance of responsibility.

Additional information, including examples of when interviewees thought disparity had occurred and their views on various issues, is provided in appendix IV.

Table 1: Problems Interviewees Identified in the Guidelines System

Number interviewed in parentheses

Problem	District judges (10)	Prosecutors (12)	Defense attorneys (19)	Probation officers (12)
Guidelines "dehumanize" consideration of offenders	0	3	6	2
Sentences are too harsh/long[a]	3	5	4	7
Inflexible/do not allow judicial discretion	5	2	12	4
Complex, difficult to use, too many amendments	1	1	3	4
Give too much power to prosecutors	3	0	3	6
Other problems[b]	3	3	7	4

[a]Some interviewees said sentences in general were too harsh. Others said sentences were too harsh for drug offenders, minor participants in conspiracies, and/or first time offenders.

[b]Several other problems were cited by from one to five interviewees. These problems included that the system was too political, that it was too time consuming (time spent on appeals and judges' difficulties in finding time for civil cases were mentioned specifically), and that it based sentences on artificial distinctions that do not make sense (such as including the weight of the sugar cubes in which a drug is found in calculating drug quantities—a practice required by statute and incorporated into the guidelines).

Table 2: Long-Term Effects Interviewees Identified as Results of the Guidelines System

Number interviewed in parenthesis

Effect	District judges (10)	Prosecutors (12)	Defense attorneys (19)	Probation officers (12)
More offenders in prison and/or more prison costs or overcrowding	4	3	15	9
Rehabilitation concerns/poor adjustment by offenders released from long prison sentences	1	0	1	5
Prison unrest	1	0	2	1
Morale problems/ difficulties recruiting or retaining criminal justice personnel	3	0	3	1
More respect for judicial system	0	1	0	0

Table 3: Benefits Interviewees Identified in the Guidelines System

Number interviewed in parenthesis

Benefits	District judges (10)	Prosecutors (12)	Defense attorneys (19)	Probation officers (12)
Certainty of sentences	3	7	6	6
Less disparity/more uniformity in sentences	2	10	7	6
Other benefits[a]	2	0	2	1
No benefits	3	0	4	2

[a]Several other benefits were cited by one or more interviewees. These included more prison time for hard-core and white-collar criminals and clearer, better-documented evidentiary decisions by district judges.

Guidelines Appeared to Increase Criminal Justice System Workload

In September 1987, just before the sentencing guidelines went into effect, we reported on the potential impact of the guidelines on the operations of the criminal justice system as required by the act.[19] We said that the full impact of the guidelines would become clear only when there was sufficient empirical evidence available on how they were implemented.

The act required us to reevaluate the impact of the guidelines on system operations as a part of this report. However, the needed empirical data were still lacking. Reliable workload measures did not exist for the period before the guidelines were implemented that would allow reasonably accurate measurements of the impact of the guidelines. Consequently, it was not possible to determine with precision the impact of the guidelines to date.

On the basis of the limited statistical data available and our interviews with criminal justice system personnel in four districts, it appears that the guidelines increased system workload, particularly that of judges, prosecutors, probation officers, and defense attorneys. Interviewees said that some parts of their jobs were new under the guidelines and most interviewees said other parts took longer to do. Conversely, a few interviewees said some parts of their jobs now took less time to do.

Interviewees identified aspects of their jobs in each part of the process in which they participated that they thought were more time consuming under the guidelines than before. These aspects included the time to

[19]Sentencing Guidelines: Potential Impact on the Federal Criminal Justice System (GAO/GGD-87-111, Sept. 10, 1987).

investigate cases, negotiate plea agreements, resolve disputes, participate in sentencing hearings, and process cases and sentencing appeals through the courts. We discuss some of these aspects below. The remainder along with a sample of the views expressed in our interviews are provided in appendix V.

Under the guidelines, the great majority of criminal cases continue to be resolved by plea agreement. According to 1990 statistics from the Administrative Office of the U.S. Courts, more than 70 percent of all federal criminal cases were closed as a result of a plea agreement. The majority of prosecutors and defense attorneys we interviewed said that plea negotiations were more time consuming for them under the guidelines system, and half the district judges interviewed said that reviewing plea agreements was more time consuming for them. They said the additional time taken by the negotiations was a result of the more direct link between the offense agreed to in the plea and the sentence imposed under the guidelines.

The process of informally resolving disputes among prosecutors, defense attorneys, and probation officers over the contents of presentence investigation reports was new under the guidelines. As one probation officer explained, in the preguidelines era, both counsels merely submitted their own versions of the offense to the court, probation officers confidentially submitted presentence reports with sentencing recommendations to judges, and judges made their findings at formal sentencing hearings. Under the guidelines, probation officers are to submit drafts of presentence investigation reports to prosecutors and defense attorneys, and the parties are to attempt to resolve any disputes before the sentencing hearing. Disputes that cannot be resolved are to be decided by district judges at formal sentencing hearings. Table 4 presents the results of our interviews on this issue.

Table 4: Dispute Resolution Under the Guidelines

Interviewees	More time-consuming	Less time-consuming	Equally time-consuming	Total
District judges[a]	6	0	4	**10**
Federal prosecutors	11	0	1	**12**
Defense attorneys	17	2	0	**19**
Probation officers	12	0	0	**12**

[a]Some district judges said that they did not become involved in dispute resolution at this stage, while others said that they reviewed objections to presentence investigation reports and probation officers' responses as part of their preparation for sentencing.

The probation officers we interviewed said other aspects of their jobs related to researching and writing presentence investigations, in addition to assessing the impact of plea agreements and resolving disputes, were more time-consuming under the guidelines. The entire approach to completing presentence investigations had changed because of the guidelines. According to guidance provided by the Administrative Office before the guidelines, probation officers were to concentrate on assessing an offender's potential for rehabilitation and understanding the circumstances that caused the offender to commit the crime. Under the guidelines, probation officers instead are to address specific facts in cases that relate to guidelines applications.

Both defendants and prosecutors have an expanded right of appeal under the act. As might be expected, sentencing appeals have grown dramatically. In 1988, only 3 percent (about 180) of the 6,012 appeals filed in the circuit courts of appeals contained a sentencing issue. By 1991, 65 percent (about 6,400) of the 9,949 appeals filed did. See appendix V for additional information on case processing impacts, including the relative impact of the guidelines versus other recent changes to the criminal justice system.

The impact of the guidelines on case processing times was mixed. Nationally, the median processing time of criminal cases (from indictment to conviction) increased from 3.2 to 4.5 months from 1986 to 1990. The median time from conviction to sentencing rose as well, from 41 to 69 days. However, the proportion of cases that went to trial remained virtually unchanged at about 14 percent. The length of criminal trials also remained stable; more than 75 percent still took 3 days or less.

Two court components based requests for staff increases at least in part on the guidelines impact on workload. The Probation and Defender Services Divisions of the Administrative Office found that probation officers and federal and community defense attorneys could handle fewer numbers of cases under the guidelines because of the increased time spent on each case, and thus they justified requesting new positions.

The Probation Division was the only component of the Administrative Office that attempted to quantify the impact of the guidelines on its workload. The Administrative Office reported the impact to Congress in budget submissions, beginning with a request for supplemental funds in fiscal year 1988, and Congress consequently authorized 596 positions through fiscal year 1991. These positions represented 13.9 percent of the

total 4,303 positions requested for probation services that year. In fiscal year 1992, the Administrative Office requested 27 more probation positions to implement the guidelines.

The 1993 Administrative Office budget request to Congress estimated that the guidelines increased the time it took federal and community defense attorneys to defend cases by 25 to 50 percent. However, the Defender Services Chief said that because federal defense attorneys do not perform discrete tasks, as probation officers do, their work is not amenable to quantitative formulas, and the Division made no attempt to do workload studies to document the impact.

Conclusions

Data limitations precluded a definitive comparison of preguidelines and guidelines sentencing to determine the overall effectiveness of the sentencing guidelines in reducing sentencing disparity. But, for those few parts of the sentencing grid where there were sufficient data for analysis, both we and the Commission found that the amount of sentencing disparity had been reduced.

Delays in implementation of the guidelines also reduced the number of cases available for analysis in our and the Commission's efforts to evaluate current disparity under the guidelines. However, the analysis that could be done suggests that the guidelines had not been successful in eliminating unwarranted disparity in sentencing.

Similarly, the empirical data necessary to quantify the impact of the guidelines on the criminal justice system also did not exist. But, the limited data available and the perceptions of those involved in implementing the guidelines indicated that workload had increased.

The Commission faced enormously complex and contentious tasks in both establishing and evaluating the guidelines. Given the fundamental changes wrought by the guidelines, it should not be surprising that they were controversial in their conception and remain controversial in their implementation. The controversies continue to be fueled by the lack of empirical data resulting from limited experience with the guidelines. Also, a lack of data on the impact of charging, plea negotiation, and other presentence practices make it difficult to provide more definitive answers about the guidelines' impact on reducing disparity in sentencing. The absence of such evidence, coupled with the lack of data necessary to

evaluate the impacts of the guidelines on justice system operations, imply that the controversy surrounding the guidelines is unlikely to soon abate.

With the limited data currently available for analysis, it is worthwhile to consider how future analytical resources could best be used. As a practical matter, it is not possible to rectify the shortcomings of the preguidelines data to develop a more meaningful baseline for comparing sentencing outcomes before and after the guidelines. Thus, it would not be an effective or efficient use of resources to continue trying to develop such a baseline, or to perform such a comparison. A better focus would be to use the limited available resources to improve the data available for analysis of the impacts of sentencing that has occurred and is occurring under the guidelines.

In this respect, there are three primary areas where resources should be focused. First, additional data are needed on sentencing under the guidelines to allow a determination of the nature and extent that the sentences involve unwarranted disparity. This data will come with time and more experience with the guidelines. It is important to have data that will permit the Commission to evaluate whether that disparity is increasing, decreasing, or remaining stable.

Second, but equally important, data on the impact of presentencing decisions and policies should also be collected and analyzed to determine the impact of the guidelines on the criminal justice system in total. The exercise of discretion at each presentence stage can potentially affect the final sentence imposed and thus whether similar offenders are treated similarly. Baseline measures of presentence practices and of their impact on sentencing are needed to obtain a better understanding of the existing sentencing disparity under the guidelines and whether it continues. For example, continued evaluation of these practices could be used to identify differences among prosecutors and judges within a district, among districts within a circuit, and among circuits. An analysis of such differences would provide critical data for assessing the nature and extent of the unwarranted sentencing disparity that currently exists under the guidelines, the sources of the disparity, and the results of changes made to address whatever unwarranted disparity is found.

As with the case of pre- and postmeasures of disparity, there are no reliable baselines for measuring the pre- and postimpact of the guidelines on criminal justice system operations. Neither does it seem worthwhile to spend resources trying to develop such baselines. It does, however, seem

worthwhile to develop current baselines for use in measuring future changes in workload attributable to the guidelines. Such baseline data might include the workload impacts of the guidelines on investigators, prosecutors, judges, probation officers, and defense attorneys.

Recommendations

Congress should direct the Commission to continue its efforts to collect the data and perform the analyses needed to determine whether sentencing disparity is increasing, decreasing, or remaining stable under the guidelines and the amount of the disparity that is unwarranted. Given that it is not practical to rectify the shortcomings in preguidelines data and develop a more meaningful baseline for comparing sentencing outcomes before and after the guidelines, the Commission's focus should shift from preguidelines and guidelines comparisons to future trends in sentencing under the guidelines. Congress should also instruct the Commission to include in its evaluation and analyses the impact of the guidelines on presentencing decisions and policies which affect sentencing outcomes.

Similarly, Congress should direct the Commission, as a part of its overall evaluation responsibilities, to measure the ongoing impact of the guidelines on the operations of the criminal justice system. The Commission needs to work with other components of the system to assure that the appropriate measures are established and the requisite data is collected and analyzed.

Agency Comments

We provided a draft of this report for comment to the Commission, the Judicial Conference's Committee on Criminal Law, the Department of Justice, and the Administrative Office of the U.S. Courts. Their comments and our responses are summarized below. The full text of the comments of each are found in appendixes VI-IX.

Commission

In general the Commission was pleased that our report arrived "at the same basic findings and conclusions as the Commission." The Commission did raise specific points regarding issues related to the disparity analysis and our characterization of data limitations.

The Commission said that we "mischaracterized" the Commission's evaluation study and conclusions regarding the sentencing impact of such inappropriate or generally inappropriate factors as race and gender. According to the Commission, it is incorrect for us to characterize as

evidence of "unwarranted disparities" the statistically significant association found between such offender characteristics and the imposition of a sentence at the top, middle, or bottom of the guidelines range for a cell. The Commission believes such findings represent patterns of permissible judicial sentencing discretion, not unwarranted disparity. The Commission also stated that Congress did not define judicial sentencing variation within the guidelines range as disparity, much less unwarranted disparity.

Congress has defined disparity as defendants with similar criminal histories who commit similar criminal acts receiving dissimilar sentences. The Commission maintains that as long as sentences given for similar criminal acts fall within the same guidelines range, they cannot be dissimilar. We have taken a broader view than the Commission as to what constitutes dissimilar sentences and, accordingly, unwarranted disparity. Our view provides Congress with information on sentencing patterns within the guidelines ranges that are associated with the personal characteristics of offenders; the Commission's rationale does not. Sentencing variations within a given guidelines range can result in quite different sentencing outcomes. We have elaborated on this point on page 13.

The Commission said that despite the implication in our report that meaningful comparisons between preguidelines and guidelines sentencing could have been made if the preguidelines data contained additional cases, no amount of additional cases would have permitted this.

This is a misinterpretation of our position. We clearly noted that differences in much of the offender data available on those sentenced under the old and new systems made it difficult to reliably match and compare groups of preguidelines and guidelines offenders. It is precisely because available data prohibited a definitive comparison of sentences under the old and new sentencing systems that we recommended that all future evaluation focus solely on sentencing under the guidelines.

Judge Vincent L. Broderick, Chairman, Committee on Criminal Law of the Judicial Conference

While agreeing with our conclusion that additional research into the question of disparity is needed, Judge Broderick suggested that the research objective be narrowed to unwarranted disparity and that we give direction to the types of questions that should be addressed.

Our report does not address changes in unwarranted disparity under the guidelines compared to the preguidelines system because, as we noted, there were insufficient data to analyze changes in unwarranted disparity. Our report recommends that the Commission, working with the criminal justice community, identify and gather the data necessary to evaluate the degree of unwarranted disparity under the guidelines, including the effect of presentencing decisions on unwarranted disparity in sentencing.

Judge Broderick believed that comparing preguidelines and guidelines sentences of offenders in the same cells tells us little about overall disparity if there are differences in the information or in the application principles that are used to place cases within those cells.

We agree, but such data were not generally available for comparison. We relied, of necessity, on the data available. As we cautioned in our report, our analyses and conclusions were quite limited—i.e., to the cells within the guidelines grid we analyzed and for which we found that the dispersion of sentences around the average was less under the guidelines. It was not possible to generalize from the cells we examined to the guidelines as a whole.

According to Judge Broderick, in characterizing variation with respect to factors Congress and the Commission deemed to be either inappropriate, not generally appropriate, or not ordinarily relevant, the draft report presupposed the correctness of the determinations by Congress and the Commission. Judge Broderick urged that data should be collected on such factors to determine their impact on sentencing.

Our statutory mandate was limited to an examination of the impact of the guidelines, not the desirability of the policy choices Congress and the Commission have made in enacting and implementing the guidelines. As described in appendix IV, a number of those interviewed questioned some of these policy choices. We agree that data should be collected to determine the impact on sentencing of factors that are supposed to be inappropriate or generally inappropriate in sentencing.

Judge Broderick further stated that the report noted the difficulties in undertaking empirical research into the extent and causes of disparity related to prosecutors' decisions. However, he said that such research is critical and that analyzing the extensive literature on the guidelines could help to crystallize the issues that are being debated and turn them into

questions for empirical analysis that could be addressed by future evaluations. We agree.

Finally, Judge Broderick stated that the Commission does not have any particular expertise in determining the impact of the guidelines on the various agencies directly involved. The agencies involved should report on the effects that the guidelines have on them.

We did not recommend that the Commission unilaterally determine the impact of the guidelines on various federal agencies affected by them. We recommended that the Commission work with these agencies to ensure that each agency establishes appropriate measures and collects and analyzes the requisite data.

Department of Justice

The Department of Justice provided several comments on our work in the broad areas of sentencing disparities and prosecutorial discretion.

Justice said that our discussion of racial disparities in the letter was not an adequate summary of the work we did. According to Justice, we failed to clarify in the letter that our findings regarding the effect of race on sentences imposed were inconclusive. As a result, Justice thought that we might contribute to the perception of widespread antiblack discrimination in the justice system, when in fact, the evidence is not clear.

We were careful to note the limitations of our findings in both the letter and the appendixes. Specifically, there are racial disparities under the guidelines, but the effects of race are not manifested in a consistent manner. Given the data, it is too early to conclude that blacks received some sort of sentencing advantage or disadvantage relative to whites. Given the complexity of the guidelines, it made sense to focus on the specific manifestation of the effects of race, rather than draw general conclusions.

We stated that our analysis confirmed the existence of sentencing disparities, whether they are racial, gender, or socioeconomic, and that not all such disparities in sentencing have been eliminated under the guidelines. We clearly noted that our findings were preliminary, that additional data and research were needed to determine the extent of disparities in different contexts, and that measures of disparities should be monitored as the guidelines continue to be implemented. All of these statements are consistent with the findings in the appendixes.

In reporting the views of some of the court and criminal justice personnel we interviewed that the guidelines were too harsh (primarily for drug and first-time offenders), Justice said that we should note that there are mechanisms in the guidelines for adjusting sentences based on offenders' roles in crimes. It also said that we should note that amendments the Commission had proposed to the guidelines to take effect in November 1992 could reduce concerns about harshness of sentences.

In the Background section of our report, we briefly discussed adjustments allowed by the guidelines to account for aggravating or mitigating circumstances as part of our discussion of how the guidelines work. In reviewing the guidelines amendments that the Commission submitted to Congress, we found that while some proposed amendments could reduce sentences for some offenders, other amendments were likely to increase sentence lengths for some offenders. For example, the Commission proposed an amendment stating that lack of guidance as a youth and similar circumstances indicating a disadvantaged upbringing are not relevant grounds for departing from applicable guidelines ranges in sentencing. This proposed amendment could increase sentences for some offenders. As such, we did not find it appropriate to qualify statements made by interviewees on problems in the guidelines system with a discussion of how proposed amendments may or may not change their views.

Justice also said we needed additional information on steps it had taken to ensure that federal prosecutors apply discretion in charging and plea negotiation decisions according to uniform policies and procedures. In reporting that some of our interviewees thought the guidelines gave prosecutors too much power over sentencing outcomes based on their charging and plea bargaining decisions, Justice said that we should note that it has policies and procedures for prosecutors to follow in their decisionmaking. It also commented that compliance with these procedures was reviewed in evaluations of U.S. Attorneys offices.

We believe the judges, probation officers, and attorneys we interviewed were aware of Justice policies when they voiced their concerns and that the existence of the policies did not change their views. Justice's policy is to charge the most serious readily provable offense or offenses in criminal indictments, and charges are not to be bargained away or dropped unless the prosecutor doubts the government's ability to readily prove them. However, individual U.S. Attorneys offices have wide discretion within the national policy. Exceptions to the policy are authorized with supervisory

approval and notations in the case files. For example, prosecutors could drop readily provable charges due to a backlog of cases in the office or because the case would be time-consuming to try and would significantly reduce the number of cases disposed in the office.

With reference to Justice's comment that it reviews offices' compliance with national charging and plea bargaining procedures, as these reviews were described to us, they appeared to be very limited. An official of the Executive Office of U.S. Attorneys said the reviews consisted of asking several questions about compliance within the context of broader performance reviews. The official said that no reviews of case files had been done. Justice Evaluation and Review staff ask prosecutors if they are familiar with the national policies and if they are following them. They also ask prosecutors if they feel their guidelines training has been adequate and if they feel that the guidelines are being applied properly in their offices. According to the official, there was no indication of problems based on the interview results.

Justice was concerned that we inappropriately implied impropriety in prosecutors' decisionmaking when we referred to a Commission finding that in 17 percent of cases, decisions made by prosecutors on charge reductions and other plea bargains had an effect on sentencing.

We recognize that prosecutors must judge each case on its individual merits and that there are many instances when it is appropriate to reduce or dismiss charges because of proof problems, new evidence, or other reasons. However, our principal point in citing the study is a valid one—prosecutorial decisions can have a major impact on the ultimate sentence imposed. Understanding the impact of those decisions on sentencing disparity is essential to an evaluation of any continuing sentencing disparity under the guidelines.

Administrative Office of the U.S. Courts

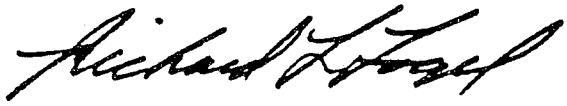
The comments of the Administrative Office primarily addressed the workload section of our report. The Administrative Office said that we did not use all available data in our workload analysis and was concerned that we said because workload data were limited, our analysis could not be conclusive. They said that we could have made a more definitive statement on how the guidelines affected workload, especially for probation officers.

We made a significant effort to collect and analyze data related to the guidelines' effect on the workload of court and other criminal justice

personnel. In our report, we said that the guidelines appeared to increase the workload of most criminal justice personnel but that empirical data necessary to quantify the effect of the guidelines were not available. Many other changes occurred in the criminal justice system during the same time that the guidelines were being implemented, and we believe none of the available data made it possible to isolate the effects of these changes from those of the guidelines. We also noted that the Probation Division was the only Administrative Office component to attempt to quantify the effects of the guidelines and that they did indeed receive additional staff based on their workload study.

Copies of this report will be made available to the Commission, the Judicial Conference, the Administrative Office of the U.S. Courts, the Department of Justice, and other interested parties. It will also be made available to others upon request.

Major contributors are listed in appendix X. If you have any questions please contact me on (202) 275-6059 or Lynda Willis on (202) 566-0026.



Richard L. Fogel
Assistant Comptroller General

Contents

Letter		1
Appendix I Changes in Disparity Between Preguidelines and Guidelines Sentences		40
	Sentencing Commission's Analysis of Sentencing Disparities	40
	GAO's Analysis of Sentencing Dispersion	41
Appendix II Disparity in Sentencing Under the Guidelines		62
	Background	62
	The Analysis	65
	Effects of Factors on Within Range Sentences	85
	Summary	110
Appendix III Racial Differences in Robbery Sentencing Patterns Under the Guidelines		111
	Objectives and Approach	111
	Results	111
	Racial Differences in Mean Lengths of Sentences Imposed on Convicted Bank Robbers	112
	Explaining the Racial Gap in Sentences Imposed: Testing for Racial Differences in Offense and Offender Characteristics and in the Effects of Characteristics on Sentence Lengths	117
	General Effects	133
Appendix IV Debate on Guidelines' Problems, Benefits, and Effects Continues		143
	Objectives, Scope, and Methodology	143
	Problems, Benefits, and Effects Interviewees Saw as a Result of the Guidelines	148
	Majority of Interviewees Thought Judges Should Have More Grounds to Depart From the Guidelines	150
	Other Studies of the Guidelines' Impact Found Similar Benefits, Problems, and Effects	154
	Most Interviewees Thought Unwarranted Sentencing Disparity Did Not Occur Frequently	155

Appendix V Guidelines' Impact on Staff, Workloads, and Case Processing Times of the Courts and Court-Related and Investigative Agencies 161

Objectives, Scope, and Methodology 161
Interviewees Said Some Aspects of Their Work Took Longer Under the Guidelines 162
Data Did Not Show Increases in Number or Length of Trials Under the Guidelines 170
Data Showed Link Between Guidelines' Implementation and Increases in Case Processing Time Frames 171
Interviewees Said the Guidelines Had the Most Significant Effect on Workload of All Recent Interventions 172
Two Components Based Staffing and Budget Increases on the Guidelines' Impact on Workload 174
Conclusion 175

Appendix VI Comments From the United States Sentencing Commission 177

Appendix VII Comments From Judge Vincent L. Broderick, Chairman, Committee on Criminal Law of the Judicial Conference of the United States 183

Appendix VIII Comments From the U.S. Department of Justice 189

Appendix IX
Comments From the Administrative Office of the U.S. Courts 193

Appendix X
Major Contributors to This Report 195

Tables

Table 1: Problems Interviewees Identified in the Guidelines System 19
Table 2: Long-Term Effects Interviewees Identified as Results of the Guidelines System 19
Table 3: Benefits Interviewees Identified in the Guidelines System 20
Table 4: Dispute Resolution Under the Guidelines 21
Table I.1: Number of Records Deleted From GAO Analysis Due to Missing Information 44
Table I.2: Derivation of Criminal History Category and Offense Level Variables for the Guidelines File 45
Table I.3: Summary of Weighted Records Analyzed by Offense Group 49
Table I.4: Results of the Analysis of Coefficient of Variance of Estimated Time to Be Served—All Groups Analyzed 57
Table I.5: Results of the Analysis of Coefficient of Variance of Estimated Time to Be Served—Drugs and Theft and Fraud 58
Table I.6: Standard Deviations in Months of Imposed Sentence and Time Served, Controlling for Guidelines Offense Level. Heroin and Cocaine Category 1 Offenders 59
Table I.7: Guidelines Levels of Drug Amounts 60
Table I.8: Tests for Disparity in Mean Sentence Imposed and Mean Time Served Based on Demographic Variables. Combined Heroin and Cocaine Category 1 Offenders, Offense Level 26 61
Table II.1: Frequencies and Percentages in the Cross-Classification of Race and Sentence Location 66
Table II.2: Race by Within Range Sentences 67

Table II.3: Race by Departure Vs. Within Range Sentences 67
Table II.4: Chi-Square Values for the Full and Partitioned Two-Way Tables Involving Sentence Location and Selected Factors 69
Table II.5: Additional Partitionings of the Two-Way Departure Vs. Within Range Subtables Involving Circuit, Criminal History, and Seriousness Level 70
Table II.6: Alternative Models for the Two-Way Departure Vs. Within Range Subtables Involving Circuit, Criminal History, Seriousness Level, Type of Offense, and Mode of Disposition 71
Table II.7: Expected Frequencies Under the Preferred Models for the Two-Way Departure Vs. Within Range Subtables, and Odds and Odds Ratios Derived From Them 74
Table II.8: Models for the Three-Way Tables in Which Departure Vs. Within Range Sentences Were Cross-Classified by Circuit and Criminal History, Seriousness Level, Type of Offense, and Mode of Disposition 83
Table II.9: Odds Ratios Estimating the Association Between Circuit and Sentence Location, Before and After Controlling Four Factors 83
Table II.10: Additional Partitionings of the Within Range Subtables 86
Table II.11: Final Partitionings of the Two-Way Within Range Subtables Involving Circuit, Criminal History, and Seriousness Level 88
Table II.12: Alternative Models for the Collapsed Two-Way Within Range Subtables 89
Table II.13: Expected Frequencies Under the Preferred Models for the Two-Way Within Range Subtables, and Odds and Odds Ratios Derived From Them 90
Table II.14: Models for the Three-Way Tables in Which Within Range Sentences Are Cross-Classified by Race and Criminal History, Seriousness Level, Type of Offense, and Mode of Disposition 104
Table II.15: Models for the Three-Way Tables in Which Within Range Sentences Are Cross-Classified by Gender and Criminal History, Seriousness Level, Type of Offense, and Mode of Disposition 105

Table II.16: Models for the Three-Way Tables in Which Within Range Sentences Are Cross-Classified by Marital Status and Criminal History, Seriousness Level, Type of Offense, and Mode of Disposition 106
Table II.17: Models for the Three-Way Tables in Which Within Range Sentences Are Cross-Classified by Employment Status and Criminal History, Seriousness Level, Type of Offense, and Mode of Disposition 107
Table II.18: Models for the Three-Way Tables in Which Within Range Sentences Are Cross-Classified by Age and Criminal History, Seriousness Level, Type of Offense, and Mode of Disposition 108
Table II.19: Models for the Three-Way Tables in Which Within Range Sentences Are Cross-Classified by Circuit and Criminal History, Seriousness Level, Type of Offense, Mode of Disposition 109
Table II.20: Odds Ratios Estimating the Effects of the Extralegal Factors on Sentence Location, Before and After Controlling for Four Legally Relevant Factors 110
Table III.1: Mean Lengths of Sentences Imposed, Robbery, Fiscal Year 1990 114
Table III.2: Mean Lengths of Sentences Imposed by Criminal History Category, Robbery, Fiscal Year 1990 116
Table III.3: Means and Standard Deviations in Models Without Guidelines Scores 127
Table III.4: Means and Standard Deviations in Models Including Guidelines Scores 129
Table III.5: Regression Results, Models Without Guidelines Scores, Dependent Variable: Natural Log of Sentence Length Imposed 131
Table III.6: Regression Results, Models With Guidelines Scores, Dependent Variable: Natural Log of Sentence Length Imposed 132
Table III.7: Analysis of the Black/White Sentence Differential: Models Without Guidelines Scores 136
Table III.8: Analysis of the Black/White Sentence Differential: Guidelines Scores Models 137
Table III.9: Details of Parameter Estimates and Decomposition: Models Without Guidelines Scores 139
Table III.10: Details of Parameter Estimates and Decomposition: Models Guidelines Scores 141

Table IV.1: Structured Interviews Done by Interviewee Group and Court District 143
Table IV.2: Interviewee Responses on Whether Factors Either Not Included in the Sentencing Guidelines or Included as Not Ordinarily Relevant Should Be Grounds for Lawful Departure From the Guidelines 151
Table IV.3: Interviewee Responses on Who Should Be Permitted to Make Motions for Downward Departures Based on Substantial Assistance 153
Table IV.4: Interviewee Responses on Changes in Sentencing Disparity in Their Districts Under the Guidelines 156
Table IV.5: Interviewee Responses on How Frequently Unwarranted Sentencing Disparity Occurred in Their Districts Under the Guidelines 156
Table V.1: Impact of the Guidelines on Time Required to Negotiate/Review Plea Agreements 163
Table V.2: Impact of the Guidelines on Time Required to Resolve Disputes 165
Table V.3: Impact of the Guidelines on Time Required to Complete Other Presentence Investigation Work 166
Table V.4: Impact of Guidelines on Time Required to Participate in Sentencing Hearings 167
Table V.5: Criminal Appeals Filed, Years Ending June 30, 1988, Through June 30, 1991 168
Table V.6: National Case Disposition Rates for All Defendants, Years Ending June 30, 1986, Through June 30, 1990 170
Table V.7: Percentage of Trials Completed by Days for All Federal Defendants, Years Ending June 30, 1986, Through June 30, 1990 170
Table V.8: Median Number of Months From Case Filing to Disposition for All Crime Types, Years Ending June 30, 1986, Through June 30, 1990 171
Table V.9: Median Number of Days for Cases to Move From Conviction to Sentencing, Years Ending June 30, 1986, Through June 30, 1990 172
Table V.10: Rank of Effects of Guidelines and Other Interventions on Workload 173

Figures

Figure 1: Sentencing Guidelines Table 8
Figure I.1: Analysis Groups Used by GAO 48
Figure I.2: Analysis Groups GAO Used for Drug Offenders 50
Figure I.3: Analysis Groups GAO Used for Theft and Fraud Offenders 52
Figure I.4: Analysis Groups GAO Used for Firearms Offenders 54
Figure II.1: Expected Odds on Downward and Upward Departures Vs. Within Range Sentences, by Circuit 77
Figure II.2: Expected Odds on Downward and Upward Departures Vs. Within Range Sentences, by Criminal History 78
Figure II.3: Expected Odds on Downward and Upward Departures Vs. Within Range Sentences, by Offense Seriousness Level 79
Figure II.4: Expected Odds on Downward and Upward Departures Vs. Within Range Sentences, by Type of Offense 80
Figure II.5: Expected Odds on Downward and Upward Departures Vs. Within Range Sentences, by Mode of Disposition 81
Figure II.6: Expected Odds on Bottom and Top of the Range Vs. Middle Range Sentences, by Race 94
Figure II.7: Expected Odds on Bottom and Top of the Range Vs. Middle Range Sentences, by Gender 95
Figure II.8: Expected Odds on Bottom and Top of the Range Vs. Middle Range Sentences, by Marital Status 96
Figure II.9: Expected Odds on Bottom and Top of the Range Vs. Middle Range Sentences, by Employment Status 97
Figure II.10: Expected Odds on Bottom and Top of the Range Vs. Middle Range Sentences, by Age 98
Figure II.11: Expected Odds on Bottom and Top of the Range Vs. Middle Range Sentences, by Circuit 99
Figure II.12: Expected Odds on Bottom and Top of the Range Vs. Middle Range Sentences, by Criminal History 100
Figure II.13: Expected Odds on Bottom and Top of the Range Vs. Middle Range Sentences, by Offense Seriousness Level 101
Figure II.14: Expected Odds on Bottom and Top of the Range Vs. Middle Range Sentences, by Type of Offense 102
Figure II.15: Expected Odds on Bottom and Top of the Range Vs. Middle Range Sentences, by Mode of Disposition 103

Figure IV.1: Criminal Cases Commenced in the Four Court Districts We Visited, Year Ending June 30, 1991 146
Figure IV.2: Criminal Cases Commenced, Years Ending June 30, 1987, Through June 30, 1991 147

Abbreviations

DEA	Drug Enforcement Administration
FBI	Federal Bureau of Investigation
FPSSIS	Federal Probation Sentencing and Supervision Information System
MONFY90	monitoring and reporting system database for fiscal year 1990
USMS	U.S. Marshals Service

Changes in Disparity Between Preguidelines and Guidelines Sentences

With the limited data available we assessed the impact of the guidelines on reducing sentencing disparity. We matched groups of preguidelines and guidelines offenders with similar criminal histories and offense levels and estimated the variances in their expected time served. We found, after adjusting for increases in time served in the guidelines period, that sentencing disparity had decreased under the guidelines for the groups for which we had adequate data to analyze. Our results were consistent with those of the Commission, which concluded in general that the spread of sentences imposed and time served had narrowed under the guidelines.

We also found that, due to data limitations, neither we nor the United States Sentencing Commission were able to address fully the question of whether the guidelines reduced unwarranted disparity. Left unexplored by us and the Commission, for example, were questions related to the spread of preguidelines and guidelines sentences for similarly situated offenders who differed only on legally irrelevant factors such as race, gender, or socioeconomic status,[1] or on the impacts of presentencing decisions on the location of offenders on the sentencing grid.

Sentencing Commission's Analysis of Sentencing Disparities

To measure the impact of the guidelines on sentencing disparities, the Commission compared sentence outcomes for similarly situated groups of preguidelines and guidelines bank robbers, bank embezzlers, heroin distributors, and cocaine distributors.[2] The Commission identified similar defendants by grouping offenders who had specific characteristics that related to offense conduct. They matched on relevant preguidelines factors that were used to develop offense groupings under the guidelines. For example, similar bank robbers were those who were similar in terms of factors such as dollar loss, weapon use, victim injury, role in the offense, and so on. By choosing to match offenders, the Commission made an explicit trade-off between the size of the samples available to analyze and the precision with which it was able to identify similarly situated preguidelines and guidelines offenders.[3]

[1]We noted that we were able to provide some preliminary data on the relationship between demographic factors and the spread of sentences. We review these findings in our concluding section.

[2]The Federal Sentencing Guidelines, Vol. II, Ch. 4, Parts I-V, pp. 269-299.

[3]The Sentencing Commission analyzed relatively small samples that were chosen purposefully. For example, the respective fractions of preguidelines and guidelines cases in their datasets that the Commission analyzed were bank robbers, 14.5/22.8 percent; bank embezzlers, 11.8/20.9 percent; heroin distributors, 7.4/7.6 percent; and cocaine distributors, 13.3/5.0 percent. These numbers can be found on pp. 278, 283, 287, 291, 295, and 298 of the Sentencing Commission's evaluation report, The Federal Sentencing Guidelines.

In each of its analyses, the Commission concluded that the variances in sentences imposed under the guidelines had decreased, but for only three of eight groups did it conclude that reductions in the variances of time served under the guidelines were statistically significant.

Because of the trade-off between precision in defining similarly situated offenders and sample sizes, the Commission's results were not generalizable to other groups of similarly situated offenders or even to other offenders within the specific offense types analyzed. Moreover, the particular groups of offenders chosen were not chosen because they were representative or typical of offenders generally but because they were the largest groups of a large number of groups of offenders with similar criminal histories and offense characteristics.

GAO's Analysis of Sentencing Dispersion

Like the Commission, we sought to make a general determination of whether the sentencing guidelines reduced disparities in sentences. To do this, we compared the dispersion[4] of criminal sentences—expected prison time to be served—for samples of similarly situated offenders sentenced under the preguidelines system with samples of offenders sentenced under the guidelines. Expected time to be served is the amount of time a defendant can expect to spend in prison at the time of sentencing. We analyzed changes in dispersion by estimating preguidelines and guidelines variances and coefficients of variation in expected time served. We estimated coefficients of variation to account for large increases in time served.

We defined similarly situated preguidelines and guidelines offenders by their criminal history categories and offense severity levels as defined by the guidelines. We matched preguidelines and guidelines offenders who had similar sets of criminal history scores and offense severity levels. These scores are used to locate an offender's position on the sentencing table. That position determines the presumptive sentence range available to judges. The sentencing guidelines table is composed of 6 criminal history categories and 43 offense levels, forming 258 individual cells. Adjacent cells have overlapping sentence ranges. Offenders are to receive a prison sentence commensurate with their criminal history category and offense level.

However, judges are not constrained to sentence offenders to sentences that fall within the prescribed ranges. Judges may legitimately give

[4]By dispersion, we mean the spread of sentences around the average sentence.

offenders sentences that fall outside the sentence range indicated by the offenders' criminal history categories and offense levels for any of the following reasons:

- departures from the guidelines for mitigating or aggravating circumstances where the judge believes that the guidelines do not accurately reflect the offender's criminal history or behavior,
- departures from the guidelines in cases where the defendant has substantially assisted the prosecution,
- situations where consecutive sentences are given for multiple convictions, or
- situations where statute overrules the guidelines, such as in the case of mandatory minimum sentences or where statutory maximum sentences are below the guidelines.

Because the data we used did not contain information on departures, we were unable to determine the specific reasons why sentences fell outside of ranges.

For each of our pairs of preguidelines and guidelines groups of offenders, we estimated the mean and variance in expected time to be served in prison. For each group showing a larger variance after the guidelines and surrounding groups, a coefficient of variation in expected time to be served in prison was calculated. We used "expected time to be served" in prison because some of the defendants in our analysis had not served their entire sentences. Therefore, actual time served was not available and some estimate was needed in order to measure sentencing dispersion.

The act changed the fundamental nature of sentencing in the federal system by abolishing parole. A preguidelines sentence to a large extent represents the maximum term of imprisonment assuming no reduction for good conduct or parole. Few individuals served or expected to serve such sentences; rather, most served between one-third and two-thirds of the original sentence imposed. Under the act, individuals must serve their full sentence less a maximum reduction of 54 days per year for good behavior.[5]

For the preguidelines period the expected time to be served estimate we used was the presumptive parole date. This date is established by the Parole Commission at the beginning of the service of an offender's sentence. It is a date on which it is presumed that release will take place, provided the prisoner maintains a good institutional conduct record and

[5]Life sentences and sentences of 1 year or less do not qualify for this reduction.

has developed adequate release plans. Similarly, we used the guidelines sentences imposed, less the maximum amount of credit for good behavior, to estimate the expected time to be served for the guidelines period. Both our guidelines and preguidelines measures of expected time served were the same measures used by the Commission in its analysis of disparity.

We examined patterns of dispersion in sentences, as measured by the variance and coefficient of variation. Less dispersion in the sentences of guidelines offenders indicated that disparity was reduced, while greater dispersion indicated that disparity was not reduced.

Available data made it possible to conclude only that for the offense categories we were able to analyze, sentencing dispersion generally decreased under the guidelines. In those instances where variances increased under the guidelines, the increase was tied to increases in estimated mean time served. When we adjusted for this, dispersion—as measured by the coefficient of variation—decreased in all categories analyzed. Our dispersion analyses did not identify which factors determined the length of the sentence, nor did it directly answer the question of whether unwarranted sentencing disparity had been reduced.

Sources of Data

We obtained the data for our analyses from the same two agencies' databases that the Commission used in its study. Our preguidelines data were drawn from an augmented Federal Probation Sentencing and Supervision Information System of the Administrative Office of the U.S. Courts (FPSSIS) dataset constructed by the Commission representing offenders sentenced in 1985. This dataset was developed to assess the impact of the guidelines on the federal prison population in a prison impact model. This model calculated the appropriate guideline criminal history score and offense level for each offender in a stratified sample of 10,575 (27,761 weighted) offenders sentenced in fiscal year 1985, taken out of a universe of approximately 40,000 offenders. To our knowledge this file is the only dataset that has current guidelines information on criminal history scores and offense levels for offenders sentenced prior to the sentencing guidelines.

Our data on the guidelines sample of offenders were obtained from the Commission's monitoring and reporting system database for fiscal year 1990 (MONFY90). This dataset contained documentation on 29,011 cases sentenced during fiscal year 1990 and received by the Commission. For

each case it had information on the sentence imposed, the guideline offense level, and the guideline criminal history category.

Variables Used in the Analyses and Data Verification

We performed a variety of data validity checks on each of our variables. We identified missing, invalid, or illogical data. If our checks showed out-of-range or invalid values, we contacted staff at the Federal Bureau of Prisons and the Commission to obtain an explanation and to determine how they handled it in their analysis; otherwise, we deleted the record from our analysis.

We found no significant problems with the preguidelines data file. We dropped approximately .5 percent unweighted (58 records) or 1 percent weighted (301 records) of the preguidelines offenders. However, we discovered several problems in the guidelines data file. We deleted approximately 5 percent (1,469) of the guidelines records from our analysis because they were missing data on key elements, such as offense level, criminal history category, or estimated time to be served. Table I.1 shows the number of records deleted from both the preguidelines and guidelines data files.

Table I.1: Number of Records Deleted From GAO Analysis Due to Missing Information

	Weighted	Unweighted
Preguidelines file		
Number of records	27,761	10,575
Number deleted	301	58
Number of records remaining	27,460	10,517
Guidelines file		
Number of records	29,011	29,011
Number deleted	1,469	1,469
Number of records remaining	27,542	27,542

The preguidelines dataset was relatively complete. It had no missing values for offense level scores, and we were able to compute criminal history categories from the criminal history scores in the dataset for all our observations.

For the guidelines data, determining the criminal history category and final offense severity level was more involved. We took advantage of three alternative measures of each. First, we used a variable calculated by the Commission that denoted either criminal history categories or final

offense severity levels. Second, if an observation was missing on that variable, we used the measure of criminal history or offense severity level that was recorded in the statement of reasons. We used this replacement method because where the data existed, there was a 100 percent agreement between the Commission's and the statement of reason's variables for criminal history category and offense level.

Third, if observations were missing from either of those variables, we used the criminal history or offense severity contained in the presentence investigation report. If the calculated variable was missing, we replaced it with the variable from the statement of reasons. We used this replacement strategy because where data on both variables existed there was 97 percent agreement between the Commission's and the presentence report's criminal history variable, and there was an 82 percent agreement between the Commission's and the presentence report's offense level variable. Table I.2 shows the number of times we used each one of the three variables.[6]

Finally, if measures on either criminal history or offense severity were missing on all three variables, we deleted the record from our analysis. We omitted 895 records for this reason.

Table I.2: Derivation of Criminal History Category and Offense Level Variables for the Guidelines File

Variable used	Number of times used
Criminal history category	
Commission	19,837
Statement of reason	7
Presentence report	7,698
Offense level	
Commission	19,945
Statement of reason	7
Presentencing	7,590
Total	**27,542**

Our dependent variable was an estimate of the expected time to be served by offenders in prison. The preguidelines dataset contained estimates of expected time to be served, but the guidelines dataset did not; therefore, we estimated time to be served for guidelines cases. We estimated time to

[6]The Sentencing Commission created its calculated variable by first using the statement of reason variable if it was available. As a result, we expected that there would not be any Statement of Reason variable available when there was a missing calculated variable. However, in seven records where the calculated variable was missing, we had a value for the Statement of Reason variable.

be served by subtracting the maximum of 54 days of "good time" per year for each year served. Good time is subtracted from sentences of more than 1 year and less than a life sentence. An offender may actually receive less than 54 days of good time for bad behavior, but for our purpose we assumed the maximum amount of good time was given (i.e., 54 days per year).

We performed two final checks. First, because the FPSSIS data file coded a life sentence as 360 months, we recoded the Commission's value for life sentence of 996 months to 360 months for comparability.

Finally, because our guidelines dependent variable was derived from observations on sentences imposed, we checked that where an offender was coded as receiving a prison sentence, there was a positive value for prison time. If there was no value, the offender was deleted from our analysis. We deleted 188 offenders for this reason. When the prison sentence value was missing or indeterminable, we tried to verify the reliability of the record by checking to see if there was a valid value for probation; if there was, we kept it. Our reasoning was that an offender cannot be given probation and a prison sentence. So if there was a valid value for probation, we set the missing value for the prison sentence equal to zero. We were able to retain 14 offenders in our analysis as a result of this recoding; however, 386 records could not be recoded and were deleted.

Division of Sentencing Table Into Smaller Groups

To compare dispersion in sentences between preguidelines and guidelines offenders, we matched offenders on criminal history and offense severity scores. However, we did our analysis on a cell-by-cell basis. We compared sentence outcomes of preguidelines offenders with those corresponding to the same cell on the sentencing table for guidelines offenders. We conducted four separate analyses. First, we analyzed all offenders, regardless of the offense for which they were convicted. Then we separately analyzed offenders convicted for drug offenses, then theft and fraud, and finally, firearms offenses.

We limited our analyses to cells with relatively large numbers of preguidelines and guidelines cases. In the analysis of all types of crimes, single cells were required to have a minimum of 19 records (50 weighted)[7] in the preguidelines data file and 50 in the guidelines data file. In the 3

[7]We went below our standard of 19 (50 weighted) cases in one preguidelines group—group 48. It had 18 unweighted cases. On our standard, see Herman Burstein, Attribute Sampling (New York: McGraw-Hill, 1971), pp. 41-42.

analyses of individual crime types, each single cell was required to have a minimum of 15 records (25 weighted) in the preguidelines and 25 in the guidelines. We reduced the number of records required for the individual crime types because they contained fewer overall observations and we were dealing with fewer offense types.

If an individual cell contained the minimum number of observations, it remained in our analysis. Cells that did not were deleted from our study, with the exception of cells at the very top and very bottom of the sentencing table. Where single cells had insufficient numbers of cases at the very top and very bottom of the sentencing table, we aggregated across individual cells. Cells at the top of the table were aggregated because they included sentences of probation or imprisonment of up to 6 months. These cells were not combined with other cells that contained harsher sentences. Similarly, all cells with a sentence of 360 months to life at the bottom of the table were combined to form a single group. Cells were combined only if there were enough cases within the cells of the same sentencing range to meet the case number requirement.

The rules we used for grouping cells were developed as a result of sensitivity analysis we did on the preguidelines and guidelines data. The sensitivity analysis showed that grouping cells with different sentencing ranges affected the results of the sentencing dispersion analysis. However, offenders in the top and bottom of the chart could be grouped together without affecting the measure of sentencing dispersion because they contained the same sentencing range.

All Offenders

The first analysis we performed was done on all offenders regardless of crime type. After all data verification was completed, the preguidelines dataset contained 10,517 records (27,460 weighted) and the guidelines dataset contained 27,542 records. Using our rules for grouping cells, the sentencing table was divided into 68 groups (see figure I.1). The 68 groups analyzed contained a total of 7,976 preguidelines records (22,375weighted) and 21,782 guidelines records. Together these 68 groups accounted for 76 percent and 79 percent of the total weighted preguidelines and guideline records, respectively. Table I.3 summarizes the number of records analyzed in our analyses.

Figure I.1: Analysis Groups Used by GAO

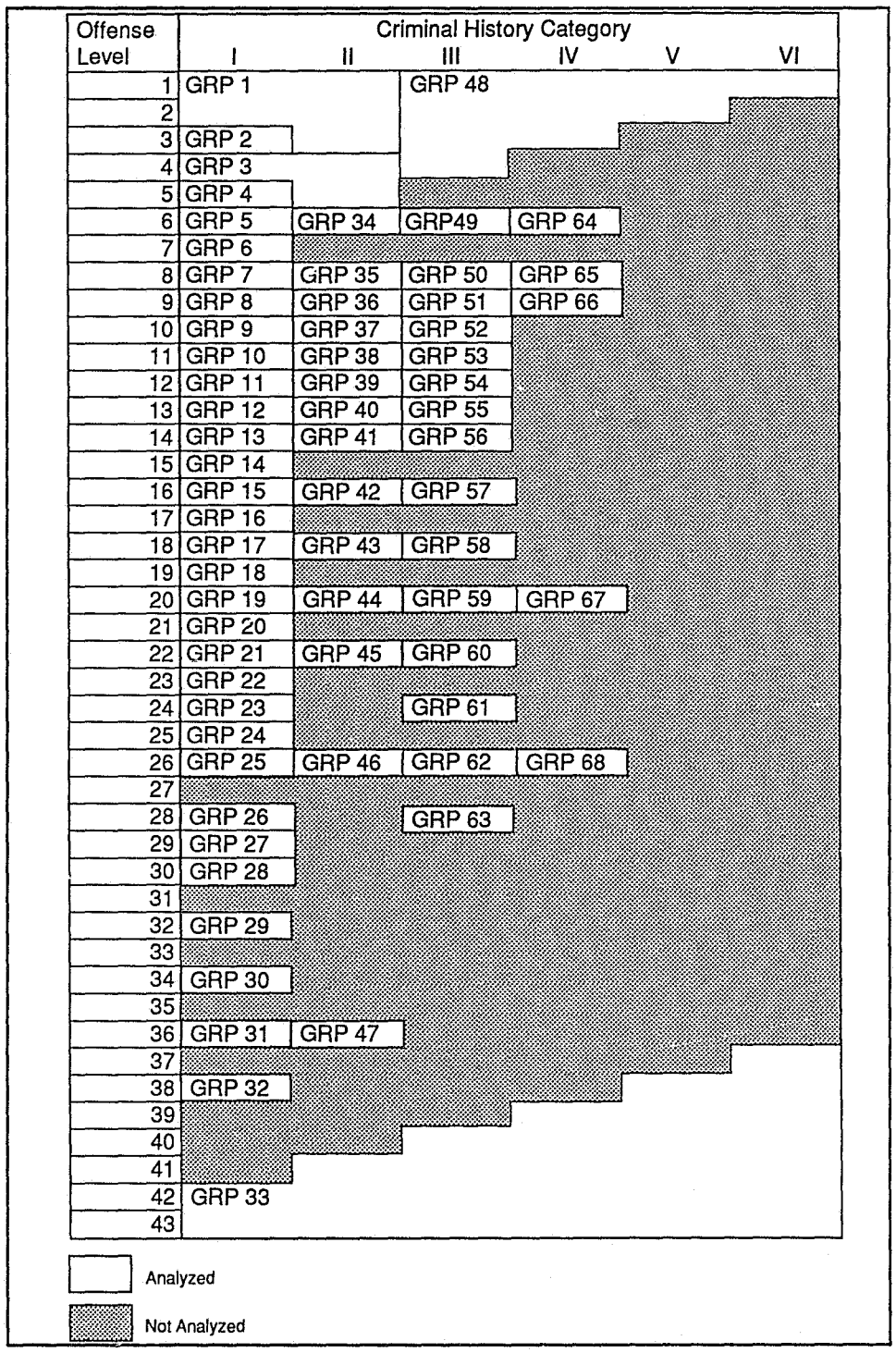


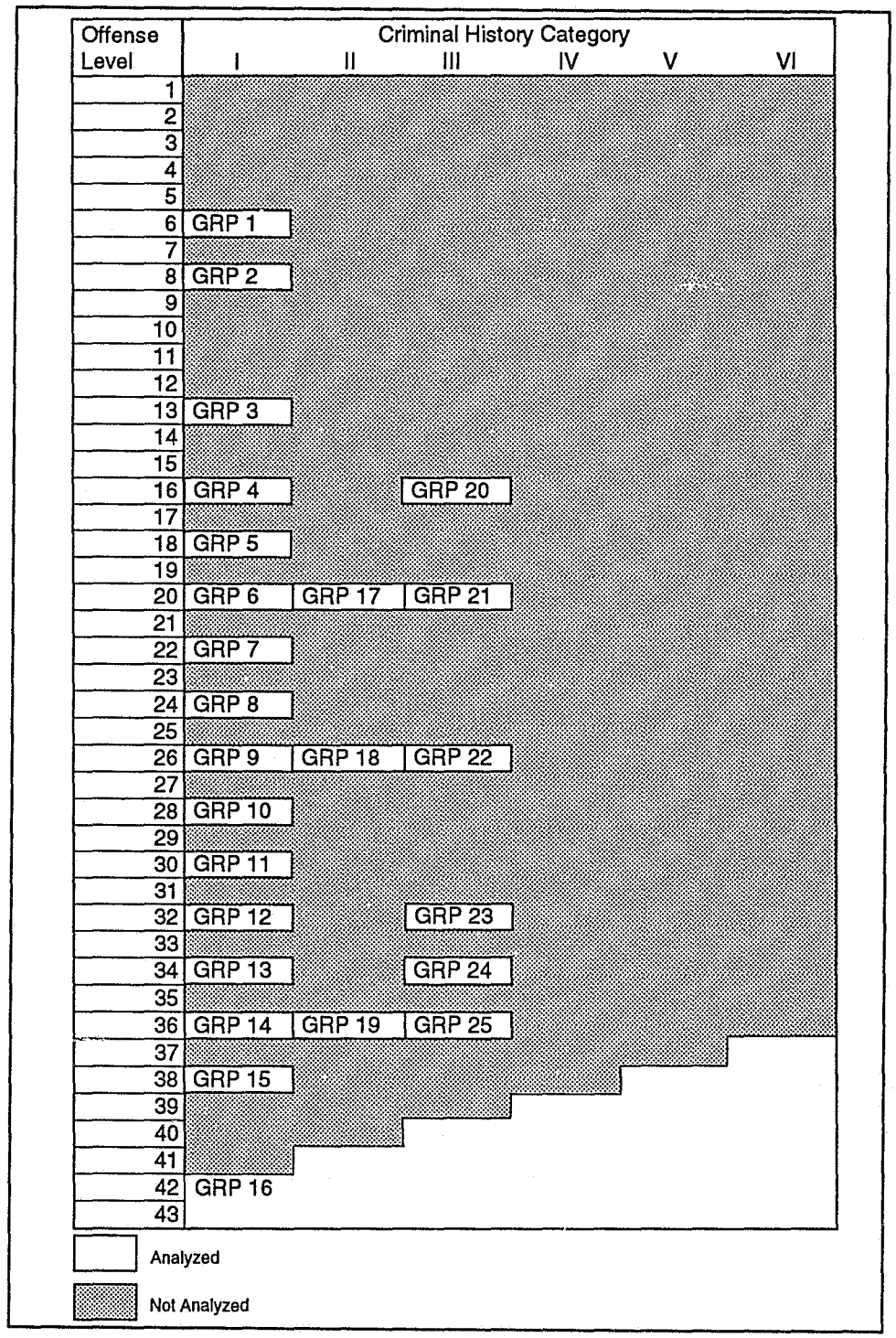
Table I.3: Summary of Weighted Records Analyzed by Offense Group

Sample	Analysis	Total number	Number analyzed	Percent analyzed
Preguidelines	All offenses	**10,517**	7,976	75.83
Guidelines	All offenses	**27,542**	21,782	79.09
Preguidelines	Drug	**2,875**	1,682	58.50
Guidelines	Drug	**11,326**	6,434	56.81
Preguidelines	Theft	**2,968**	2,295	77.32
Guidelines	Theft	**5,966**	5,043	84.53
Preguidelines	Firearms	**384**	66	17.19
Guidelines	Firearms	**1,409**	60	4.26

Drug Offenders

Our second analysis was done on offenders who committed drug offenses. We included the three following general offense types in our drug analysis: (1) importation and distribution, (2) simple possession, and (3) use of a communication facility in committing a drug offense. The two drug data files contained 2,875 preguidelines records (8,215 weighted) and 11,326 guidelines records. Figure 1.2 shows how the preguidelines and guidelines drug data files were divided into 25 groups. The 25 groups we analyzed contained a total of 1,682 preguidelines records (4,836 weighted) and 6,434 guidelines records. Together these 25 groups accounted for 59 percent and 57 percent of the total weighted preguidelines and guidelines records, respectively. Due to a lack of records, we did not analyze certain areas of the sentencing table. Most of the drug offenders in the following areas were not analyzed: (1) criminal history categories IV-VI, (2) offenders eligible for probation without conditions, and (3) offenders eligible for probation with conditions or a split sentence of prison and supervised release.

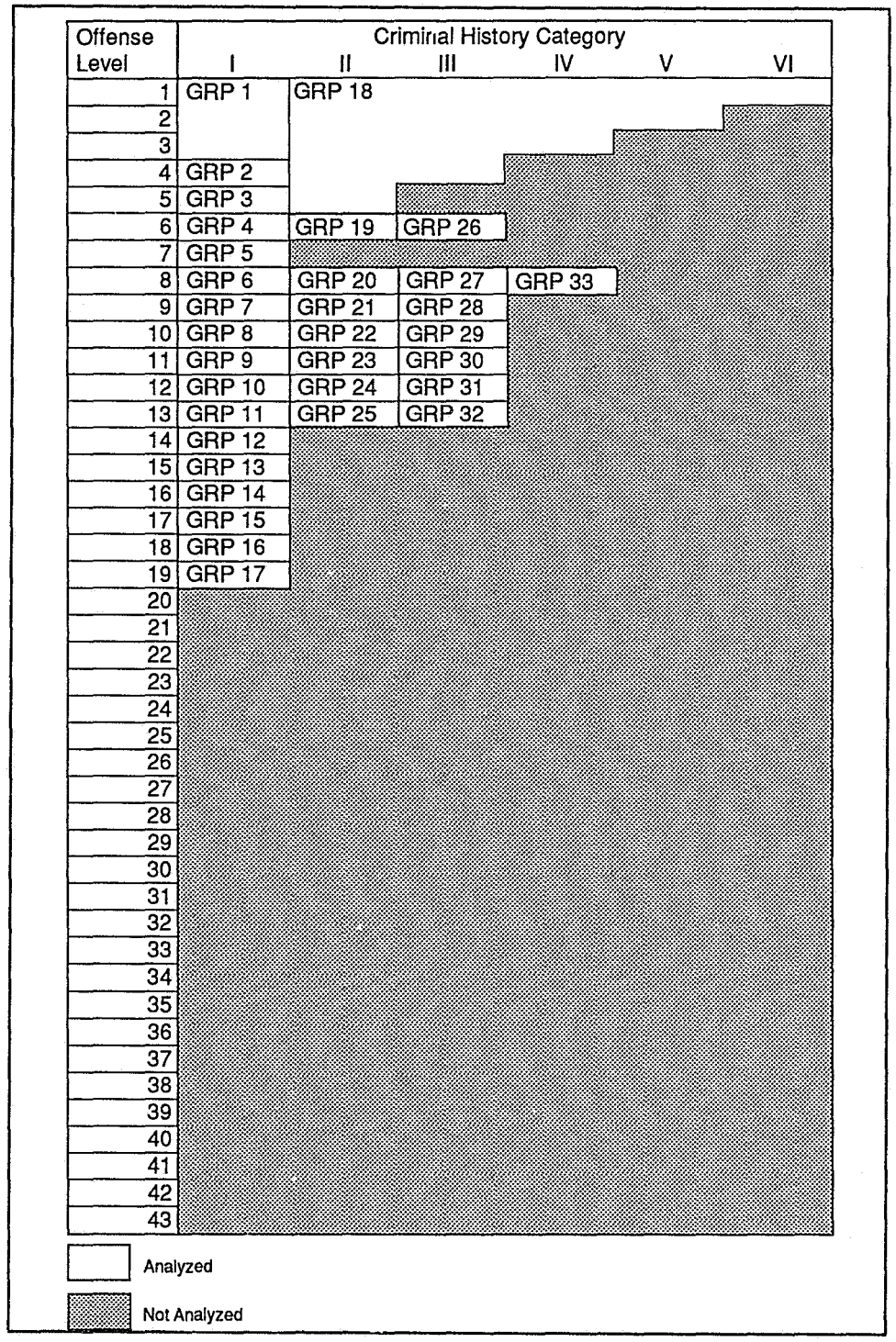
Figure I.2: Analysis Groups GAO Used for Drug Offenders



Theft and Fraud Offenders

In our analysis of theft and fraud offenses, we included the five following general offense types: (1) larceny, (2) embezzlement, (3) fraud, (4) auto theft, and (5) forgery and counterfeiting. There was a total of 2,968 preguidelines records (9,692 weighted) and 5,966 guidelines records. The preguidelines and guidelines theft and fraud data files were divided into 33 groups. The 33 groups contained 2,295 preguidelines records (8,017 weighted) and 5,043 guidelines records. Together these 33 groups accounted for 77 percent and 85 percent of the total weighted preguidelines and guidelines records, respectively. As with the drug offenses, a lack of data prevented us from analyzing large sections of the sentencing table. As figure I.3 shows, all offenders in offense levels below 19 were not analyzed.

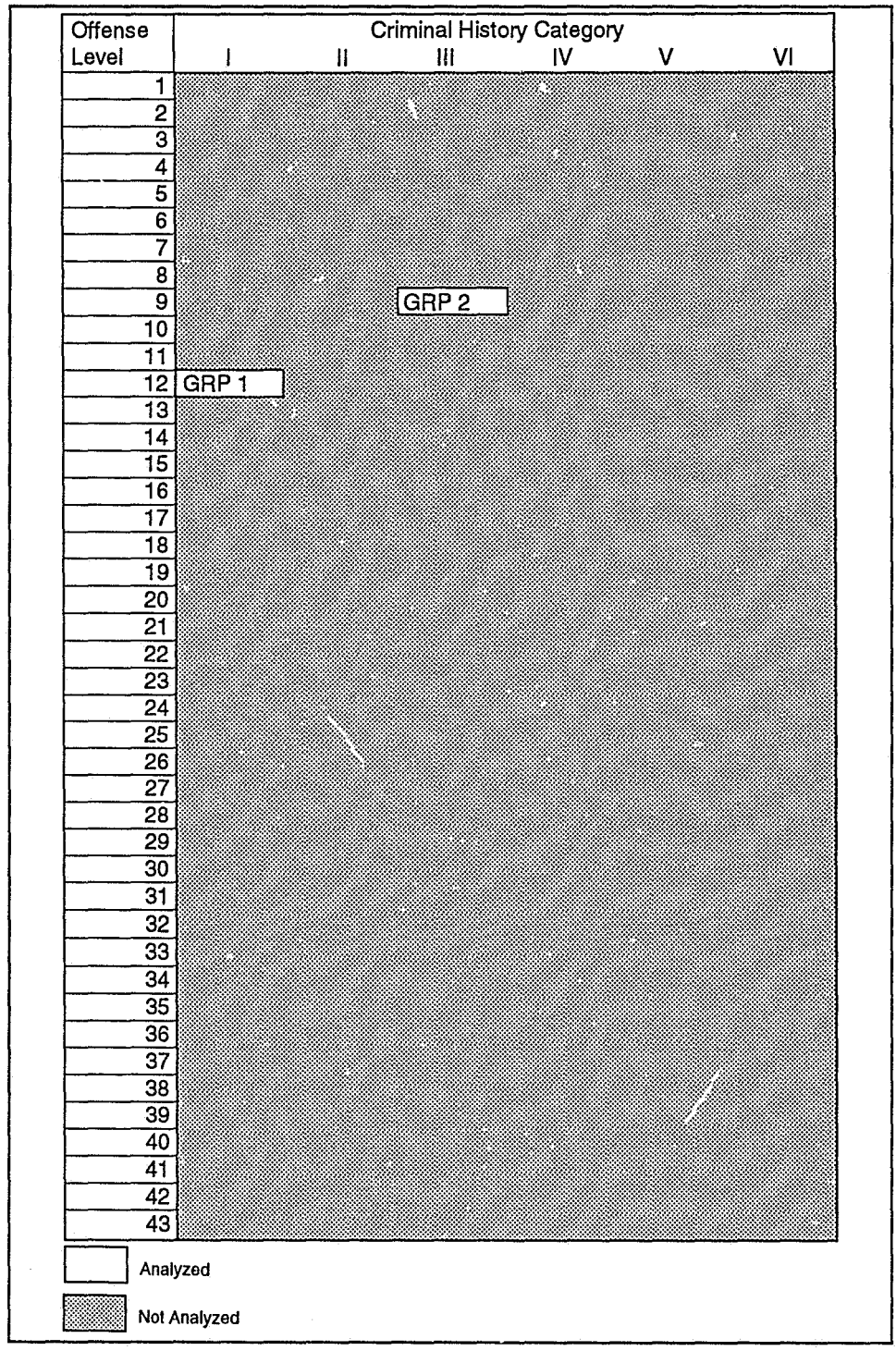
Figure I.3: Analysis Groups GAO Used for Theft and Fraud Offenders



Firearms Offenders

Last, we analyzed firearm offenses. We included five general offense types in the analysis: (1) receipt, possession, or transportation; (2) trafficking; (3) receiving, transporting, shipping, or transferring with intent to commit another crime, or knowledge that the firearm will be used in committing another crime; (4) use during or in relation to certain crimes; and (5) possession in a federal facility. Our analysis contained a total of 384 preguidelines records (1,236 weighted) and 1,409 guidelines firearms records. Only two cells in both the preguidelines and guidelines data files contained sufficient data to perform our analysis. The two groups analyzed contained 66 preguidelines records (251 weighted) and 60 guidelines firearm records. These two groups accounted for 17 percent and 4 percent of the total weighted preguideline and guideline records, respectively. Figure I.4 shows the two groups.

Figure I.4: Analysis Groups GAO Used for Firearms Offenders



Methodology Used

Our statistical analysis of the dispersion of estimated time to be served involved determining whether the variance of this quantity increased or decreased under the guidelines. This involved testing for variance homogeneity. Usually, this is done by calculating the ratio of the sample variances and looking up the value for this ratio in a table of the F-distribution. However, this procedure relies heavily on the assumption that the data from which the variances have been computed follow a normal distribution. We found that the data on estimated time to be served generally did not meet this requirement.

To test for variance homogeneity under these conditions, we used the bootstrap resampling technique.[8] The bootstrap method does not rely on assumptions about the underlying distribution, such as normality. It involves sampling with replacement from the original data and computing the value of the parameter or test statistic of interest, such as the variance or the variance ratio.[9] The resampling procedure is repeated many times, and it is assumed that the resulting sample distribution may be used as a valid estimate of the underlying population distribution. We generated 1,000 bootstrap samples for each of the preguidelines and guidelines groups, and we calculated the variances of those groups. We pooled the data, and using the preguidelines variance as the numerator of a variance ratio and the guidelines variance as the denominator, we also estimated bootstrapped variance ratios. We then conducted appropriate tests for inequality of preguidelines and guidelines variances.

Results of GAO's Analyses

The overall analysis, as well as the analyses of drug and theft and fraud offenses generally showed that variances in time served decreased after the guidelines took effect. The number of firearms groups was too small to make a conclusion about the results of our analysis.[10] Of the 68 groups

[8]The bootstrap resampling technique has been widely used to estimate parameters and confidence intervals. For an explanation and overview of applications, see B. Efron and B. Tibshirani, "The Bootstrap Method for Assessing Statistical Accuracy," Division of Biostatistics, Technical Report No. 101 (Stanford, CA: Stanford University, Mar. 1985) or R. Stine, "An Introduction to Bootstrap Methods: Examples and Ideas" Sociological Methods and Research, Vol. 18, Nos. 2 & 3 (Nov. 1989/Feb. 1990), pp. 243-291.

[9]The bootstrap method we used was similar to that used by the Commission in its distributional analysis. The major difference in our approaches was in how offender data were grouped and the resulting impact that different grouping had in the interpretation of results. For a discussion of methods for variance homogeneity, see Denis D. Boos and Cavell Brownie, "Bootstrap Methods for Testing Homogeneity of Variances," Technometries, Vol. 31, No. 1 (Feb. 1989), pp. 69-82.

[10]Our analysis of two firearms groups showed that one group had a smaller variance under the guidelines and the other group had a larger variance.

analyzed in the overall analysis, 57 groups had smaller variances under the guidelines, and 11 had larger.

The results of our analysis of drug offenders showed that, of the 25 groups analyzed, 20 groups had smaller variances under the guidelines and 5 had larger. Only 1 of the 33 theft and fraud groups analyzed had a larger variance under the guidelines.

After the analyses of the variances were completed, we further analyzed the groups in each of the four analyses that showed a higher variance after the guidelines took effect. We also looked at any group with a lower variance that was positioned on the sentencing table above or below the group with the higher variance or was positioned diagonally upward to the right or downward to the left of a higher variance group.[11] For each of these groups we calculated the coefficient of variation,[12] a relative measure of the variance. We found that in all of the higher and lower variance groups analyzed the relative measure of the variance showed that sentencing dispersion was less under the guidelines. Tables I.4 and I.5 show the results of our relative variance analysis.

[11]We chose to look at the groups diagonally to the upper right and lower left because these groups generally had the same or similar sentence ranges as the associated higher variance groups.

[12]The coefficient of variation is equal to the standard deviation divided by the mean then multiplied by 100 percent.

Table I.4: Results of the Analysis of Coefficient of Variance of Estimated Time to Be Served—All Groups Analyzed

Analysis group	Preguidelines			Guidelines		
All offenses	AVG	VAR	C.V.	AVG	VAR	C.V.
9	1.7	105.8	574.6	5.3	95.4	182.6
14	9.3	669.0	279.7	16.3	199.0	84.6
16	16.4	850.4	178.2	24.9	463.3	85.7
23	22.3	818.2	128.5	46.2	652.5	54.8
24	17.0	1,226.4	210.7	61.5	2,313.0	71.4
25	18.0	910.0	168.5	54.7	690.4	47.6
27	22.2	1,553.5	179.6	71.4	817.8	39.9
33	79.3	3,758.8	76.7	304.4	9,535.3	32.1
35	4.8	166.5	270.2	5.2	18.1	81.8
36	5.4	171.0	246.6	9.0	258.6	153.6
37	6.5	233.1	254.2	9.5	133.0	112.1
40	11.9	563.3	200.6	16.2	145.6	69.1
41	11.9	379.4	179.9	19.6	464.4	101.3
42	19.6	783.2	141.2	24.6	1,058.9	108.4
50	9.8	247.1	157.1	7.4	18.4	58.4
51	8.6	354.8	223.0	13.1	605.8	164.6
52	7.6	365.7	250.6	12.9	150.0	92.2
53	6.4	228.3	263.2	12.5	19.5	34.9
54	15.2	341.9	121.3	20.3	471.7	103.5
55	13.5	351.9	139.0	17.6	34.1	33.0
57	18.4	516.4	124.3	30.6	743.6	87.3
60	42.9	943.8	71.9	51.0	1,016.3	59.9
65	15.0	485.0	144.7	12.3	111.4	74.5
66	10.9	479.8	201.6	19.9	867.3	136.2
67	24.0	333.1	76.4	54.8	730.4	48.6

Table I.5: Results of the Analysis of Coefficient of Variance of Estimated Time to Be Served—Drugs and Theft and Fraud

	Preguidelines			Guidelines		
Analysis group	AVG	VAR	C.V.	AVG	VAR	C.V.
Drugs						
8	15.1	300.2	116.1	45.1	329.0	40.2
11	24.7	841.7	119.9	77.8	1,253.9	45.5
16	75.8	8,170.1	117.5	299.2	9,733.8	33.0
18	31.5	317.6	56.6	62.1	524.1	36.4
24	42.0	1,232.7	82.2	146.0	3,841.1	42.4
Theft and fraud						
7	1.0	30.7	570.2	1.7	8.9	178.5
8	.8	17.4	563.2	3.3	19.6	135.3
9	1.8	49.8	406.4	6.8	21.7	68.3
21	2.3	28.0	228.5	5.6	19.7	79.2

Additional Analysis of Sentence Dispersion

In a separate analysis, we reanalyzed the Commission's data on the sentences of preguidelines and guidelines bank robbers, bank embezzlers, heroin distributors, and cocaine distributors. We used the data the Commission used in their "Distributional Analysis," but we used different techniques to analyze the data. Our purpose was to determine if changing techniques would change the Commission's results.

Briefly, in its analysis of the distribution of sentences for groups of similarly situated offenders convicted and sentenced for these four crime types, the Commission concluded that overall disparity in the prison sentences imposed and time in prison had declined under the guidelines. (See pp. 277-299 of the Commission's study.) In addition, in its study of sentencing patterns under the guidelines (pp. 300-339 of the study), the Commission also concluded that there was little evidence of differences in the sentences or time served of similar offenders of different races, genders, or other demographic categories.

In our reanalysis of the Commission's data, we limited ourselves to the following. First, we tested for change in disparity among all heroin and cocaine distributors with criminal history category I. We did not limit ourselves to offenders within a single offense level.[13] Second, we pooled the data for cocaine and heroin offenders to increase the statistical power that was available to detect changes in disparity.

[13]We used analysis of variance to control statistically for the effect of offense severity.

Further, we reanalyzed the data on drug offenders belonging to specific demographic categories by pooling the data for cocaine and heroin offenders of offense level 26 in order to increase statistical power, and we transformed the data using logarithms in order to improve the assumptions of the statistical tests.

Our results for the first analysis are displayed in tables I.6 and I.7. Our results of the analysis of demographic categories are displayed in table I.8. Our results support the Commission's original conclusions: overall disparity declined, and the data provided little evidence of disparate treatment by race, gender, or other demographic categories.

Table I.6: Standard Deviations in Months of Imposed Sentence and Time Served,[a] Controlling for Guidelines Offense Level.[b] Heroin and Cocaine Category 1 Offenders

Preguidelines and guidelines samples (Commission Data Set 1). Degrees of Freedom (DF) in parentheses.

		Standard deviation[a]				Test of
Measure	Type of offense	Preguidelines		Guidelines		change
Imposed sentence						
	Heroin	2.0	(68)	1.4	(154)	3.6[c]
	Cocaine	1.9	(58)	1.4	(285)	3.3[c]
	Combined	2.0	(140)	1.6	(523)	2.4[c]
Time served						
	Heroin	1.6	(68)	1.5	(155)	1.4
	Cocaine	1.7	(58)	1.4	(308)	2.3[c]
	Combined	1.7	(140)	1.6	(548)	1.3

[a]The disparity measure is the within-level standard deviation based on a fixed effects one-way ANOVA, i.e., the square root of the within-level mean-square. To stabilize the variance across levels, we used the logarithms of imposed sentence and time served as dependent variables. A small number of observations with zero sentences or durations were omitted. The standard deviations are presented on the original scale (months).

[b]The guidelines levels are defined in terms of the weight of seized drugs. See table I.7.

[c]Statistically significant change in disparity, alpha = .05. (Compare to the F distribution with df as indicated for the preguidelines and guidelines samples.)

Table I.7: Guidelines Levels of Drug Amounts

	Weight (Grams)	
Level	**Heroin**	**Cocaine**
12	0-4	0-24
14	5-9	25-49
16	10-19	50-99
18	20-39	100-199
20	40-59	200-299
22	60-79	300-399
24	80-99	400-499
26	100-399	500-1,999
28	400-699	2,000-3,499
30	700-999	3,500-4,999
32	1,000-2,999	5,000-14,999
34	3,000-9,999	15,000-49,999
36	10,000-29,999	50,000-149,999

Note: Each analysis used all levels with two or more observations in each sample.

Table I.8: Tests for Disparity in Mean Sentence Imposed (Months) **and Mean Time Served (Months) Based on Demographic Variables. Combined Heroin and Cocaine Category 1 Offenders, Offense Level 26**

Preguidelines and guidelines samples (Commission Data Set 1). Sample base in parentheses.

		Mean sentence (N)				Mean time (N)			
		Preguidelines		Guidelines		Preguidelines Guid		Guidelines	
Variable	**Subclass**								
Education									
	Low	34	(39)	60	(82)	20	(39)	52	(82)
	High	37	(45)	61	(68)	23	(45)	53	(68)
	Difference	–3		–1		–3		–1	
Married									
	No	32	(46)	60	(87)	21	(46)	52	(87)
	Yes	41	(38)	60	(66)	22	(38)	52	(66)
	Difference	–9		0		–1		0	
Gender									
	Female	23	(25)	58	(14)	16	(25)	50	(14)
	Male	41	(59)	61	(139)	24	(59)	53	(139)
	Difference	–18[a]		–3		–8[a]		–3	
Race[b]									
	White	38	(20)	60	(53)	18	(20)	52	(53)
	Black	35	(28)	60	(47)	23	(28)	52	(47)
	Difference	3		0		–5		0	
Race[c]									
	Non-Hispanic	38	(20)	60	(53)	18	(20)	52	(53)
	Hispanic	33	(34)	61	(47)	22	(34)	53	(47)
	Difference	5		–1		–4		–1	
Employed									
	No	30	(20)	59	(32)	21	(20)	51	(32)
	Yes	38	(64)	60	(112)	22	(64)	53	(112)
	Difference	–8		–1		–1		–2	

[a]Statistically significant difference, alpha = .05. Satterthwaite two-sample t-test (assumes unequal variances).

[b]Hispanics, American Indians, Alaskan natives, and Asian and Pacific Islanders were excluded from this analysis.

[c]Blacks, American Indians, Alaskan natives, and Asian and Pacific Islanders were excluded from this analysis.

Disparity in Sentencing Under the Guidelines

Background

This appendix describes research that attempted to refine and extend the United States Sentencing Commission's analysis as found in its evaluation report.[1] The Commission's study, described in its report as preliminary and the basis for future research efforts, presented information on sentencing in the form of 35 2-way tables. These 35 tables were formed by cross-classifying a 6-category sentence location variable by 7 factors (one at a time), using 5 different samples of offenders.

The five samples of offenders included "similarly situated" categories of (1) bank robbers, (2) embezzlers, (3) heroin traffickers, and (4) cocaine traffickers, as well as (5) a 25-percent sample of all offenders (of all types) sentenced under the guidelines from April 1990 through September 1990. The sentence location variable employed in these tables contrasted offenders who received sentences that were downward departures, at the bottom of the guidelines range, below the midpoint, at or above the midpoint, at the top of the guidelines range, or upward departures. The seven factors by which this sentence location variable was cross-classified were race, gender, marital status, employment status, education, and judicial circuit.

The numbers of similarly situated robbery, embezzlement, heroin, and cocaine cases represented in the Commission's tables were roughly 111, 93, 86, and 161, respectively. The exact number in each table depended on the factor by which sentence location was cross-classified and the amount of missing data associated with it. The number of cases cross-classified when the data from the 25-percent sample of all offenders were used ranged from roughly 1,500 to 2,100, with again the exact number depending on the factor involved in the cross-classification and the amount of missing data associated with it.

With respect to the 28 offense-specific tables that were presented, the Commission found that the race of the offender was significantly related to sentence location for the particular group of heroin traffickers they considered, but not for cocaine traffickers nor for bank robbers or embezzlers. Moreover, the Commission was unable to establish whether any of the other factors (gender, age, etc.) were related to sentence location for any of the four groups of similarly situated offenders they looked at. In summarizing the results for these specific categories of offenders, the Commission concluded that either cell sizes were too small to test or no significant differences in sentence location were found with

[1]"Judicial Sentencing Patterns Under the Guidelines", The Federal Sentencing Guideline, Vol. II, Ch. 4, Section A of Part VI.

respect to gender, age, marital status, employment status, education, and judicial circuit.

In analyzing the seven two-way tables involving the 25-percent sample of all offender types, the Commission reported that the location of the sentence meted out to offenders was significantly related to race, gender, employment status, and judicial circuit but not to age, marital status, or education. Little was made of these aggregate findings except the observations that (1) racial differences in sentence location involved differences across racial groups in the tendency to receive differing within range sentences but not differences in the tendency to receive departure sentences and (2) any of the variations found might be correlated with (and presumably accounted for by) factors not addressed in its preliminary study.

There is little discussion in the Commission's report of the results which suggest that the sentences offenders received under the guidelines may be partly a function of personal characteristics that were clearly irrelevant to sentencing (such as race and gender) or potentially irrelevant to sentencing (such as employment status and judicial circuit). It is true that the reported results involved bivariate relationships which might be diminished substantially, or accounted for totally, by controls for legally relevant characteristics, such as seriousness level and criminal history. It is also true that the reported result ran contrary to the previously reported results which found largely no effects of personal characteristics on sentences handed down to specific categories of bank robbers and embezzlers and heroin and cocaine traffickers. But the Commission offered no evidence that legally relevant factors accounted for these relationships, and the failure to find significant relationships in the samples of specific offender types used might, as the Commission clearly recognized, have resulted from nothing more than the small cell sizes that its selection methods produced.

It is possible that the null results the Commission found with respect to the specific categories of bank robbers, embezzlers, and drug traffickers is not what they would have found had they chosen to look at other categories of robbers, embezzlers, and drug traffickers or at other types of offenders more generally. The particular categories of specific offenders that the Commission analyzed were not chosen because they were typical, or representative, of broader categories of offenders but because they were the largest groups of a very large number of groups of similar offenders with similar criminal histories and similar offense

characteristics that the Commission could obtain given their selection procedures. Whether inferences can be made from these specific offender types to other types of offenders is questionable.

Because of the problem of generalizing from the results involving the small groups of similarly situated offenders and because of the potential importance of the preliminary results which the Commission obtained in their work with the 25-percent sample of all offenders, we used the latter data in the present report. We began by examining the seven two-way tables it created from this data set by cross-classifying sentence location with race, gender, marital status, employment status, education, age, and circuit. We also considered similar two-way tables formed by cross-classifying sentence location by criminal history, seriousness level, type of offense, and mode of disposition (plea or trial). We then looked to see whether controlling for these latter, legally relevant factors, accounted for the effects we found of the former, legally irrelevant factors, on the location of sentences which offenders received. (Each variable used in this analysis is defined later in this section.)

We employed simple loglinear techniques in our analyses and looked separately at the effects of these aforementioned factors on the tendency for offenders to receive departure versus within range sentences and, for those who received within range sentences, at the tendency for offenders to receive sentences at the top of the guidelines range rather than at the bottom of the range or somewhere in between. In looking at the clearly or potentially legally irrelevant variables, we found that only judicial circuit affected whether offenders received sentences that departed from the range prescribed by the guidelines. All of these legally irrelevant variables, however, with the exception of education, affected whether offenders received sentences at the top or the middle or the bottom of the range. None of the effects that we found of these variables, either on the tendency to receive departing sentences or on the tendency to receive shorter or longer within range sentences, appeared to be accounted for in very great measure by simple controls for the legally relevant factors we considered. We provide details regarding the nature and magnitude of these effects in the analysis section of this appendix.

The Analysis

Our analysis began with a consideration of 11 2-way tables formed by cross-classifying sentence location with 7 extralegal factors[2] and 4 legal factors. The sentence location variable employed was the same variable used by the Commission in its report and contrasted sentences that involved downward departure, bottom of the range, below the midpoint, at or above the midpoint, top of the range and upward departure sentences.[3] The 7 extralegal factors were all categorical variables and involved the following contrasts:

1. race (white, black, Hispanic),

2. gender (male, female),

3. marital status (married, unmarried),

4. employment status (unemployed, partially employed, fully employed over the 12 months prior to sentencing),

5. education (less than high school, high school, more than high school),

6. age (18-25, 26-35, 36 and older), and

7. circuit (12 categories, including D.C. and the First through Eleventh Circuits).

The legal factors considered were similarly categorical and involved the following contrasts:

[2]We use the term "extralegal" throughout to denote factors which are clearly irrelevant to guidelines sentencing from a legal standpoint or which are potentially irrelevant. The guidelines seem to us to be framed in such a way as to make it clear that race, sex, and social class should not play a part in sentencing under the guidelines and that, under most conditions, marital status, employment status, education, age, and personal characteristics of this sort should not be pertinent as well. In the Commission's 25-percent departure study sample with which we were working, there was no clear measure of social class, though education and employment status might be regarded as crude proxies for that variable.

[3]While the Commission's report was not entirely clear about how this variable was created or what these categories correspond to exactly, conversations with analysts at the Commission suggest the following. Upward departures were any sentences of greater length than the guidelines range entails, while downward departures involved sentences of shorter length for reasons other than a defendant's substantial assistance to the government. (Cases involving downward departures for substantial assistance were deleted from the file.) Top of the range and bottom of the range sentences were precisely what they imply. If the guidelines range for a particular offender, given the seriousness of the offense committed and the offender's criminal history, was 41-51 months, then a sentence of 41 months would be a bottom of the range sentence and a sentence of 51 months would be a top of the range sentence. The midpoint of this range (46 months) would be the point that separates sentences below the midpoint from sentences at or above the midpoint.

1. criminal history (guidelines table criminal history categories I through VI),

2. seriousness level (final offense level from the guidelines table, ranging from 1 through 43),

3. offense type (violent, economic, drug, firearms, immigration, other),[4] and

4. mode of disposition (guilty plea, trial).

Our analysis of these tables proceeded by first partitioning them into two complementary subtables which revealed, more directly than the expanded cross-classification, the separate effects of these factors on the tendency for offenders to receive departing versus within range sentences and, for those who received the latter, sentences that were at the top of the range rather than at the bottom or somewhere in between. Table II.1, for example, shows the expanded cross-classification of race and sentence location, while tables II.2 and II.3 show the subtables formed from this cross-classification by partitioning.

Table II.1: Frequencies and Percentages in the Cross-Classification of Race and Sentence Location

Percentages in brackets

	Sentence location						
Race	**Downward departure**	**Bottom of the range**	**Below the midpoint**	**Above the midpoint**	**Top of the range**	**Upward departure**	**Total**
White	83	484	99	150	135	32	**983**
	[8.4]	[49.2]	[10.1]	[15.3]	[13.7]	[3.3]	**[100.0]**
Black	46	283	52	59	79	11	**530**
	[8.7]	[53.4]	[9.8]	[11.1]	[14.9]	[2.1]	**[100.0]**
Hispanic	34	149	51	55	41	7	**337**
	[10.1]	[44.2]	[15.1]	[16.3]	[12.2]	[2.1]	**[100.0]**
Total	**163**	**916**	**202**	**264**	**255**	**50**	**1,850**
	[8.8]	[49.5]	[10.9]	[14.3]	[13.8]	[2.7]	**[100.0]**

[4]Our categorization of the aggregated categories of violent, economic, and drug offenders paralleled the scheme used by the Commission elsewhere in its report. Violent offenses included homicide, robbery, kidnapping and assault. Economic offenses included larceny, embezzlement, tax offenses, fraud, and forgery and counterfeiting. Drug offenses involved all drug-related offenses excluding simple possession. The "other" category included all offenses other than these and firearms and immigration offenses, which were large enough categories of offenders to retain without aggregation.

Table II.2: Race by Within Range Sentences

	Sentence location				
Race	Bottom of the range	Below the midpoint	Above the midpoint	Top of the range	Total
White	484	99	150	135	**868**
	[55.8]	[11.4]	[17.3]	[15.6]	**[100.0]**
Black	283	52	59	79	**473**
	[59.8]	[11.0]	[12.5]	[16.7]	**[100.0]**
Hispanic	149	51	55	41	**296**
	[50.3]	[17.2]	[18.6]	[13.9]	**[100.0]**
Total	**916**	**202**	**264**	**255**	**1,637**
	[56.0]	[12.3]	[16.1]	[15.6]	**[100.0]**

Table II.3: Race by Departure Vs. Within Range Sentences

	Sentence location			
Race	Downward departure	Within range	Upward departure	Total
White	83	868	32	**983**
	[8.4]	[88.3]	[3.3]	**[100.0]**
Black	46	473	11	**530**
	[8.7]	[89.2]	[2.1]	**[100.0]**
Hispanic	34	296	7	**337**
	[10.1]	[87.8]	[2.1]	**[100.0]**
Total	**163**	**1,637**	**50**	**1,850**
	[8.8]	[88.5]	[2.7]	**[100.0]**

Partitioning in this fashion allowed us to test independently hypotheses that assert that these various factors are unrelated to whether offenders received departing sentences versus within range sentences on the one hand and longer versus shorter within range sentences on the other. With respect to the racial factor, for example, the chi-square values associated with the model of independence given in the first row of table II.4 suggest very clearly that racial groups did not differ with respect to the tendency to receive departing versus within range sentences but did differ significantly with respect to the tendency to receive certain categories of within range sentences.[5] That is, we fail to reject the hypothesis that race

[5]The model of independence being tested here and elsewhere in this appendix is a model which asserts that in the population from which this sample was drawn the variables which are cross-classified in the table to which the model was applied were independent of, or unrelated to, one another. The value of chi-square associated with the model of independence informs us as to how greatly the observed frequencies in the table depart from the expected frequencies under this model. Large and improbable values of chi-square, relative to the number of degrees of freedom (df) associated with it, indicate that the model of independence can be rejected, while small and probable values indicate that the model cannot be rejected.

was independent of the tendency to receive departing versus within range sentences, given the likelihood ratio chi-square (L^2) value of 3.23, which with 4 degrees of freedom (df) is a highly probable value (P = .52) under the model of independence for that subtable. But we can easily reject the hypothesis that race was independent of the tendency to receive different categories of within range sentences, given the L^2 of 16.39 with 6 df, P = .01, associated with the model of independence for that subtable.[6]

Similar partitionings of the other expanded two-way tables reveal that, among the seven extralegal factors considered, only circuit affected whether offenders received departure versus within range sentences (L^2 = 35.53 with 22 df, P = .03), while gender, employment status, and circuit, in addition to race, affected whether offenders received varying within range sentences. It turns out, as will be shown below, that marital status and age also affected the type of within range sentences that offenders received, though before we report findings pertaining to that outcome we develop more fully our understanding of these differences across circuits in departure sentences.

Effects of Factors on Departure Sentences

Table II.4 shows that, in addition to circuit, each of the four legal factors (criminal history, seriousness level, offense type, and mode of disposition) affected whether offenders received departure versus within range sentences. We would like to know, of course, whether the significant differences across circuits in the tendency to receive departure versus within range sentences resulted from differences across circuits in these legal variables. Before investigating that issue, we first attempted to provide a description of the effects of these different variables, considered separately. This required additional partitionings of the two-way subtables involving circuit, criminal history, and seriousness level and the fitting of models other than the simple model of independence to all of these two-way subtables in which independence was rejected.

[6]We note that the sum of the L^2 values for the two subtables formed by partitioning equals the L^2 value for the expanded cross-classification, or full table (e.g., 3.23 + 16.39 = 19.62, etc.). In that sense, the partitioned subtables we worked with retain all of the information about the association between sentence location and each factor that was contained in the original two-way tables.

Table II.4: Chi-Square Values for the Full and Partitioned Two-Way Tables Involving Sentence Location and Selected Factors

	Full table			Departure vs. within range subtable			Within range subtable		
Factor	df	Chi-square	p	df	Chi-square	p	df	Chi-square	p
Race	10	19.62	.03	4	3.23	.52	6	16.39	.01
Gender	5	47.40	<.001	2	3.90	.14	3	43.49	<.001
Marital status	5	9.86	.08	2	3.57	.17	3	6.29	.10
Employment status	10	33.99	<.001	4	4.14	.39	6	29.85	<.001
Education	10	9.10	.52	4	3.56	.47	6	5.53	.48
Age	10	11.65	.31	4	1.17	.88	6	10.48	.11
Circuit	55	86.92	<.01	22	35.53	.03	33	51.39	.02
Criminal history	25	146.11	<.001	10	20.70	.02	15	125.40	<.001
Seriousness level	205	473.61	<.001	82	178.12	<.001	117	295.49	<.001
Type of offense	25	126.74	<.001	10	71.13	<.001	15	55.61	<.001
Mode of disposition	5	14.16	.01	2	6.96	.03	3	7.20	.07

Table II.5 shows that the differences across categories of judicial circuit, offender's criminal history, and seriousness level could be simply described, without loss of significant information, by collapsing (or aggregating) certain categories of those variables. With respect to circuit, for example, we found in moving from step 1 to step 2 that the Third, Fourth, Seventh, Tenth, and Eleventh Circuits were not significantly different from one another in terms of the tendency for offenders to receive departing versus within range sentences. This is established by the difference between L^2 values for independence models fitted to the 12 X 3 table in which all circuits are contrasted and to the 8 X 3 table in which those five circuits are combined and contrasted with the other seven circuits. This "difference" L^2, which equals 35.53 - 31.77 = 3.76 with 22 - 14 = 8 degrees of freedom, is effectively the L^2 testing independence in the 5 X 3 table in which these five circuits are contrasted with one another. The high probability associated with this chi-square value ($P > .50$) indicates that independence cannot be rejected in that table, which is to say that there were no significant differences across those five circuits, and no good reason, statistically speaking, not to combine them.

Table II.5: Additional Partitionings of the Two-Way Departure Vs. Within Range Subtables Involving Circuit, Criminal History, and Seriousness Level

Factor	Step	Categories combined on step	Independence model			Difference from previous step		
			df	Chi-square	p	df	Chi-square	p
Circuit	1	None	22	35.53	.03			
	2	[3,4,7,10,11]	14	31.77	<.01	8	3.76	>.75
	3	[2,5,6,8]	8	30.15	<.001	6	1.62	>.95
	4	[1,9]	6	29.43	<.001	2	0.72	>.50
Criminal history	1	None	10	20.70	.02			
	2	[1,2]	8	19.98	.01	2	0.72	>.50
	3	[3,4]	6	18.63	<.01	2	1.35	>.50
	4	[5,6]	4	12.72	<.05	2	5.91	>.05
Seriousness level	1	None	82	178.12	<.001			
	2	[1-8]	68	156.30	<.001	14	21.82	>.05
	3	[29-43]	42	126.95	<.001	26	29.35	>.25
	4	[9-18]	24	108.27	<.001	18	18.68	>.25
	5	[19-28]	6	83.29	<.001	18	24.98	>.10

Moving from step 2 to step 3, we found in similar fashion that the Second, Fifth, Sixth, and Eighth Circuits could be combined with one another, and in moving from step 3 to step 4 we further found that the First and Ninth Circuits were alike with respect to departure versus within range sentences and could be similarly combined.[7] The outcome of these four steps was the creation of a statistically justifiable simplification of the circuit variable that contrasted these three aggregated circuits and the D.C. Circuit, which was unlike any of the others with respect to this particular outcome and could not be combined with any of them.

In parallel fashion, we found that criminal history could be reduced to a three-category variable which contrasted those in categories I-II, III-IV and V-VI, without significant loss of information about the tendency for offenders to receive departing versus within range sentences. The 43 levels of seriousness level could be similarly reduced without significant loss of information on this outcome, to four levels which distinguish those at levels 1-8, 9-18, 19-28 and 29-43. For these factors, like circuit, the values of the difference chi-squares given in table II.4 attest to the fact that the particular categories combined did not differ significantly from one

[7]Here too it was the difference L^2s which statistically justified these collapsings. The high probabilities associated with the L^2 values of 1.62 and 0.73 with 6 and 2 df, respectively, indicated that the circuits we combined were not significantly different with respect to the tendency for offenders in them to receive departing versus within range sentences.

another in terms of the tendency for offenders within those categories to receive downward departures, within range sentences, or upward departures.

With these three variables collapsed in this fashion, we were in a better position to describe their effects, as well as the effects of type of offense and mode of disposition, on the likelihood that offenders received departing versus within range sentences. We did this by fitting and comparing reasonably simple loglinear models to these two-way tables, models that placed different simplifying constraints on the associations between these factors and this dependent variable. These models are given in table II.6.

Table II.6: Alternative Models for the Two-Way Departure Vs. Within Range Subtables Involving Circuit, Criminal History, Seriousness Level, Type of Offense, and Mode of Disposition

Factor	Model	Marginals/effects fitted	df	Chi-square	p
Circuit	1	[C] [L]	6	29.43	<.001
	2	[CL']	3	27.57	<.001
	3	[CL_2]	3	2.26	.324
Criminal history	1	[H] [L]	4	12.73	.013
	2	[H'L']	3	3.73	.292
Seriousness level	1	[S] [L]	6	83.29	<.001
	2	[SL']	3	44.89	<.001
	3	[S'L]	4	28.52	<.001
	4	[SL_1]	3	14.72	.002
	5	[SL_2]	3	17.00	.001
	6	[SL_3]	3	69.52	<.001
Type of offense	1	[T] [L]	10	71.13	<.001
	2	[TL']	5	21.96	.001
	3	[TL_1]	5	26.39	<.001
	4	[TL_2]	5	37.15	<.001
	5	[TL_3]	5	43.96	<.001
Mode of disposition	1	[D] [L]	2	6.96	.031
	2	[DL']	1	1.65	.198

Legend: C=Circuit, H=Criminal history, S=Seriousness level, T=Type of offense, D=Mode of disposition, L=Sentence location.

Note: Apostrophes indicate linear constraints; L_1, L_2, and L_3 indicate associations with sentence location that involve the first (downward), second (within), and third (upward) categories of location.

With respect to the 4 X 3 table in which our collapsed circuit categories are cross-classified by the three sentencing outcomes (downward departure, within range, upward departure) being compared, we fitted and contrasted three simple models. The first of these was the simple model of independence which we knew did not fit the data in the table acceptably, but which served as a useful baseline model to which we could compare the other two. Model 2 was a model that allowed circuit and this sentence location variable to be associated and assumed (or required) that location was linear in its relationship with circuit. That is, it constrained the differences between circuits in the tendency for offenders to receive downward departure sentences rather than within range sentences to be the same as the differences between circuits in the tendency for offenders to receive within range sentences rather than upward departures. (The location variable, in other words, was treated as an ordered variable in which the within range category resided equidistant between the downward and upward departure categories.) Model 3, by comparison, asserted that the association between circuit and these sentence location categories involved a difference across circuits in the tendency to receive within range sentences versus departing sentences (either upward or downward) but no differences across circuits in the tendency to receive downward rather than upward departures. That Model 3 unlike Model 2 both improved significantly upon Model 1[8] and fit the data acceptably[9] allowed us to choose Model 3 unequivocally as the preferred model to describe the association in that table.

In a similar fashion, the relative fit of models fitted to the other two-way tables that is established by the chi-square values given in table II.6 allowed us to choose a linear by linear association model (Model 2) to describe the association in the criminal history by location table and a linear model (Model 2) to describe the association in the disposition by location table.[10] For the tables in which this sentence location variable was cross-classified by seriousness level and type of offense, none of the models which placed simplifying constraints on the association present in

[8]The significant improvement of Model 3 over Model 1 is established by the significant difference in L^2 values between the two (i.e., 29.42 - 2.26 = 27.16 with 6 - 3 = 3 df, P <.001).

[9]The acceptable fit of Model 3 to the data is established by the chi-square value of 2.26 with 3 df associated with that model, which represents a goodness of fit statistic. The probability of .324 associated with that value implies that the difference between the expected frequencies under this model and the observed frequencies in the table can reasonably be assumed to be due to chance or to sampling fluctuations.

[10]An excellent discussion of the types of constrained models we fit to these tables involving linear associations and linear by linear associations is found in Leo A. Goodman's The Analysis of Cross-Classifications Having Ordered Categories (Cambridge: Harvard University Press, 1984).

the tables fit the data acceptably, so our only recourse for those tables was to choose the trivial model[11] as preferred.

We used these preferred models to estimate the associations between the variables in the tables to which they were fitted. This was done by taking the expected frequencies under the different preferred models for the various tables, which are given in table II.7, and calculating from them various odds and odds ratios, the latter of which served as estimates of the associations of the different factors with this location variable.

[11]The trivial model for tables of this sort is one which has expected frequencies that equal the observed. It fits the data not only acceptably, but exactly. While not shown among the various models given in table II.6, it has an L^2 value of 0.00 with 0 df.

Table II.7: Expected Frequencies Under the Preferred Models for the Two-Way Departure Vs. Within Range Subtables, and Odds and Odds Ratios Derived From Them

	Sentence location						
Factor	Downward departure	Within range	Upward departure	Odds on downward vs. within	Odds ratios	Odds on upward vs. within	Odds ratios
Circuit							
D.C.	0.00	30.00	0.00	0	[a]	0	[a]
3,4,7,10,11	52.64	749.00	17.36	.0703		.0232	
2,5,6,8	77.46	728.00	25.54	.1064	1.51	.0351	1.51
1,9	54.90	354.00	18.10	.1551	2.21	.0511	2.21
Criminal history							
I-II	142.90	1,304.49	36.61	.1095		.0281	.54
III-IV	29.35	364.72	13.93	.0805	.73	.0382	.73
V-VI	10.75	181.80	9.46	.0591	.54	.0520	
Seriousness level							
1-8	12	563	9	.0213		.0160	
9-18	69	683	32	.1010	4.74	.0469	2.93
19-28	58	371	17	.1563	7.34	.0458	2.86
29-43	44	237	3	.1856	8.71	.0127	.79
Type of offense							
Violent	20	93	7	.2151	5.16	.0753	3.29
Economic	20	480	11	.0417		.0229	
Drug	100	709	11	.1410	3.38	.0155	.68
Firearms	6	108	12	.0556	1.33	.1111	4.85
Immigration	4	100	6	.0400	.96	.0600	2.62
Other	23	208	7	.1106	2.65	.0337	1.47
Mode of disposition							
Plea	162.24	1,534.52	46.24	.1057	1.52	.0301	
Trial	20.76	299.48	13.76	.0693		.0459	1.52

[a]Ratios cannot be calculated for D.C.

Note: See footnote 13 for the derivation of odds ratios and figures II.1 through II.5 for plots of the odds on downward departures versus within range sentences and upward departures versus within range sentences across categories of the various factors.

Taking the circuit by sentence location table first, we calculated the odds on receiving downward departures versus within range sentences and then the odds on receiving upward departures versus within range sentences,

within each of the four circuit groups.[12] We then calculated odds ratios by choosing one circuit group as the reference group and dividing the odds for the other groups by the odds for that one. Here we chose the circuit category which included the Third, Fourth, Seventh, Tenth, and Eleventh Circuits as the reference category and calculated odds ratios that compared the odds for other circuit categories to the odds for that one.[13] In so doing, we found that the odds on receiving downward departure sentences versus within range sentences and the odds on receiving upward departure sentences versus within range sentences, were greater in the Second, Fifth, Sixth, and Eighth Circuits than in the Third, Fourth, Seventh, Tenth, and Eleventh Circuits, by a factor of 1.51 (i.e., .1064/.0703 = .0351/.0232 = 1.51). The odds on receiving departure sentences in the First and Ninth Circuits, again either upward or downward, were greater still. Offenders in those two circuits were more than twice as likely (2.21 times as likely, that is) to receive departing sentences than those in the referent circuit group. Offenders in D.C., obviously, were the least likely to receive departure sentences, and differences between them and other circuits were inestimable given the fact that no offenders in the D.C. Circuit sample received such sentences.

While it is customary to describe associations by calculating odds and odds ratios, as we have just shown, another way to simply convey differences across circuits in the tendency to receive these differing types of sentences is to take the odds that are calculated for each circuit and plot them on a logarithmic scale.[14]

In figure II.1, the odds on receiving downward and upward departures versus within range sentences were plotted across the three circuit groups

[12]Odds were calculated by simply dividing the expected frequency in the one category by the expected frequency in the other. For example, in D.C. the odds on receiving downward departures versus within range sentences, and upward departures versus within range sentences were both 0/30.00 = 0. In the First and Ninth Circuits the odds on receiving downward departures versus within range sentences were 54.90/354.00 = 0.1551, while the odds on receiving upward departures versus within range sentences were 18.10/354.00 = 0.0511. These odds, while somewhat different from the more customarily calculated proportions (P), are nonetheless directly related to those proportions (Odds = P/(1-P)), and can be given a fairly straightforward interpretation. The odds of .1551 calculated above can be interpreted as meaning that for every one person in the First and Ninth Circuits who received a within range sentence, .16 persons received a downward departure or, more sensibly, for every 100 persons who received a within range sentence, 16 received downward departures.

[13]In this table, and later in table II.13, odds ratios are given in the row of the table which corresponds to the factor category that serves as the numerator in calculating the ratio. The denominator for all ratios is the factor category contained in the row of the table in which no odds ratio is given. For the circuit by sentence location table, 1.51 = .1064/.0703 = .0351/.0232, and 2.21 = .1551/.0703 = .0511/.0232.

[14]Because odds, unlike percentages, are compared by division rather than subtraction, the appropriate way to plot them is on a multiplicative scale. Such a scale equates the difference between values of 1.0 and 2.0 with the difference between 2.0 and 4.0 (rather than 2.0 and 3.0), since both of these differences would involve a doubling in the odds or an increase in the odds by a factor of 2.0.

(the D.C. circuit, with odds equaling zero, could not be plotted), and the differences across circuits are apparent.[15] In all circuits, the odds on receiving downward departures were greater than the odds on receiving upward departures, but differences across these circuit groups in the tendency for offenders to receive downward departures mirrored differences in the tendency to receive upward departures. Offenders in Second, Fifth, Sixth and Eighth Circuits, had higher odds than offenders in Third, Fourth, Seventh, Tenth and Eleventh Circuits, on receiving a departing sentence, and offenders in First and Ninth Circuits had even higher odds on receiving departing sentences.

The associations between the four legally relevant variables we considered and sentence location can be similarly described by the odds and odds ratios given in table II.7 and depicted graphically by plotting odds on a logarithmic scale, as is done in figures II.2 through II.5. Those calculations and figures reveal the following:

[15]While figures of the sort we present here may be somewhat unfamiliar to many readers, the information they convey can be interpreted in much the same way as plotted regression lines. The slopes of the "lines" in these figures indicate how much the odds on receiving certain sentences rather than others increase or decrease as we move from one category of the various factors to another.

Figure II.1: Expected Odds on Downward and Upward Departures Vs. Within Range Sentences, by Circuit

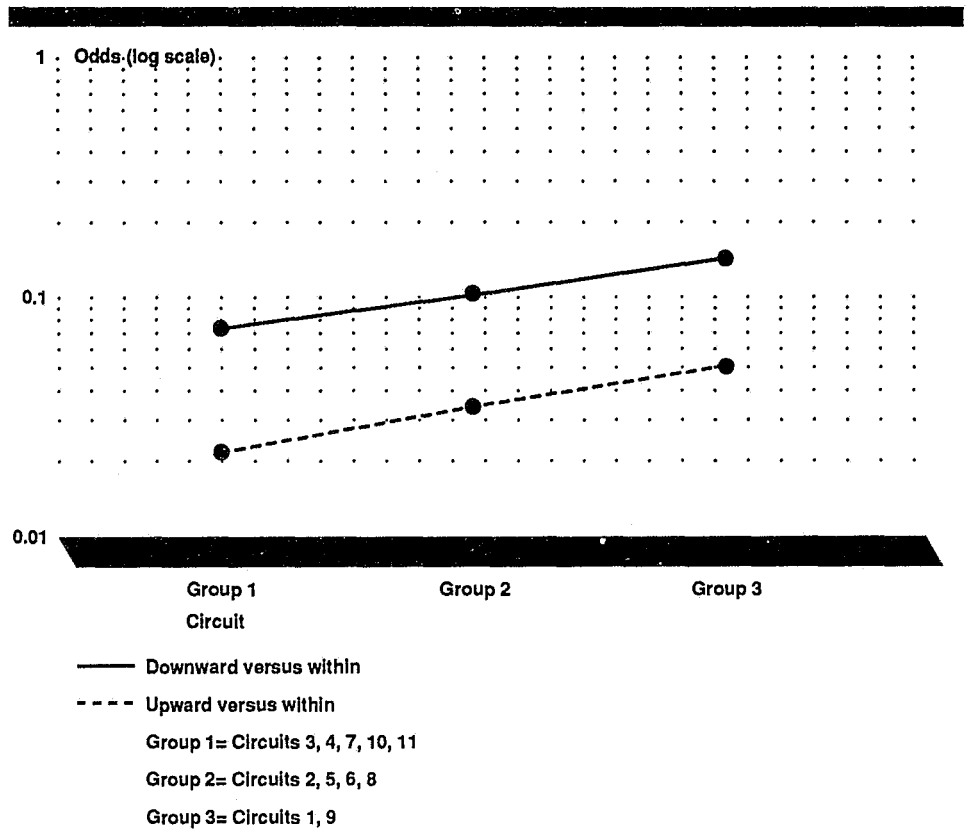


Figure II.2: Expected Odds on Downward and Upward Departures Vs. Within Range Sentences, by Criminal History

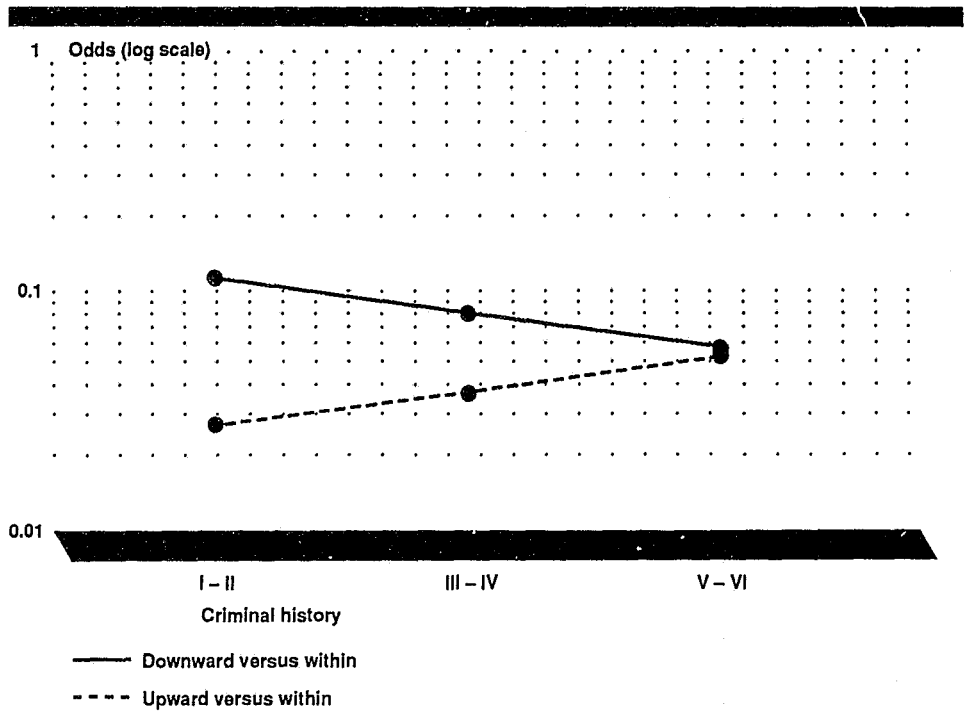


Figure II.3: Expected Odds on Downward and Upward Departures Vs. Within Range Sentences, by Offense Seriousness Level

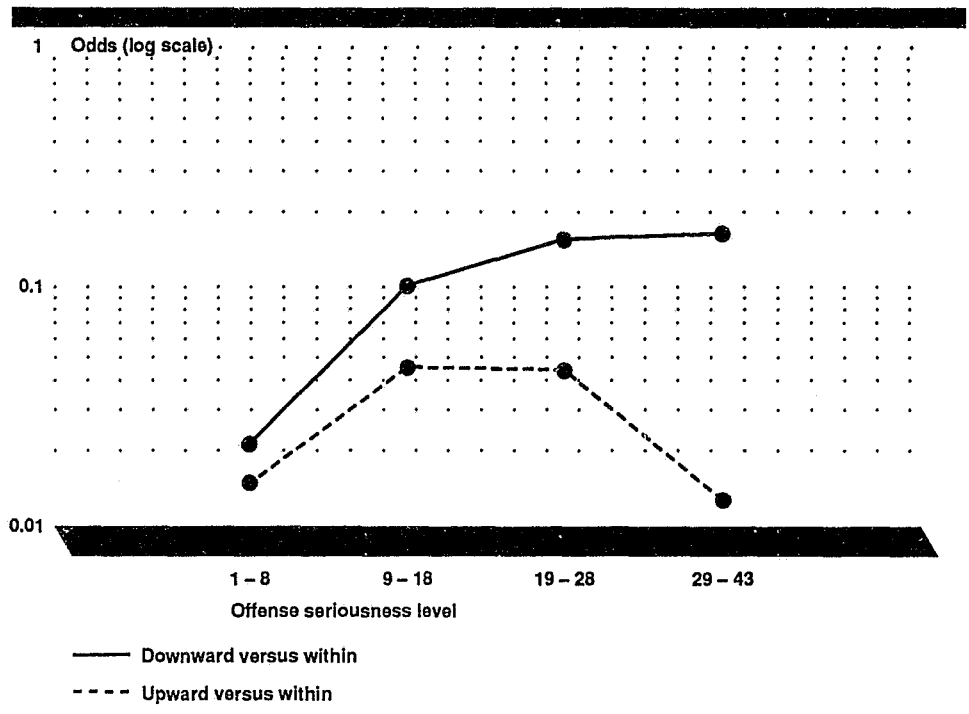


Figure II.4: Expected Odds on Downward and Upward Departures Vs. Within Range Sentences, by Type of Offense

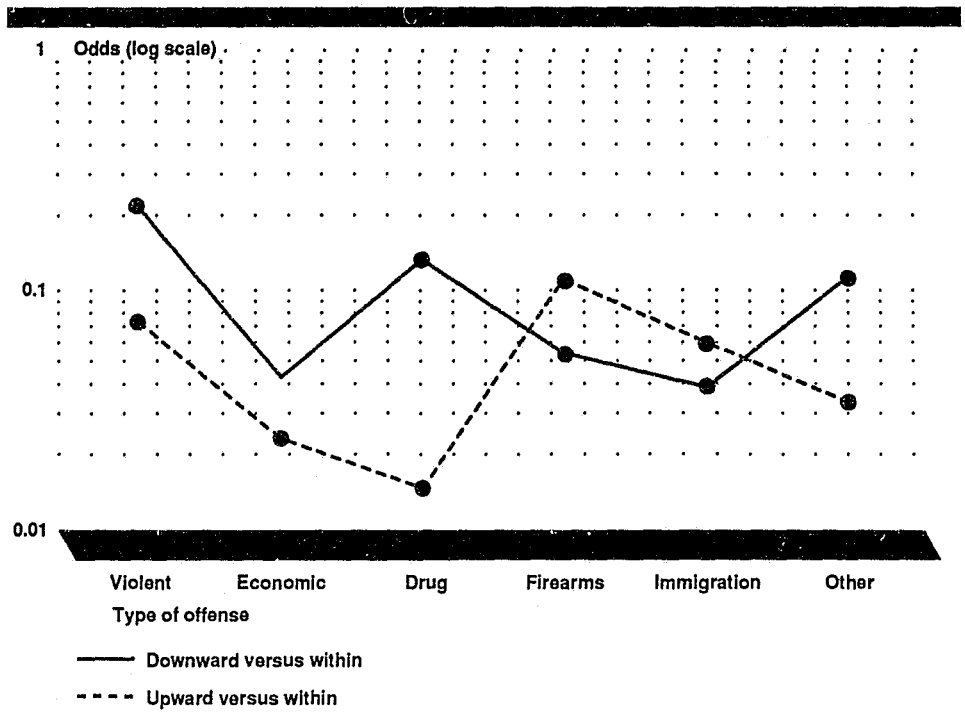
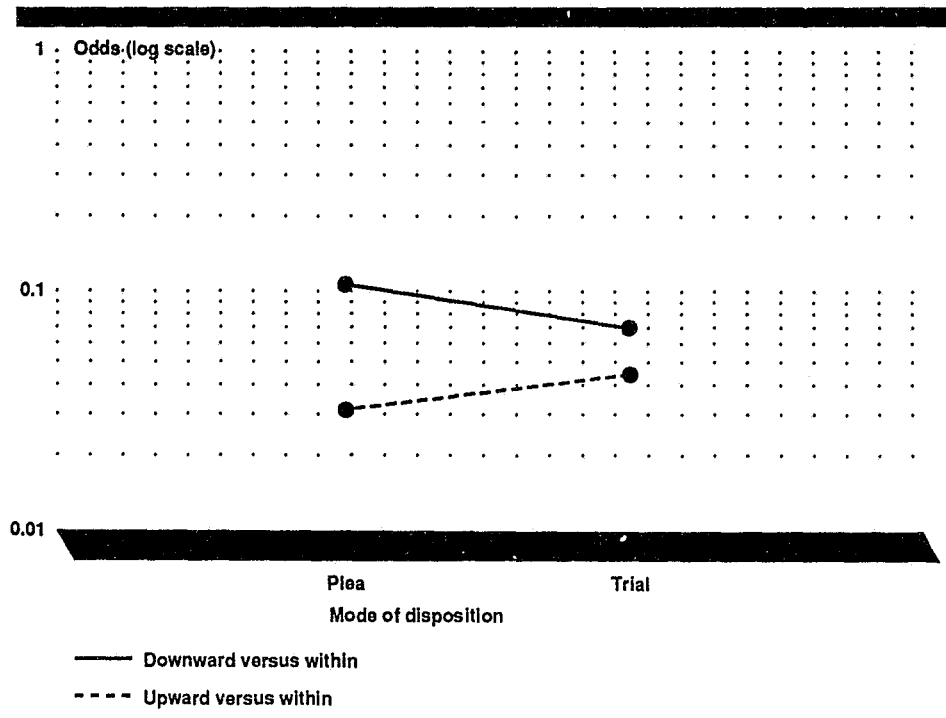


Figure II.5: Expected Odds on Downward and Upward Departures Vs. Within Range Sentences, by Mode of Disposition



1. Offenders in higher criminal history categories had lower odds on receiving downward departures than offenders in lower criminal history categories and higher odds on receiving upward departures. (See fig. II.2). The odds on receiving downward departures diminish by a factor of 0.74 as we move from criminal history category I-II to III-VI, and from category III-IV to V-VI, while the odds on receiving upward departures go down by a like amount as we move from category V-VI to III-IV and from III-IV to I-II.[16]

2. Offenders in higher seriousness categories were progressively more likely than offenders in lower seriousness categories to receive downward departures rather than within range sentences. Differences in these odds are monotonic, though not linear. (See fig. II.3.) Offenders in the lowest and highest seriousness levels, however, were less likely to receive sentences that involved upward departures than offenders in the middle seriousness levels.

[16]The odds ratio of 0.54 given in table II.7, which contrasts the odds for the highest criminal history category versus odds for the lowest criminal history category, is simply 0.74^2. In a linear model of the sort fitted here, the comparison of categories that are separated by two levels requires taking the linear parameter to the second power.

3. Considerable variation in the tendency to receive departure versus within range sentences existed across types of offenders, though differences across offender types defy any easy generalizations. Violent offenders and offenders of firearms statutes had higher odds on receiving sentences that departed upwards from the guidelines range than the other types of offenders, but violent offenders also had the highest odds on receiving sentences that depart downward. Drug offenders were more likely than any group except violent offenders to receive downward departures, and they also had the lowest odds on receiving upward departures. (See fig. II.4.)

4. Offenders convicted by plea were 1.52 times as likely as offenders convicted by trial to receive downward departures as opposed to within range sentences, and offenders convicted by trial were 1.52 times as likely as offenders who plead guilty to receive sentences that departed upward. (See fig. II.5.)

While an understanding of the effects of these legally relevant factors on sentence location is useful in its own right, our primary interest in them here is to see whether they account in any fashion for the differences we found across circuits in the tendency for offenders to receive departure versus within range sentences. Information bearing on that issue is provided in tables II.8 and II.9.

Table II.8: Models for the Three-Way Tables in Which Departure Vs. Within Range Sentences Were Cross-Classified by Circuit and (1) Criminal History, (2) Seriousness Level, (3) Type of Offense, and (4) Mode of Disposition

Additional factor	Model	Marginals/effects fitted	df	Chi-square	p
Criminal history	1	[CH] [L]	16	39.16	.001
	2	[CH] [HL]	12	26.45	.009
	3	[CH] [HL] [CL_2]	10	7.94	.634
	4	[CH] [HL] [CL]	8	6.16	.630
Seriousness level	1	[CS] [L]	22	121.62	<.001
	2	[CS] [SL]	16	37.74	.002
	3	[CS] [SL] [CL_2]	14	17.40	.236
	4	[CS] [SL] [CL]	12	15.52	.214
Type of offense	1	[CT] [L]	34	106.14	<.001
	2	[CT] [TL]	24	35.50	.061
	3	[CT] [TL] [CL_2]	22	21.82	.471
	4	[CT] [TL] [CL]	20	20.37	.435
Mode of disposition	1	[CD] [L]	10	31.58	<.001
	2	[CD] [DL]	8	24.50	.002
	3	[CD] [DL] [CL_2]	6	5.94	.430
	4	[CD] [DL] [CL]	4	3.64	.457

Legend: C=Circuit, H=Criminal history, S=Seriousness level, T=Type of offense, D=Mode of disposition, L=Sentence location.

Note: L_2 indicates an association with sentence location that involves the second (within range) category vs. the other two (upward and downward departures).

Table II.9: Odds Ratios Estimating the Association Between Circuit and Sentence Location, Before and After Controlling Four Factors

	Odds ratios	
Control variable	a	b
None	1.51	2.21
Criminal history	1.49	2.17
Seriousness level	1.59	2.25
Type of offense	1.44	2.03
Mode of disposition	1.48	2.17

Note: Odds ratio "a" represents the effect of being in Circuits Two, Five, Six, and Eight versus Three, Four, Seven, Ten, and Eleven on the odds on receiving departure (upward or downward) versus within range sentences. Odds ratio "b" represents the effect of being in Circuits One and Nine versus Three, Four, Seven, Ten, and Eleven on those same odds.

In table II.8, we show the results of fitting models to four three-way tables formed by cross-classifying the departure versus within range sentence

location variable by circuit[17] and (one at a time) (1) criminal history, (2) seriousness level, (3) type of offense, and (4) mode of disposition. We fit four hierarchical[18] models to each of the tables, all of which were logit-specified models that fit or allowed for the association between circuit and the additional factor in each table (i.e., criminal history, seriousness level, etc.), and varied only in terms of the effects they specify of circuit and these other factors on sentence location.

The first model fitted to each table was the logit-specified model of independence, which asserts that the sentence location variable is unrelated to either of the factors in the table. The second model fitted to each table allowed sentence location to be related to the additional factor in each table (i.e., the legally relevant factor), but not to circuit. It is hardly surprising that in all tables Model 2 improves significantly Model 1 inasmuch as the difference between them is formally testing the significance of the associations between sentence location and the four legally relevant characteristics that were revealed in the two-way tables just considered. It is noteworthy, however, that this second model did not fit the observed data acceptably in any of the four tables to which it was fitted, which suggested that there was something more "going on" in these tables than the simple associations between circuit and the four legally relevant factors and between these four factors and sentence location.

What more was going on here is revealed by the significant improvement in fit upon Model 2 afforded by Model 3. Model 3 added to Model 2 the constrained association between circuit and sentence location that we found in the two-way table involving those variables, which allowed differences to exist across circuit categories in the tendency for offenders to receive departure sentences versus within range sentences. The significant improvement of Model 3 over Model 2 established that these significant differences in sentencing across circuits persisted even after we allowed for the differences across circuits in offenders' criminal history, seriousness level, and so on, and after we allowed for the effect of those legal factors on sentence location. Moreover, the fact that Model 3 fit the data acceptably in all tables implied that these circuit differences were similar across all offender types, regardless of the seriousness of their

[17]In the three-way tables we analyzed the D.C. Circuit is omitted. Our bivariate work suggested that this circuit was unlike any of the others in terms of the tendency for offenders to receive departure sentences, and the small number (30) of cases in D.C. did not permit further cross-classification.

[18]The term hierarchical means that the models we fit here, considered in sequence, build upon previous models by adding additional associations or by relaxing constraints imposed upon those associations. The notation used in the tables to describe the various models we fit, closely follows the notation found in Leo A. Goodman's Analyzing Qualitative/Categorical Data, (Cambridge, MA: Abt Associates, Inc., 1978).

offense, their criminal history, and whether they pled guilty or went to trial, which is to say that there was no clear evidence of any three-way interactions in any of these tables. And finally, the fact that Model 3 was not improved upon by Model 4 established that this simplifying constraint on the circuit by sentence association adequately accounted for that association, and that no differences existed across circuits in the tendency for offenders to receive upward versus downward departures.

The fact that none of these legally relevant factors accounted completely for the differences in sentences across circuits does not necessarily imply that they do not account for those associations in part, and in table II.9 we show what happened to the odds ratios estimating the circuit by sentence location association when they were re-estimated from the expected frequencies under these models for the three-way tables. As table II.9 clearly shows, the effect of circuit on sentence location, or the differences across circuits in the tendency for offenders to receive departing sentences, was not even accounted for in a small way by differences across circuits in these legal factors. Odds ratios estimating those effects remained very similar after controlling for these factors (one at a time) to what they looked like before, and in fact our estimate of this association appeared slightly more pronounced when we controlled for differences across circuits in seriousness level.

Effects of Factors on Within Range Sentences

We now return to the two-way tables involving the cross-classifications of these factors by within range sentence categories. Chi-square values associated with models of independence for these tables are given in the right panel of table II.4 and repeated in the left panel of table II.10. We began our analysis of these tables by considering whether the four within range sentence location categories (bottom of the range, below the midpoint, at or above the midpoint, top of the range) could be collapsed without significant loss of information. Table II.10 reveals that, with respect to the tables in which within range sentence categories were cross-classified by the seven extralegal factors, no significant information was lost when we collapsed the two middle within range categories that involve the below the midpoint and above the midpoint sentences.[19] We found, in fact, that independence could be rejected in six of the seven

[19]As with previous partitionings reported, it was the difference chi-square values that alerted us as to whether significant information was lost as a result of collapsing. For the race by within range sentence table, for example, the difference chi-square of 16.39 - 13.55 = 2.84 with 6 - 4 = 2 df, which has a probability greater than .10, informed us that racial differences in below the midpoint versus above the midpoint sentences were insignificant or that independence holds in the 3 X 2 table in which the three racial categories were cross-classified by those two sentence categories. Similar interpretations could be given to the other insignificant difference chi-square values shown in table II.10.

collapsed tables involving the extralegal factors; only education remained unrelated to the location of the within range sentences. We found also that the location of within range sentences was significantly related to each of the four legally relevant factors we considered, though the significance of the relationship between mode of disposition and sentence location was not clearly established until this location variable was collapsed. While some information was lost when we collapsed these middle two sentence location categories in the tables involving criminal history and seriousness level, the bulk of the variation in those tables was retained after collapsing.[20] Moreover, because our primary interest in the legally relevant factors involved using them as controls in reestimating the significance and effects of the extralegal factors, the information lost with respect to these two tables as a result of collapsing would not diminish their capacity to account for the associations involving the extralegal factors (because the latter associations didn't involve differences in below and above the midpoint sentences). Our analysis proceeded, therefore, using the three within range categories (bottom of the range, middle of the range, top of the range) rather than the original four.

Table II.10: Additional Partitionings of the Within Range Subtables

	Full table			Collapsed table			Difference		
Factor	df	Chi-square	P	df	Chi-square	P	df	Chi-square	P
Race	6	16.39	.01	4	13.55	.009	2	2.84	>.10
Gender	3	43.49	<.001	2	43.46	<.001	1	0.03	>.75
Marital status	3	6.29	.10	2	6.07	.048	1	0.22	>.50
Employment status	6	29.85	<.001	4	29.43	<.001	2	0.42	>.75
Education	6	5.53	.48	4	4.07	.37	2	1.46	>.25
Age	6	10.48	.11	4	9.79	.044	2	0.69	>.50
Circuit	33	51.39	.02	22	40.54	.009	11	10.85	>.25
Criminal history	15	125.40	<.001	10	113.77	<.001	5	11.63	<.05
Seriousness level	117	295.49	<.001	78	206.28	<.001	39	89.21	<.001
Offense type	15	55.61	<.001	10	48.61	<.001	5	7.00	>.10
Mode of disposition	3	7.20	.07	2	6.26	.044	1	0.94	>.25

Note: Collapsed tables were formed by aggregating below the midpoint and above the midpoint sentence location categories.

[20]The ratio of the chi-square value for the collapsed table to the chi-square value for the full or uncollapsed table informed us directly about how much of the variation, or departure from independence, was retained after collapsing. That is, 113.76/125.40 = .907 and 206.28/295.49 =.698 implied that roughly 91 percent and 70 percent of the variation in within range sentences across categories of criminal history and offense seriousness, respectively, were retained after collapsing.

As we did in our analysis of departures, here too we attempted to establish which of the large number of categories of the circuit, criminal history, and seriousness level variables were related to sentence location before providing a simple description of those relationships. The results reported in table II.11 suggest that each of these polytomous variables could be collapsed without significant loss of information, though the aggregated categories that were retained to reveal differences in the location of within range sentences were different than the categories that were retained in our analyses of departure versus within range sentences. Variation across circuits in the location of within range sentences involved differences across three aggregated circuit categories that contrasted the D.C., Third, and Sixth Circuits with the Second and Fourth Circuits and with, in the third category, all remaining circuits.[21] Variation in within range sentences across criminal history categories involved differences between offenders in criminal history category I versus offenders in criminal history category II versus offenders in categories III-VI. Finally, the sequence of steps undertaken in attempting to collapse the 43 seriousness levels suggested that certain adjacent categories of that variable could be collapsed without significant loss of information, the result being a seven category variable that contrasted offenders in levels 1-6, 7-8, 9-11, 12-19, 20-28, 30-32, and 33-43.[22]

[21]As before, the categories that were "lumped together" were not significantly different from one another with respect to the sentencing outcome being considered, so the collapsing or aggregation of those categories was statistically justifiable.

[22]With respect to the location of within range sentences they received, offenders in offense seriousness level 29 and level 37 were significantly different from the categories of offenders both above them and below them. There were, however, too few offenders in the levels to retain them in subsequent analyses, and so our only recourse was to treat them as outliers and delete them from our analyses.

Table II.11: Final Partitionings of the Two-Way Within Range Subtables Involving Circuit, Criminal History, and Seriousness Level

Factor	Step	Categories combined on step	Independence model			Difference from previous step		
			df	Chi-square	p	df	Chi-square	p
Circuit	1	None	22	40.54	.009			
	2	[0,3,6]	18	34.67	.010	4	5.87	.10
	3	[1,5,7,8,9,10,11]	6	24.88		12	9.79	.50
	4	[2,4]	4	20.02		2	4.86	.05
Criminal history	1	None	10	113.77				
	2	[3,4,5,6]	4	108.34		6	5.43	.25
Seriousness level	1	None	78	206.28				
	2	[1-6]	68	193.49		10	12.79	.10
	3	[7-8]	66	192.47		2	1.02	.50
	4	[9-11]	62	190.28		4	2.19	.50
	5	[12-19]	48	176.65		14	13.63	.25
	6	[20-28]	32	155.68		16	20.97	.10
	7	[30-32]	28	146.80		4	8.88	.05
	8	[33-36,38-43]	16	131.41		12	15.39	.10

To describe the associations present in these collapsed tables, we fitted and compared as before models that placed different simplifying constraints on these associations. Table II.12 doesn't give results for the full range of models that we fitted to the 11 2-way tables we are now considering, but the results reported do serve to establish which models were chosen as preferred to describe the relationships in these 2-way tables. Models which placed linear constraints on the associations were chosen for the tables in which within range categories were cross-classified by gender, employment status, age, criminal history, and mode of disposition.[23]

[23]In all of these linear models the categories of the sentence location variable were regarded as ordered and equidistant, which implied that these five factors had a similar effect on the tendency for offenders to receive middle range rather than top of the range sentences as they had on the tendency for offenders to receive bottom of the range versus middle range sentences. The models chosen as preferred to describe the associations involving employment status and criminal history also imposed linear constraints on these two factors. The models for those two tables were, that is, linear by linear association or uniform association models.

Table II.12: Alternative Models for the Collapsed Two-Way Within Range Subtables

Factor	Model	Marginals/effects fitted	df	Chi-square	p
Race	1	[R] [L]	4	13.55	.009
	2	[RL']	2	12.09	.002
	3	[RL_2]	2	0.01	.997
Gender	1	[X] [L]	2	43.46	<.001
	2	[XL']	1	0.24	.622
Marital status	1	[M] [L]	2	6.07	.043
	2	[ML']	1	3.75	.053
	3	[ML_2]	1	0.60	.439
Employment status	1	[E] [L]	4	29.43	<.001
	2	[E'L']	3	1.21	.751
Age	1	[A] [L]	4	9.79	.044
	2	[A'L']	3	5.62	.132
	3	[AL']	2	0.66	.720
	4	[A'L]	2	5.23	.073
Circuit	1	[C] [L]	4	20.02	<.001
	2	[CL']	2	16.27	<.001
	3	[CL_2]	2	2.71	.258
Criminal history	1	[H] [L]	4	108.34	<.001
	2	[H'L']	3	0.89	.828
Seriousness level	1	[S] [L]	12	116.25	<.001
	2	[S'L']	11	113.60	<.001
	3	[S'L]	10	104.50	<.001
	4	[SL']	6	42.45	<.001
	5	[SL_1]	6	24.97	<.001
	6	[SL_2]	6	49.46	<.001
	7	[SL_3]	6	81.59	<.001
Type of offense	1	[T] [L]	10	48.61	<.001
	2	[TL']	5	13.74	.017
	3	[TL_1]	5	20.83	.001
	4	[TL_2]	5	34.69	<.001
	5	[TL_3]	5	17.09	.004
Mode of disposition	1	[D] [L]	2	6.26	.044
	2	[DL']	1	0.12	.730

(Table notes on next page)

Legend: R=Race; X=Gender, M=Marital status, E=Employment status, A=Age, C=Circuit, H=Criminal history, S=Seriousness level, T=Type of offense, D=Mode of disposition, L=Sentence location.

Note: Apostrophes indicate linear constraints; L_1, L_2, and L_3 indicate associations with sentence location that involve the first, second, and third categories of sentence location, respectively.

Models which constrained the association to involve a difference across factor categories in the odds on receiving middle range sentences versus sentences at either the bottom or top of the range were chosen as preferred to describe the associations in the tables involving race, marital status, and circuit. For the tables in which these three location categories were cross-classified by seriousness level and type of offense, none of the models that placed simplifying constraints on the data provided an acceptable fit, so for them the trivial model (not shown in table II.12) was chosen as preferred.

As before, to describe these associations we calculated odds and odds ratios from the expected frequencies under the preferred models. These are given in table II.13. The odds calculated informed us as to what the likelihood of receiving bottom versus middle range and top versus middle range sentences were for the various categories of offenders, and the ratios informed us in a direct fashion how much more or less likely to receive varying sentences some categories of offenders were than others. Figures II.6 through II.15, in which these odds are plotted on logarithmic (i.e., multiplicative) scales, help to describe these associations. From these calculations and graphs, the following results are apparent:

Table II.13: Expected Frequencies Under the Preferred Models for the Two-Way Within Range Subtables, and Odds and Odds Ratios Derived From Them

	Sentence location						
Factor	Bottom range	Middle range	Top range	Odds on bottom vs. middle	Odds ratios	Odds on top vs. middle	Odds ratios
Race							
White	484.20	249.00	134.80	1.94		0.541	
Black	283.17	111.00	78.83	2.55	1.31	0.710	1.31
Hispanic	148.63	106.00	41.37	1.40	0.72	0.390	0.72
Gender							
Male	742.32	419.36	246.32	1.77		0.587	
Female	204.68	60.64	18.68	3.38	1.91	0.308	0.52
Marital status							
Married	315.05	137.00	89.95	2.30	1.33	0.657	1.33

(continued)

Factor	Sentence location			Odds on bottom vs. middle	Odds ratios	Odds on top vs. middle	Odds ratios
	Bottom range	Middle range	Top range				
Unmarried	451.95	262.00	129.05	1.73		0.493	
Employment status							
Unemployed	211.84	141.50	96.66	1.50		0.683	
Partially employed	177.75	95.15	52.10	1.87	1.25	0.548	0.80
Fully employed	378.41	162.35	71.24	2.33	1.55	0.439	0.64
Age							
18-25	255.33	111.35	52.33	2.29	1.28	0.470	0.78
26-35	375.19	209.61	126.19	1.79		0.602	
36+	355.41	190.04	109.48	1.87	1.04	0.576	0.96
Circuit							
0, 3, 6	143.92	42.00	40.08	3.43	1.94	0.954	1.94
1, 5, 7-11	722.74	409.00	201.26	1.77		0.492	
2, 4	178.34	74.00	49.66	2.41	1.36	0.671	1.36
Criminal history							
I	708.40	282.55	111.06	2.51	2.07	0.393	.48
II	112.83	64.66	36.51	1.74	1.44	0.565	.70
III-VI	214.77	176.80	143.43	1.21		0.811	
Seriousness level							
1-6	278	61	46	4.56		0.754	
7-8	70	79	29	0.89	0.20	0.367	0.49
9-11	177	88	52	2.01	0.44	0.591	0.78
12-19	174	118	84	1.47	0.32	0.712	0.94
20-28	199	103	59	1.93	0.42	0.572	0.76
30-32[a]	87	27	7	3.22	0.71	0.259	0.34
33-43[a]	52	43	7	1.21	0.27	0.163	0.22
Offense type							
Violent	41	24	28	1.71	0.69	1.167	1.70
Economic	286	115	79	2.49		0.687	
Drug	403	218	88	1.85	0.74	0.404	0.59
Firearms	44	34	30	1.29	0.52	0.882	1.28
Immigration	45	39	16	1.15	0.46	0.410	0.60
Other	132	52	24	2.53	1.02	0.461	0.67
Mode of disposition							
Plea	882.81	427.38	230.81	2.07	1.23	0.540	0.81
Trial	147.19	87.62	58.19	1.68		0.664	

[a]Levels 29 and 37 were deleted.

1. Blacks were more likely than whites to receive sentences at the bottom or top of the guidelines range rather than in the middle of the range, by a factor of roughly 1.3. Hispanics conversely were somewhat less likely than whites (by a factor of 0.7) and by implication considerably less likely than blacks (by a factor of 0.7/1.3 = 0.5) to receive bottom or top of the range sentences.

2. Females were nearly twice as likely (i.e, 1.91 times as likely, to be exact) as males to receive sentences at the bottom rather than in the middle of the range and in the middle of the range versus at the top of the range.[24] The latter implies, of course, that females were roughly half as likely as males (i.e., 0.52 times as likely) to receive top of the range versus middle of the range sentences, or middle versus bottom range sentences.[25]

3. Married offenders were more likely than unmarried offenders to receive sentences at the bottom and top of the range rather than in the middle of the range, by a factor of 1.33.

4. Offenders employed part-time were more likely than unemployed offenders and offenders employed full-time were more likely than offenders employed part-time to receive bottom versus middle range sentences. They were, at the same time, less likely to receive sentences at the top of the range rather than in the middle of the range.[26]

5. Offenders between the ages of 18 and 25 were 1.28 times as likely as offenders aged 26-35 to receive sentences at the bottom of the range rather than in the middle and less likely to receive sentences at the top of the range versus in the middle, by a factor of 1/1.28 = 0.78. Offenders older

[24]The equality of the effect of sex on the odds on bottom versus middle and middle versus top range sentences is a constraint of the linear model chosen to describe this relationship. In one sense, the single odds ratio of 1.91 "understates" the magnitude of the association between sex and location. The tendency to receive sentences at the bottom rather than at the top of the range was 1.9 x 1.9 = 3.6 times as great for females as for males.

[25]The odds ratio of 0.52 in table II.13 is redundant, in the sense that it is simply the reciprocal of (and as such can be derived from) the odds ratio of 1.91. Related to this, the dotted line in figure II.7, which represents the effect of being female versus male on the odds on receiving top of the range versus mid-range sentences, is the reverse image of the solid line that represents the effect of being female versus male on the odds on receiving bottom range versus mid-range sentences.

[26]In this table the uniform association or linear by linear association model implies that the single odds ratio of 1.25 describes the whole of the association in the table. All of the other odds ratios given in table II.13 can be derived from this one (i.e., 1.55 = 1.25 x 1.25, apart from rounding, and 0.80 and 0.64 are the reciprocals of 1.25 and 1.55, respectively). Since this uniform association parameter estimates the effect of being in higher adjacent categories of the employment status variable on the odds on being in lower adjacent categories of sentence location, the ratio describing the difference between fully employed offenders and unemployed offenders on the odds on receiving sentences at the bottom rather than top of the range would be 1.25 to the fourth power, or 2.44, apart from rounding.

than 36 were not much different from offenders aged 26 through 35, which is to say that the ratio of 1.04 (and its reciprocal, 0.96) was not much different from unity, which would imply no difference.

6. Offenders in the D.C., Third, and Sixth Circuits, were 1.94 times as likely and offenders in the Second and Fourth Circuits were 1.36 times as likely as offenders in the remaining circuits to receive sentences at either the bottom or top of the guidelines range rather than in the middle of the range.

7. Offenders in criminal history category I were more likely than offenders in criminal history category II, who in turn were similarly more likely than offenders in criminal history categories III-VI, to receive bottom range sentences rather than middle of the range sentences. Offenders in lower adjacent criminal history categories were, in like manner, less likely than those in higher categories to receive top of the range sentences.[27]

8. Those in the lowest seriousness level, where bottom range sentences could mean no prison time, were considerably more likely than offenders whose crimes were more serious to receive sentences at the bottom rather than middle of the guidelines range. They also, however, had the highest odds on receiving sentences at the top of the range. While the former odds were nonlinearly and nonmonotonicly related to seriousness level, the latter odds (on receiving top of the range sentences) did diminish in a roughly linear fashion across the seriousness categories beginning with level 12 (see fig. II.13).

9. While differences across offense types in sentence location were apparent, they were not systematic and did not reveal any clear pattern of some types of offenders being treated more or less severely than others. Offenders sentenced for economic crimes and offenders in the "other" category had the highest odds on receiving sentences at the bottom of the range, but they did not have the lowest odds on receiving top of the range sentences, which were exhibited by drug offenders. Violent offenders had the highest odds on receiving sentences at the top of the range versus the middle and, relative to all of the clearly defined types of offenders in table II.13 except for offenders of economic crimes, the lowest odds on receiving sentences at the bottom of the range.

[27]All of the ratios given in table II.13 can be derived from the single uniform association parameter of 1.44, which also implies that the odds on receiving sentences at the bottom of the range versus the top of the range were 1.44 x 1.44 x 1.44 x 1.44 = 2.07 x 2.07 = 4.3 times as likely for those in criminal history category I as for those in categories III-VI, apart from rounding.

10. Offenders convicted by plea were more likely than offenders who were convicted by trial to receive sentences at the bottom of the guidelines range rather than in the middle of the range and less likely to receive sentences at the top versus in the middle of the range. The linear parameter equal to 1.23 describing this association implied that offenders who pled guilty were 1.23 x 1.23 = 1.51 times as likely as other offenders to receive sentences at the bottom rather than top of the guidelines range.

Figure II.6: Expected Odds on Bottom and Top of the Range Vs. Middle Range Sentences, by Race

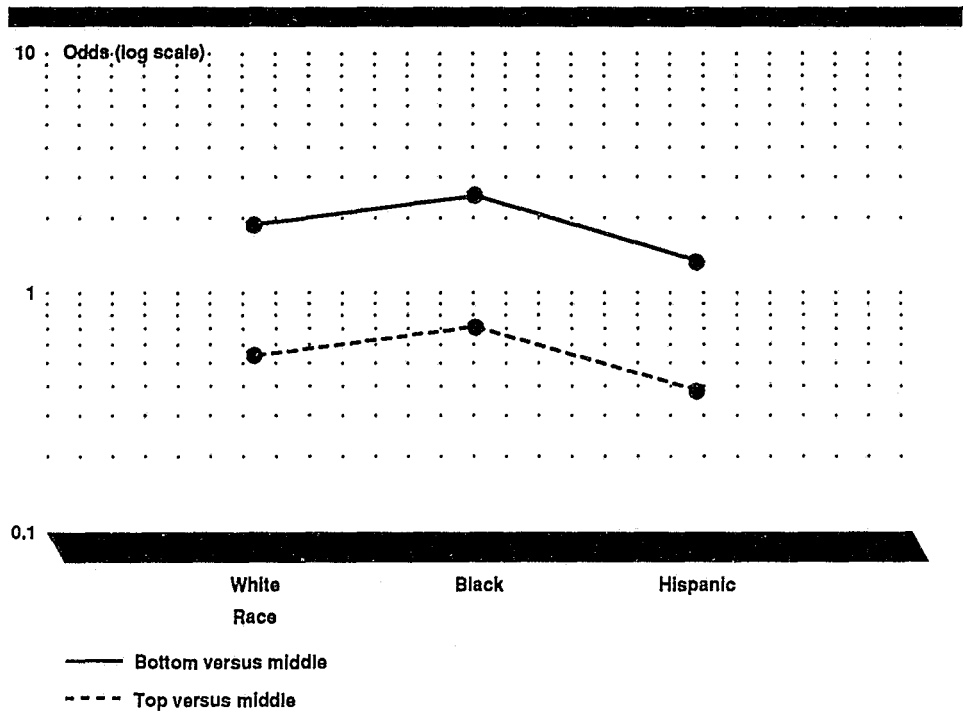


Figure II.7: Expected Odds on Bottom and Top of the Range Vs. Middle Range Sentences, by Gender

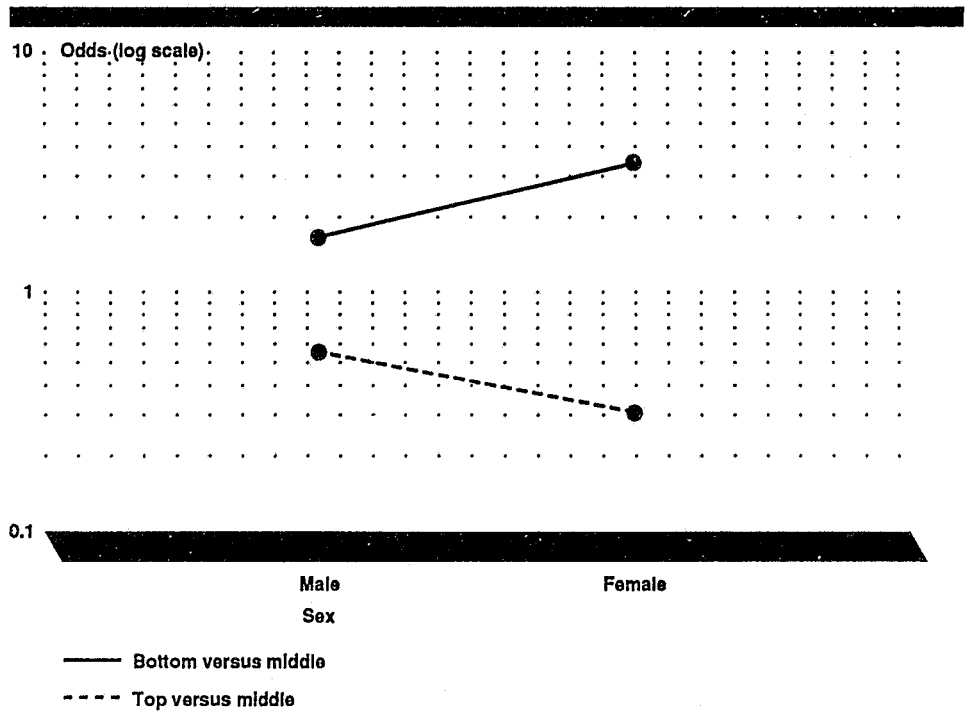


Figure II.8: Expected Odds on Bottom and Top of the Range Vs. Middle Range Sentences, by Marital Status

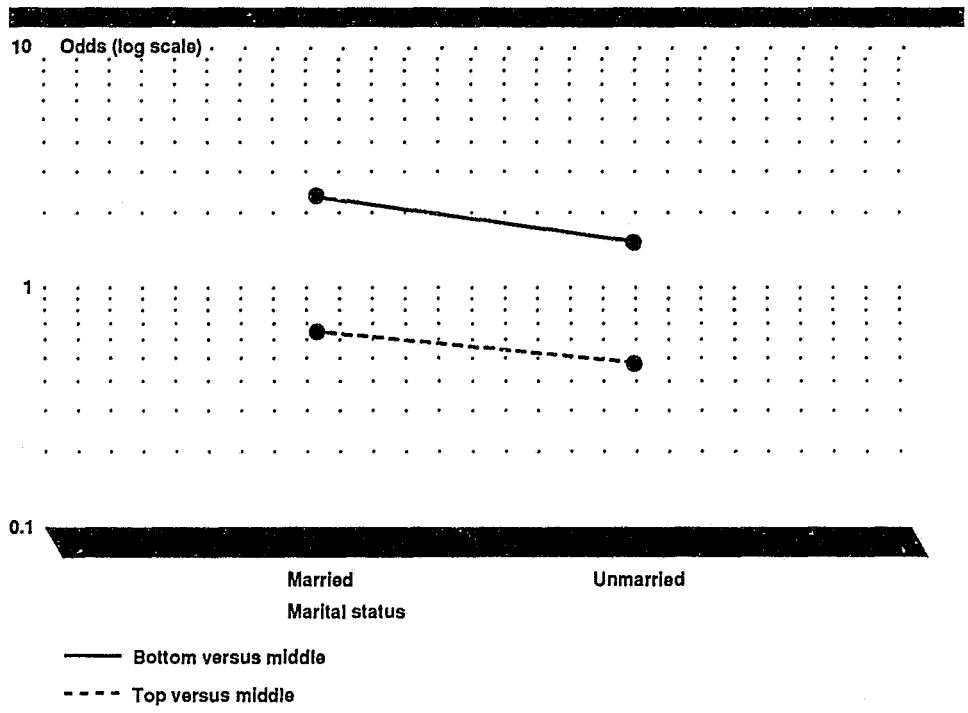


Figure II.9: Expected Odds on Bottom and Top of the Range Vs. Middle Range Sentences, by Employment Status

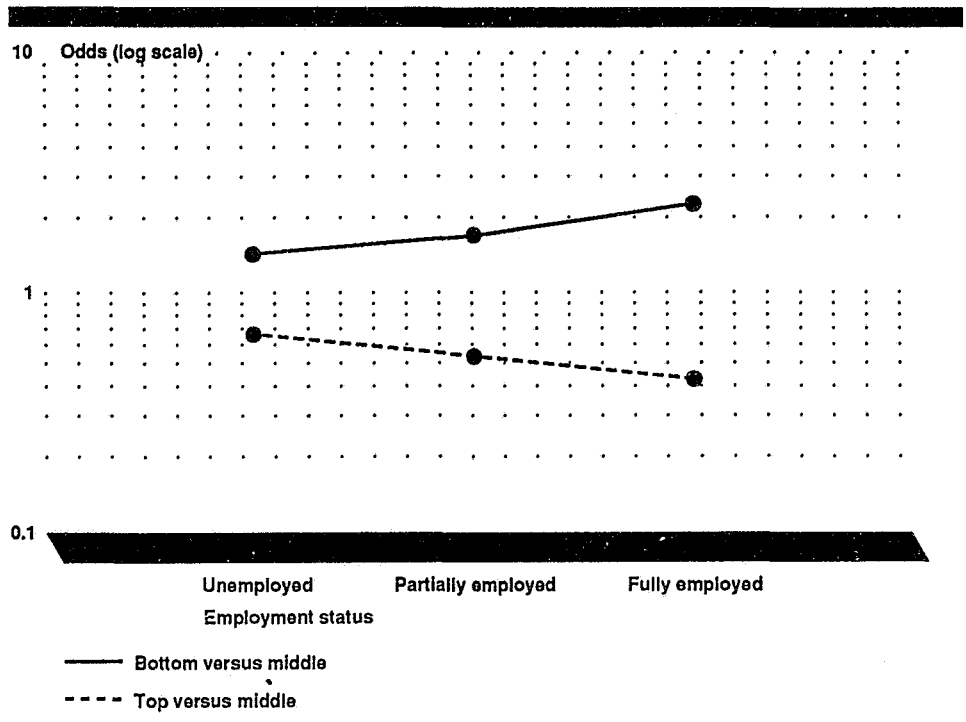


Figure II.10: Expected Odds on Bottom and Top of the Range Vs. Middle Range Sentences, by Age

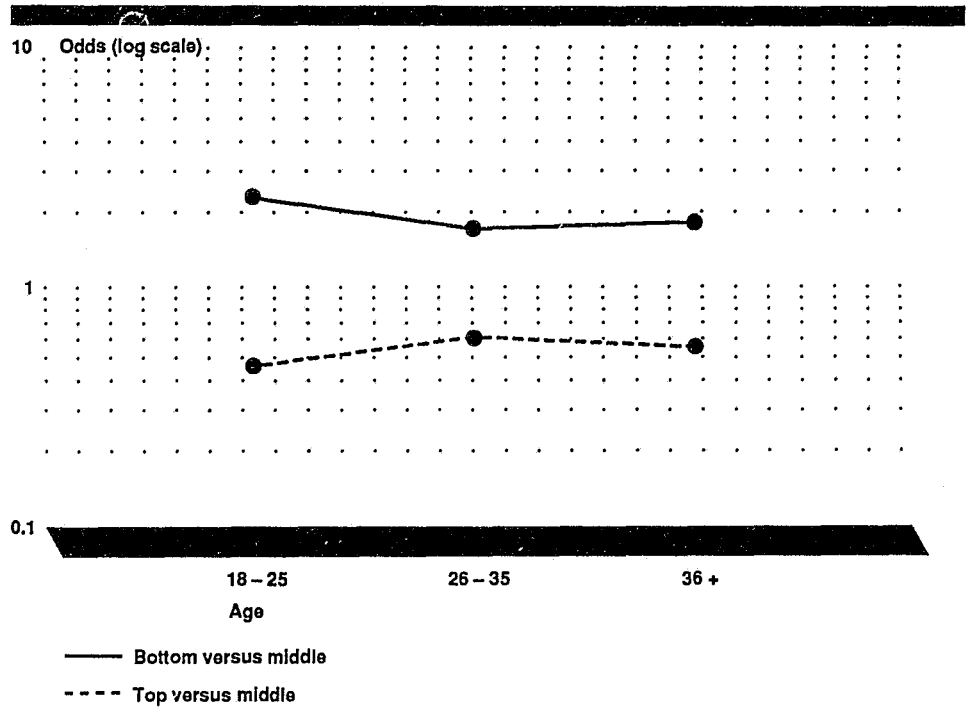


Figure II.11: Expected Odds on Bottom and Top of the Range Vs. Middle Range Sentences, by Circuit

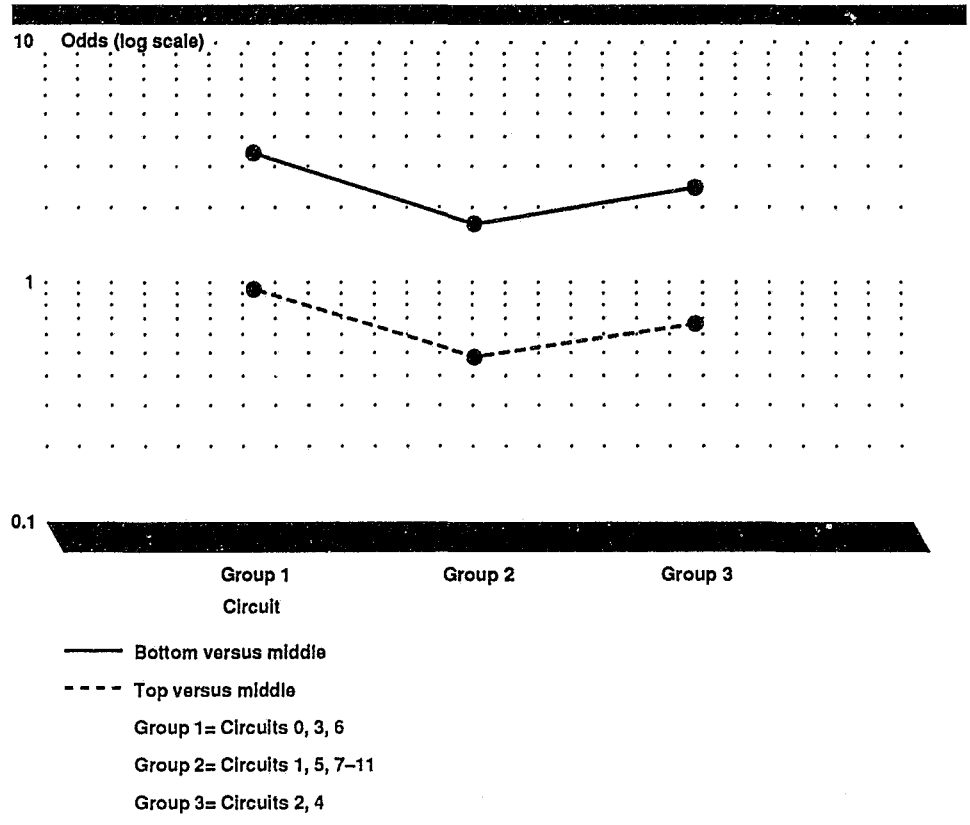


Figure II.12: Expected Odds on Bottom and Top of the Range Vs. Middle Range Sentences, by Criminal History

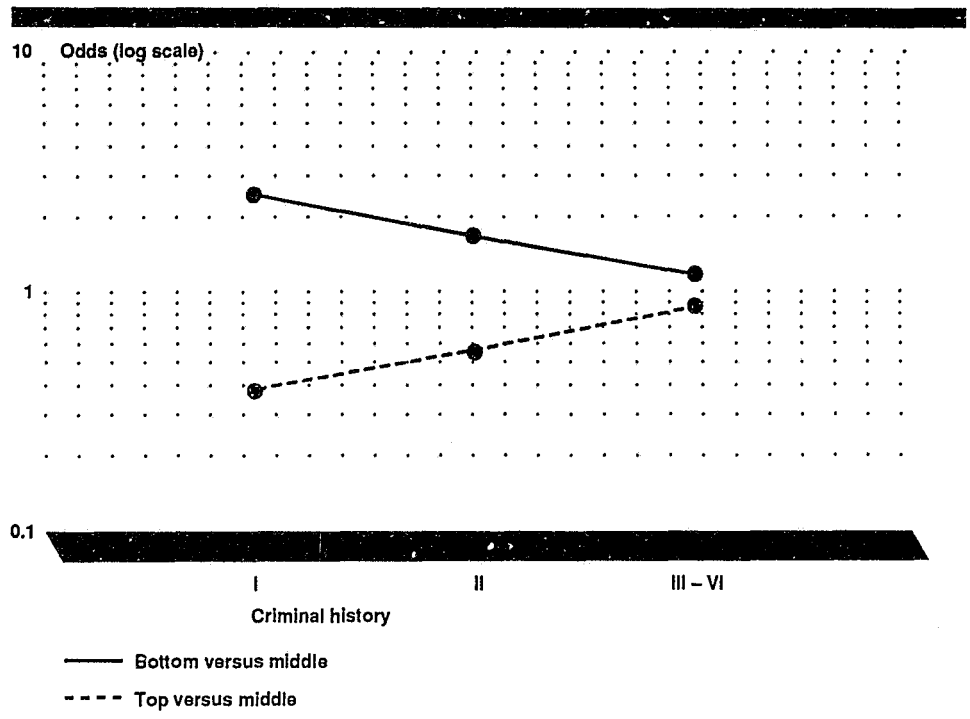
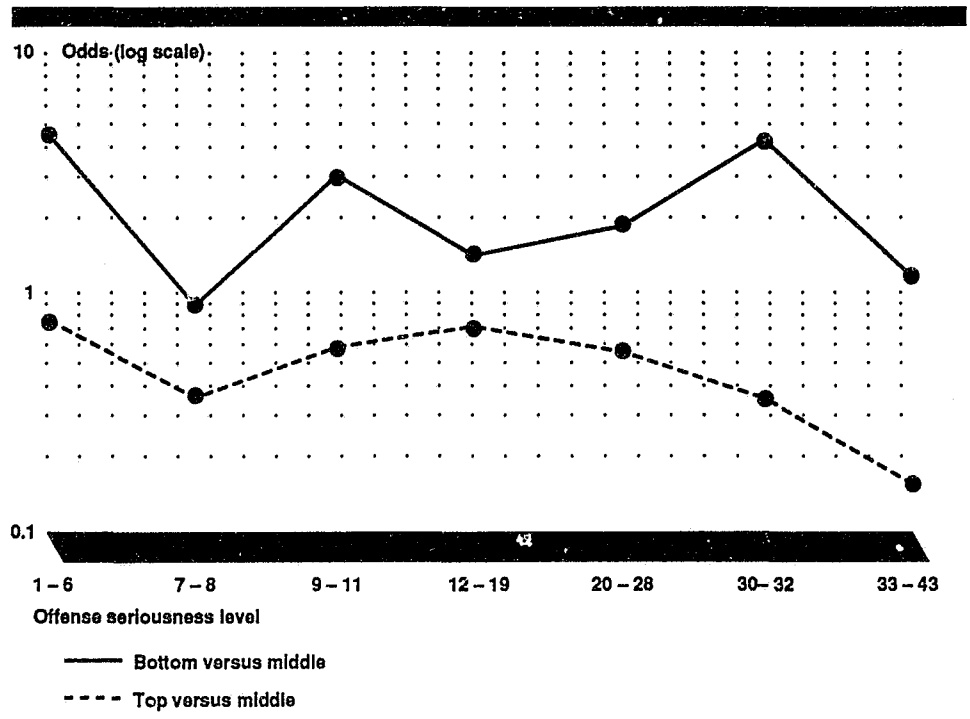


Figure II.13: Expected Odds on Bottom and Top of the R nge Vs. Middle Range Sentences, by Offense Seriousness Level



Note: Offenders at seriousness levels 29 and 37 were deleted from the analysis. See footnote 22 for explanation

Figure II.14: Expected Odds on Bottom and Top of the Range Vs. Middle Range Sentences, by Type of Offense

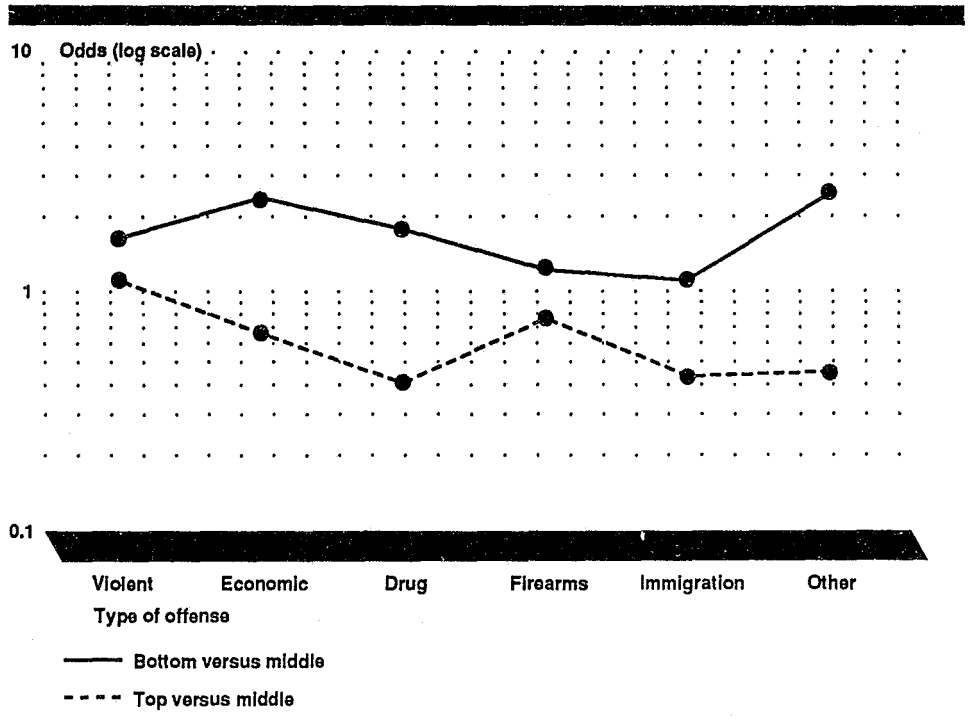
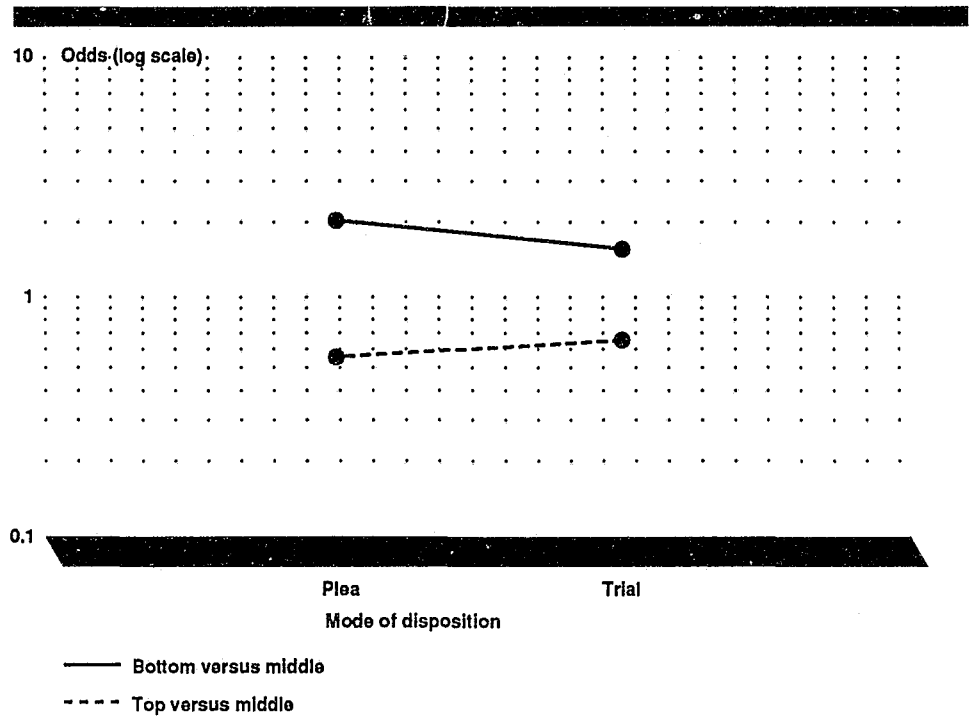


Figure II.15: Expected Odds on Bottom and Top of the Range Vs. Middle Range Sentences, by Mode of Disposition



To discern, as we did in our investigation of departure sentences, whether the associations between these extralegal factors and sentence location were accounted for in whole or in part by legal factors that were jointly related to them, we reestimated the odds ratios describing these associations using expected frequencies under models fitted to three-way tables formed by cross-classifying sentence location by each extralegal factor (considered one at a time) and each legally relevant factor (again considering one at a time).[28] While we have not provided each of the three-way tables in this report, we show in tables II.14 through II.19 the results of fitting a set of four similar models to these tables. In analyzing each of the 24 three-way tables formed by cross-classifying sentence location by the 6 extralegal and 4 legal factors, we fit (1) a logit specified model of independence, (2) a model which allowed the legal factor to be

[28]It would have been preferable, of course, to estimate these associations after controlling all legally relevant factors simultaneously, but sample sizes did not permit working with a table that cross-classified all of these variables simultaneously. It appeared, at any rate, that for the most part controlling for many of the extralegal factors did not alter our estimates of these associations appreciably, so we have some reason to believe that a larger sample and an expanded cross-classification might produce very similar results. Nonetheless we admit that the results from our three-way tables should be viewed as preliminary, pending a more thorough investigation of a larger number of cases. We note here too that a more extensive examination should pay particular attention to interactions in the multivariate tables constructed, which we simply have not had time, nor in many instances sufficient cases, to investigate rigorously here.

related to both the extralegal factor and to sentence location,[29] (3) a model which allowed, in addition to these associations, an association between sentence location and the extralegal factor that was constrained as in our two-way tables, and (4) a model which allowed an unconstrained version of that latter association. In all of the tables we fitted these models to, we found that the third model improved significantly upon the fit of the second, which established that the association of each of the extralegal factors with sentence location remained significant after each of the factors was controlled. Moreover, the fact that the third model fitted to each table did not improve significantly upon the second implied that the simplifying constraints imposed on these associations between the extralegal factors and the location variable remained tenable and appropriate constraints after controls.

Table II.14: Models for the Three-Way Tables in Which Within Range Sentences Are Cross-Classified by Race and (1) Criminal History, (2) Seriousness Level, (3) Type of Offense, and (4) Mode of Disposition

Additional factor	Model	Marginals/effects fitted	df	Chi-square	p
Criminal history	1	[RH] [L]	16	114.86	<.001
	2	[RH] [HL]	12	18.11	.112
	3	[RH] [HL] [RL_2]	10	3.87	.953
	4	[RH] [HL] [RL]	8	3.62	.890
Seriousness level	1	[RS] [L]	40	163.61	<.001
	2	[RS] [SL]	28	74.54	<.001
	3	[RS] [SL] [RL_2]	26	62.91	<.001
	4	[RS] [SL] [RL]	24	62.04	<.001
Type of offense	1	[RT] [L]	34	104.23	<.001
	2	[RT] [TL]	24	56.66	<.001
	3	[RT] [TL] [RL_2]	22	48.17	.001
	4	[RT] [TL] [RL]	20	47.96	<.001
Mode of disposition	1	[RD] [L]	10	26.63	.003
	2	[RD] [DL]	8	23.46	.003
	3	[RD] [DL] [RL_2]	6	9.11	.168
	4	[RD] [DL] [RL]	4	9.06	.060

Legend: R=Race, H=Criminal history, S=Seriousness level, T=Type of offense, D=Mode of disposition, L=Sentence location.

Note: L_2 indicates an association with sentence location that involves the second (middle range) category vs. the other two (bottom and top of the range).

[29]While we constrained some of the associations between the legally relevant factors and sentence location in our two-way tables to simplify our description of them, in our three-way tables we left them all unconstrained. This provided a more rigorous test of the effect of the extralegal factors on sentence location after the legal factors were controlled.

Table II.15: Models for the Three-Way Tables In Which Within Range Sentences Are Cross-Classified by Gender and (1) Criminal History, (2) Seriousness Level, (3) Type of Offense, and (4) Mode of Disposition

Additional factor	Model	Marginals/effects fitted	df	Chi-square	p
Criminal history	1	[XH] [L]	10	124.52	<.001
	2	[XH] [HL]	6	29.94	<.001
	3	[XH] [HL] [XL']	5	6.41	.268
	4	[XH] [HL] [XL]	4	5.79	.215
Seriousness level	1	[XS] [L]	26	155.47	<.001
	2	[XS] [SL]	14	59.93	<.001
	3	[XS] [SL] [XL']	13	22.16	.053
	4	[XS] [SL] [XL]	12	22.08	.037
Type of offense	1	[XT] [L]	22	107.63	<.001
	2	[XT] [TL]	12	57.66	<.001
	3	[XT] [TL] [XL']	11	19.36	.055
	4	[XT] [TL] [XL]	10	19.36	.036
Mode of disposition	1	[XD] [L]	6	53.53	<.001
	2	[XD] [DL]	4	50.31	<.001
	3	[XD] [DL] [XL']	3	8.94	.030
	4	[XD] [DL] [XL]	2	8.72	.013

Legend: X=Sex, H=Criminal history, S=Seriousness level, T=Type of Offense, D=Mode of disposition, L=Sentence location.

Note: Apostrophes indicate linear constraints.

Table II.16: Models for the Three-Way Tables in Which Within Range Sentences Are Cross-Classified by Marital Status and (1) Criminal History, (2) Seriousness Level, (3) Type of Offense, and (4) Mode of Disposition

Additional factor	Model	Marginals/effects fitted	df	Chi-square	p
Criminal history	1	[MH] [L]	10	85.26	<.001
	2	[MH] [HL]	6	7.54	.274
	3	[MH] [HL] [ML_2]	5	3.42	.635
	4	[MH] [HL] [ML]	4	2.94	.568
Seriousness level	1	[MS] [L]	26	131.36	<.001
	2	[MS] [SL]	14	26.61	.022
	3	[MS] [SL] [ML_2]	13	20.12	.092
	4	[MS] [SL] [ML]	12	19.20	.084
Type of offense	1	[MT] [L]	22	62.78	<.001
	2	[MT] [TL]	12	14.39	.276
	3	[MT] [TL] [ML_2]	11	9.74	.554
	4	[MT] [TL] [ML]	10	9.64	.473
Mode of disposition	1	[MD] [L]	6	8.60	.197
	2	[MD] [DL]	4	7.26	.123
	3	[MD] [DL] [ML_2]	3	2.42	.490
	4	[MD] [DL] [ML]	2	1.76	.415

Legend: M=Marital status, H=Criminal history, S=Seriousness level, T=Type of offense type, D=Mode of disposition, L=Sentence location.

Note: L_2 indicates an association with sentence location that involves the second (middle range) category vs. the other two (bottom and top of the range).

Table II.17: Models for the Three-Way Tables In Which Within Range Sentences Are Cross-Classified by Employment Status and (1) Criminal History, (2) Seriousness Level, (3) Type of Offense, and (4) Mode of Disposition

Additional factor	Model	Marginals/effects fitted	df	Chi-square	p
Criminal history	1	[EH] [L]	16	106.92	<.001
	2	[EH] [HL]	12	30.58	.002
	3	[EH] [HL] [E'L']	11	19.33	.055
	4	[EH] [HL] [EL]	8	17.80	.023
Seriousness level	1	[ES] [L]	40	160.04	<.001
	2	[ES] [SL]	28	56.13	.001
	3	[ES] [SL] [E'L']	27	28.53	.384
	4	[ES] [SL] [EL]	24	26.22	.342
Type of offense	1	[ET] [L]	34	93.08	<.001
	2	[ET] [TL]	24	45.34	.005
	3	[ET] [TL] [E'L']	23	22.91	.466
	4	[ET] [TL] [EL]	20	20.71	.414
Mode of disposition	1	[ED] [L]	10	35.35	<.001
	2	[ED] [DL]	8	34.01	<.001
	3	[ED] [DL] [E'L']	7	7.39	.389
	4	[ED] [DL] [EL]	4	6.01	.199

Legend: E=Employment status, H=Criminal history, S=Seriousness level, T=Type of offense, D=Mode of disposition, L=Sentence location.

Note: Apostrophes indicate linear constraints.

Table II.18: Models for the Three-Way Tables In Which Within Range Sentences Are Cross-Classified by Age and (1) Criminal History, (2) Seriousness Level, (3) Type of Offense, and (4) Mode of Disposition

Additional factor	Model	Marginals/effects fitted	df	Chi-square	p
Criminal history	1	[AH] [L]	16	122.45	<.001
	2	[AH] [HL]	12	16.28	.179
	3	[AH] [HL] [AL']	10	9.93	.447
	4	[AH] [HL] [AL]	8	9.22	.324
Seriousness level	1	[AS] [L]	40	144.08	<.001
	2	[AS] [SL]	28	35.96	.144
	3	[AS] [SL] [AL']	26	28.92	.315
	4	[AS] [SL] [AL]	24	27.68	.274
Type of offense	1	[AT] [L]	34	81.19	<.001
	2	[AT] [TL]	24	27.88	.265
	3	[AT] [TL] [AL']	22	19.77	.597
	4	[AT] [TL] [AL]	20	19.07	.518
Mode of disposition	1	[AD] [L]	10	23.61	.009
	2	[AD] [DL]	8	18.60	.017
	3	[AD] [DL] [AL']	6	9.02	.173
	4	[AD] [DL] [AL]	4	8.01	.091

Legend: A=Age, H=Criminal history, S=Seriousness level, T=Type of offense, D=Mode of disposition, L=Sentence location.

Note: Apostrophes indicate linear constraints.

Table II.19: Models for the Three-Way Tables in Which Within Range Sentences Are Cross-Classified by Circuit and (1) Criminal History, (2) Seriousness Level, (3) Type of Offense, (4) Mode of Disposition

Additional factors	Model	Marginals/effects fitted	df	Chi-square	p
Criminal history	1	[CH] [L]	16	143.87	<.001
	2	[CH] [HL]	12	35.53	<.001
	3	[CH] [HL] [CL_2]	10	19.26	.037
	4	[CH] [HL] [CL]	8	15.75	.046
Seriousness level	1	[CS] [L]	40	170.93	<.001
	2	[CS] [SL]	28	54.69	.002
	3	[CS] [SL] [CL_2]	26	41.49	.028
	4	[CS] [SL] [CL]	24	37.96	.035
Type of offense	1	[CT] [L]	34	110.00	<.001
	2	[CT] [TL]	24	61.39	<.001
	3	[CT] [TL] [CL_2]	22	50.77	<.001
	4	[CT] [TL] [CL]	20	45.22	.001
Mode of disposition	1	[CD] [L]	10	29.66	.001
	2	[CD] [DL]	8	23.39	.003
	3	[CD] [DL] [CL_2]	6	5.72	.455
	4	[CD] [DL] [CL]	4	2.51	.642

Legend: C=Circuit, H=Criminal History, S=Seriousness level, T=Type of offense, D=Mode of disposition, L=Sentence location.

Note: L_2 indicates an association with sentence location that involves the second (middle range) category vs. the other two (bottom and top of the range).

Table II.20 provides estimates of these odds ratios after these controls were made. As is clear from that table, these associations we have just described in a two-way context were not greatly affected by these controls. The difference between blacks and whites in the tendency to receive bottom or top versus middle of the range sentences remained very similar after controlling for racial differences in criminal history, seriousness level, and so on, and after allowing for the effects of these legal characteristics on sentence location. The difference between Hispanics and whites was diminished somewhat by the control for type of offense (i.e., the odds ratio changed from .72 to .83, the latter being closer to unity), and differences across gender and employment categories were somewhat smaller after controlling for differences in criminal history. But these effects remained significant and roughly similar in magnitude to what they looked like prior to controls.

Table II.20: Odds Ratios Estimating the Effects of the Extralegal Factors on Sentence Location, Before and After Controlling for Four Legally Relevant Factors

	Odds ratios				
		Controlling for			
Factor	Without controls	Criminal history	Seriousness level	Type of offense	Mode of disposition
Race	1.31	1.32	1.29	1.33	1.35
	0.72	0.71	0.73	0.83	0.72
Gender	1.91	1.66	1.86	1.86	1.89
Marital status	1.33	1.29	1.38	1.30	1.31
Employment status	1.25	1.16	1.26	1.22	1.24
Age	1.28	1.22	1.25	1.25	1.29
	1.04	1.00	1.05	0.98	1.06
Circuit	1.94	1.92	1.83	1.73	1.93
	1.36	1.34	1.28	1.31	1.41

Note: Odds ratios calculated without controls are from table II.13. Other odds ratios shown are calculated in the same fashion as in that table, using the expected frequencies from models for the three-way tables.

Summary

While these results are tentative, they suggest that the differences in sentence location we found across different categories of offenders defined by extralegal characteristics were not easily or readily accounted for by controlling separately for four legally relevant factors that seem to us to be among the most proximate or important factors related to guideline sentencing. Whether they could be accounted for by other factors remains unanswered. Until an attempt to answer that question is made with a larger body of data and the more elaborate multivariate models that such data would permit, it seems reasonable to conclude that while the guidelines may have diminished in general the amount of variation in the length of sentences meted out to similarly situated offenders, they may not have removed all disparities that existed across categories of offenders defined by their race, gender, age, employment status, or marital status and by the judicial circuit in which they were sentenced.

Racial Differences in Robbery Sentencing Patterns Under the Guidelines

This appendix describes the methods and results of the analyses we did to determine whether racial differences in sentencing patterns existed under the guidelines. In our work we attempted to extend the Commission's analysis as found in its evaluation report.[1]

Objectives and Approach

Our objective in this analysis was to investigate patterns of legitimate judicial discretion and racial disparities in sentencing as they occurred for bank robbers under the guidelines. We conducted two analyses of racial differences in the sentencing of bank robbers. First, we attempted to uncover and describe race effects in lengths of sentences imposed and to determine whether the race effects were consistent across the guidelines sentencing table. To do so, we estimated white/black differences in mean lengths of sentences imposed on defendants, controlling for racial differences in criminal history categories and in offense severity.

Second, we attempted to explain the racial gap in sentences imposed. We estimated separate white and black regressions of sentence length on legally relevant and extralegal factors and used a procedure to distinguish between two portions of the racial sentencing gap: the portion attributable to racial differences in offense and offender characteristics (e.g., racial differences in offense severity levels or age) and the portion attributable to racial differences in the relation of these characteristics to sentence lengths (i.e., the portion due to the coefficients). We tested the assumption that race effects were constant throughout the ranges of relevant independent variables, and we also attempted to address issues such as which legally relevant factors other than the guidelines scores were correlated with sentence outcomes and whether the guidelines scores were correlated with disparities in outcomes.

The details of the methods and results for these analyses are discussed later in this appendix.

Results

From our first analysis, we found a variety of ways in which race was associated with sentence outcomes throughout the sentencing table. We found evidence of interactions between race and offense severity such that blacks received shorter sentences than whites at the lower range of offense severity but longer sentences at the upper range. We also found that in some portions of the sentencing table, blacks received longer

[1]"Judicial Sentencing Patterns Under the Guidelines, The Federal Sentencing Guidelines," Vol.II, Ch. 4, Sec. A of Part VI.

sentences than whites throughout the entire range of offense severity levels.

From our second analysis we found a small sentencing gap after we controlled for racial differences in offense and offender characteristics and that overall racial differences in offense characteristics accounted for most of that gap. However, we found that blacks were sentenced less harshly than whites on some extralegal factors but that they were sentenced more harshly than whites on the legally relevant guidelines offense severity scores. Overall we found relatively small racial differences in sentence lengths imposed, but we found mixed evidence on how the guideline offense severity scores affected sentences. First, they were associated with reducing the overall amount of the racial gap in sentencing that could be attributed to differences in ways the sentencing system evaluated the characteristics of whites and blacks. Second, given the overall racial sentencing gap with which they were associated, the scores also were correlated with relatively large differences in the way in which the characteristics of blacks and whites were evaluated or weighted by the sentencing system.

Racial Differences in Mean Lengths of Sentences Imposed on Convicted Bank Robbers

In its report, the Commission examined judicial sentencing patterns under the guidelines by looking, among other factors, at racial differences in the "sentence location" of similarly situated bank robbers, as well as other categories of defendants. (See the background section in app. II for a discussion of sentence location.) The Commission chose similarly situated defendants by matching offender characteristics. The Commission recognized that its analysis was based on a limited number of cases from a restricted portion of the guidelines sentencing table.

We adopted a slightly different strategy in this analysis. First, we analyzed data on all bank robbers sentenced in fiscal year 1990. Second, we defined similarly situated offenders for the specific offense of bank robbery by controlling for differences in criminal history categories and offense severity levels. We controlled for differences in criminal history categories by analyzing data within these categories and offense severity levels by using statistical techniques. Third, we analyzed the length of the sentence imposed (as compared to the categorical variable sentence location).

Approach

We estimated means and differences in means in order to explore and describe sentencing patterns associated with the race of robbery

defendants. We were not able to draw conclusions about the causes of the racial differences we uncovered, but we outlined areas that merited further investigation. In our analysis, we used data provided and used by the Commission in its report on the Sentencing Guidelines. The data included all eligible robbery cases sentenced in fiscal year 1990 (Oct. 1, 1989, through Sept. 30, 1990) that were identified in the Commission's monitoring and reporting system data file for fiscal year 1990 (MONFY90). Overall, the data consisted of 1,960 cases. For analysis, we selected defendants who were charged with bank robbery, Chapter Two Guideline Offenses 2B3.1.[2]

We divided the cases into two groups: an overall group and a group of defendants whose sentences fell within the minimum and maximum ranges prescribed by the guidelines. This latter group should not be confused with defendants who did not receive a sentencing departure. Sentences might fall outside the prescribed range associated with a particular set of criminal history and offense severity scores for reasons other than departures. For example, defendants might receive consecutive sentences, be subject to statutory maximum or minimum sentences, or receive sentence enhancements in addition to their sentences for their offense of conviction. None of these out-of-range cases would constitute a departure. Further, sentences falling outside of ranges would not, in and of themselves, constitute unwarranted disparity.

We grouped the cases by their criminal history categories. Within each category, we calculated the difference in means between sentences imposed on whites and blacks. We then repeated this analysis after adjusting sentence lengths for racial differences in offense severity levels. We repeated these steps for the entire group and the group of defendants whose sentences fell within the guidelines range.

We found, first, as table III.1 shows, racial differences in lengths of sentences imposed. For the entire group of fiscal year 1990 bank robbers, blacks received sentences in excess of 9 months more than whites. When we compared sentences for defendants convicted by plea or by jury trial, we also found differences. Blacks convicted by plea received sentences of almost 4 months more than those received by whites. Conversely, whites convicted by trial received much longer sentences than blacks. Inspection

[2]The data we received from the Sentencing Commission contained information on 1,960 bank robbery defendants sentenced in fiscal year 1990. The data had large numbers of missing values on variables of interest. For example, there were 409 cases with missing values for race, 580 missing values on final offense severity levels, and 582 with missing values for criminal history categories. We ultimately analyzed 1,451 cases. All of those defendants received prison sentences of between 1 and 995 months. The problems of missing data also affect the second analysis in this appendix, described later.

of these results shows racial differences that merit further scrutiny. For example, the differences by mode of disposition suggest that race and mode of disposition may interact, with blacks receiving longer sentences than whites on conviction by plea, but shorter ones than whites on conviction by trial.[3]

Table III.1: Mean Lengths of Sentences Imposed, Robbery, Fiscal Year 1990

	All defendants		Sentences falling within guideline ranges	
Race of defendants	Means and differences	N	Means and differences	N
White	91.51	949	80.14	481
Black	100.66	499	93.23	211
White/Black (difference)	–9.15		–13.08	
By disposition				
By plea				
White	80.57	671	70.54	332
Black	84.42	326	84.21	142
White/Black (difference)	–3.86		–13.67	
By jury trial				
White	196.01	69	177.96	27
Black	170.09	65	147.90	20
White/Black (difference)	25.92		30.06	

For robbers receiving within range sentences, blacks were given longer sentences than whites overall. This was also the case when blacks were convicted by plea; but, as with the entire group, the direction of the difference reversed for defendants convicted by jury trial.

Adjusted Means

In order to explore sentencing patterns further, we adjusted the means calculated above to control for criminal history and offense severity. We separated defendants into subgroups according to the six criminal history categories. To control for offense severity, we estimated adjusted mean sentence lengths imposed. We did so estimating a series of regressions, testing for race effects, and calculating adjusted means from our regressions. Specifically, we estimated three sets of regressions. First, we estimated a fully saturated regression of length of sentence imposed on final offense severity scores, a dummy variable for race (equal to 1 if

[3]The results of our second analysis also suggested that blacks received less favorable treatment than whites when convicted by guilty plea. On the other hand, our regressions did not produce evidence of racial differences in treatment on conviction by jury trial.

black, 0 otherwise), and an interaction term of race and final offense severity. Treating the groups' data as if they were a sample, we then tested whether to consider the white and black regression lines as coincident (i.e., that the race coefficient and the race by offense severity coefficient were simultaneously equal to zero).

If we rejected the idea of coincident lines, we then estimated a regression of sentence length on offense severity and race, testing for parallelism, that is, that the coefficient on race was equal to zero. This test could produce one of two results: parallel white and black lines or nonparallel lines. If the lines were nonparallel, further tests were performed to determine if the nonparallel lines had common or different intercepts. If we did not reject the null hypothesis of parallelism, we used the estimated sentence length equation to calculate white and black adjusted mean sentences. If we rejected the null hypothesis that the lines were parallel, we estimated regressions, separately for blacks and whites, of sentence length imposed on final offense severity score. These unrestricted models permitted us to test whether the effects of offense severity varied for blacks and whites with similar offense severity scores.

Results

Table III.2 displays the results of this analysis. We found for the entire group (the results in the first two sets of rows) relatively large racial gaps in sentences imposed in the unadjusted data in a number of criminal history categories. However, when we controlled for the effects of offense severity in the entire group, we found that the racial gap diminished by a small amount in most criminal history categories, but it diminished by about 15 months in criminal history category VI. The decrease in the racial gap in sentences in the adjusted data was consistent with the fact that blacks had, on average, slightly higher offense severity scores within categories.

Table III.2: Mean Lengths of Sentences Imposed by Criminal History Category, Robbery, Fiscal Year 1990

	Criminal history category											
	Category I		Category II		Category III		Category IV		Category V		Category VI	
	Means	N	Means	N	Means	N	Means	N	Means	N	Means	N
Unadjusted means entire group												
White	59.27	216	52.03	65	66.10	105	74.20	70	90.37	49	187.61	158
Black	67.26	84	66.34	32	70.06	51	92.63	41	118.06	18	203.66	90
White/Black	-7.99		-14.31		-3.96		-18.43		-27.69		-16.05	
Adjusted means, entire group												
White	59.54	216	52.39	15	66.99	105	74.85	70	90.30	49	189.81	158
Black	66.55	84	66.31	32	68.22	51	91.27	41	118.09	18	190.92	90
White/Black	-7.01		-13.92[a]		-1.23		-16.42[b]		-27.79[a]		-1.11	
Unadjusted means, within range sentences												
White	40.64	149	44.75	51	51.17	82	61.96	51	79.00	34	174.91	111
Black	46.96	55	47.32	22	56.41	32	70.44	25	76.29	14	182.77	61
White/Black	-6.32		-2.57		-5.24		-8.48		2.71		-7.86	
Adjusted means, within range sentences												
White	40.64	149	45.38	51	51.10	82	62.94	51	78.81	34	180.55	111
Black	46.76	55	45.84	22	56.40	32	68.40	25	76.03	14	172.28	61
White/Black	-6.12[a]		-0.46		-5.30[a]		-5.46[b]		1.98		8.27	

[a]Separate black and white slopes and intercepts.

[b]Common black and white slopes, difference intercepts.

The adjusted data for the entire group also showed that the sentencing patterns associated with race varied across criminal history categories in the sentencing table. For example, the regressions uncovered race and offense severity interactions in categories II and V, a constant race effect in category IV, and no race effects in the other categories. The separate regressions needed to estimate the effects in categories II and V indicated that blacks tended to receive shorter sentences than whites at the lower end of the offense severity scale but that they received longer sentences than whites at the higher end of offense severity. On the other hand, in criminal history category IV, race did not interact with offense severity; rather, blacks received longer sentences than whites throughout the range of offense severity.

The results for the within range sentences also indicated that the effects of race on sentencing outcomes varied across criminal history categories.

These results (reported in the third and fourth sets of rows of table III.2) of the unadjusted means showed larger racial differences in sentences imposed in the first, third, fourth, and sixth criminal history categories as compared to the second and fifth. On the other hand, the adjusted data tended to reduce the size of the racial gap slightly, and they also revealed the same types of race and offense severity interactions described before, with the notable exception of category VI. In category VI, whites tended to receive longer sentences than blacks. On the other hand, in criminal history categories I and III, blacks tended to receive shorter sentences than whites at the lower end of the offense severity scale but longer sentences at the higher end. In category IV, blacks received longer sentences throughout.

Interpretation

These results provided descriptions of patterns of racial differences in sentencing. To the extent that offense severity scores and criminal history categories equate offense conduct and criminal histories across different offenses—that is, they define similarly situated offenders—the racial disparities that appeared in various portions of the sentencing table merit further investigation for two reasons: first, to uncover the reasons for the disparities, and second, to explain why the patterns appeared to vary across categories. These disparities may be due to racial differences in other legally relevant factors, a number of which we did control for in the second analysis described in this appendix. For example, for the entire group, part of the differences might be due to legitimate factors such as departures or mandatory or consecutive sentences. For the group of defendants whose sentences fell within the range prescribed in the guidelines, other legally relevant factors, such as presentencing selection decisions (e.g., conviction by plea or trial) might explain the differences.

Further, the inconsistent race effects across criminal history categories after controlling for offense severity compounded the problem of explaining the racial differences. It was not immediately clear why the race effects should vary as they did across history categories, that is, there is no immediately apparent theory for such patterns. For example, in the adjusted data for the entire group, the race offense severity interaction in category II disappeared in category III, changed to a simple race effect in category IV, and finally reappeared in category V. Whether these patterns were due to specific aspects of offense conduct not observed in the MONFY90 data, to racial disparities, or to both, suggested that further effort was needed to understand these patterns.

Explaining the Racial Gap in Sentences Imposed: Testing for Racial Differences in Offense and Offender Characteristics and in the Effects of Characteristics on Sentence Lengths

Objectives

In this second analysis of racial differences in sentences imposed on bank robbers, we attempted to distinguish between two portions of the racial disparity in sentences imposed: the portion due to racial differences in offense and offender characteristics (characteristics) versus the portion due to differences in coefficients (i.e., weights of those characteristics). By racial differences in status on offense and offender characteristics we mean racial differences in the distribution of legally relevant factors—such as offense severity or criminal history scores—or extralegal factors—such as gender or socio-economic status. By racial differences in coefficients, we mean differences arising from the way in which the federal sentencing system evaluated or weighed the race-specific characteristics. More generally, we attempted to test whether there are racial differences in the way in which the characteristics of similarly situated offenders were weighted by the federal sentencing system, as defined by the variables in our models.

Data and Methods

We used the entire group of cases that we analyzed in our preceding analysis; we did not perform separate analyses on the within range and entire group of cases. As before, we treated the cases of bank robbers sentenced in fiscal year 1990 as if they were a sample. We estimated separate regressions of sentences imposed for blacks and whites. We tested for differences in the coefficients and then used the information from the two equations to calculate the difference between the actual sentences blacks served and those they could have been expected to serve if their coefficients were the same as whites. We borrowed our method for this test from research in the economics, sociology, and criminology

literature that has attempted to measure discrimination. Researchers have used the method—the residual difference test—because it measures discrimination as the residual left after controlling for relevant explanatory variables and assessing differences in the coefficients on those variables between the disadvantaged and advantaged groups.[4] We employed the test in our analysis as a means to isolate the various portions of racial disparities—due to average characteristics versus due to how those characteristics are weighed—and not to measure discrimination.

We were unable to estimate a definitive discrimination model for a variety of reasons. For example, omitted, legally relevant variables (such as demeanor at sentencing) might have reduced or increased the residual gap or the estimated amount attributed to differences in the coefficients further. Second, the small number of cases precluded our investigating patterns among other meaningful subgroups, such as women versus men. The logic of this method suggested separate regressions for each meaningful subgroup. We were unable to do this for groups other than blacks and whites where differences in the coefficients might have arisen. We did not look at, for example, race/gender interaction effects.[5]

At the same time, this method permitted us to distinguish between effects of characteristics and the effects of the manner in which those characteristics were weighed by sentencers. Thus, the method assessed both the differences in coefficients attributable to race and to congressional interest in whether similarly situated offenders receive similar sentences.

Overview of the Residual Difference Test

The method builds upon the logic of using separate regressions to estimate the race by offense severity interactions in our previous analysis of sentence lengths. Here, however, we attempted to control for as many legally relevant variables as possible in our regressions. We did so in order to diminish the unexplained portion of the regression and thus reduce the residual that could be attributed to racial differences.

[4]For examples of the use of this method to examine wage discrimination, see Alan S. Blinder, "Wage Discrimination: Reduced Form and Structural Estimates," Journal of Human Resources, Vol. VIII, No. 4 (1973), pp. 436-455, or, more recently, Rachel A. Rosenfeld and Arne L. Kalleberg, "A Cross-National Comparison of the Gender Gap in Income," American Journal of Sociology, Vol. 96, No. 1 (1990), pp. 69-106. For a slight variation on Blinder's approach and an example of the use of this method to examine discrimination in sentencing, see Samuel L. Myers, Jr., "Statistical Tests of Discrimination in Punishment," Journal of Quantitative Criminology, Vol. 1, No. 2 (1985), pp. 191-218.

[5]The small number of cases also precluded our estimating differences between whites and Hispanics or blacks and Hispanics.

We followed closely the procedure as outlined by Blinder (see footnote 4), for demonstration purposes. We began by estimating separate regressions for whites and blacks for the dependent variable, the natural log of the length of sentence imposed, as indicated in equations (1) and (2):

$$(1) \qquad S_i^w = b_o^w + \sum_i b_i^w X_i^w + u_i^w$$

$$(2) \qquad S_i^b = b_o^b + \sum_i b_i^b X_i^b + u_i^b$$

The respective equations for whites and blacks are denoted by the superscripts "w" and "b." The right-hand sides of equations 1 and 2 consist of vectors of the same explanatory variables, denoted by the "X"s, and parameter values, denoted by the "b"s. We included measures of legally relevant variables (such as offense conduct, victim injury, etc.) and extralegal factors (such as age or gender). By estimating the equations separately and comparing sets of coefficients, we estimated first-order interactions between race and each of the independent variables. This approach differs from methods of estimating race effects that rely on a single regression with a single dummy variable for race to measure all range effects. The single dummy variable approach, while useful for increasing the efficiency of parameter estimates, has two major drawbacks. First, it constrains the effects of race to be equal to the difference in intercepts between the two equations, that is, to the unexplained portion of the regression. The single dummy variable approach also precludes considering the complete set of first-order race interactions. Second, even if a fully saturated model were estimated, in the dummy variable approach the error variances of the separate equations are equal.[6]

Given equations 1 and 2, we computed a raw differential, the difference between equations 1 and 2:

[6]Constraining error variances across the equations to be equal has implications for hypothesis testing in the separate white and black equations. If, for example, the error variance in the pooled regression is less than the error variance in the black equation, then one is more likely to reject the null hypothesis of no effect (of a particular variable) when in fact the null hypothesis is correct. That is, there is an increased likelihood of incorrectly inferring that a variable has an effect on the sentencing outcomes for blacks when in fact it does not.

(3) $$S_i^w - S_i^b = b_o^w + \sum_i b_i^w \bar{X}_i^w - (b_o^b + \sum_i b_i^b \bar{X}_i^b)$$

This can be broken down into the portion of the difference explained by the regression:

(4) $$\sum_i b_i^w \bar{X}_i^w - \sum_i b_i^b \bar{X}_i^b$$

and the portion explained by the differences in intercepts:

(5) $$b_o^w - b_o^b$$

The differences in intercepts reflected one component of the overall racial differences in sentences. The differential due to the regression, equation 4, can be further broken down into group differences in average characteristics and differences in the coefficients, i.e.,

(6) $$\sum_i b_i^w (\bar{X}_i^w - \bar{X}_i^b) + \sum_i \bar{X}_i^b (b_i^w - b_i^b)$$

The differences in the coefficients are analogous to the concept of similarly situated offenders receiving different treatment. The components of equation 6 have the following interpretations: the first sum is the value of the advantage in characteristics for whites (perhaps, for example, shorter criminal histories) as evaluated by the whites' sentence length equation. The second sum is the difference between the sentences blacks would have received if they received whites' treatment (i.e., had the same

coefficients as whites) and the actual sentences they received. The second sum exists only if the criminal justice system evaluates differently the identical bundle of characteristics depending upon whether they are possessed by blacks or whites. If there are no differences in coefficients, that is, if the criminal justice system evaluates bundles of traits identically, then the racial gap reduces to the unexplained portion of the regression due to the difference in intercepts.

In sum, the measures used in the analysis are the raw differential, as expressed in equation 3; the portion of the raw differential attributable to racial differences in characteristics or traits, as given by the first sum in equation 6; the portion of the differential attributable to racial differences in weighting of characteristics (that is, dissimilar treatment of similar defendants), as given by the second sum in equation 6; the portion of the differential unexplained, that is, attributed to the differences in intercepts, as given by equation 5; and the total differential due to the differential weighting of variables by sentencers as given by the sum of the differences in coefficients and the differences in intercepts.

While we had theoretical and conceptual support for estimating separate equations, we performed a statistical test to determine whether we gained additional information by partitioning the data and estimating separate models as compared to pooling the data and estimating a single equation. That is, we performed a Chow test[7] to determine if there were differences in the set of coefficients of the separate regressions. Note, however, that even if the Chow test provided support for estimating separate equations by race, it did not permit us to identify differences due to the separate factors in the models. In other words, by simply performing the Chow test, it was possible to find overall differences in the coefficients but no disparity because of offsetting effects. Therefore, it was necessary to perform the decomposition we described earlier.[8]

Analytic Approaches

As stated previously, our primary objective in this analysis was to determine the extent to which racial disparities in robbery sentencing under the guidelines could be attributed to differences in the manner in

[7]See Gregory C. Chow, "Tests of Equality between Sets of Coefficients in Two Linear Regressions," Econometrica, Vol. 28 (1960), pp. 591-605. For an overview of the test, see, for example, Robert S. Pindyck and Daniel L. Rubinfeld, Econometric Methods and Economic Forecasts, 2nd Edition, (New York: McGraw-Hill, 1976), pp. 123-126.

[8]For comparison with the results of the separate equations, we also estimated a single-equation with a dummy variable for race and a complete set of race interactions. We performed F-tests to determine whether the race effects in the pooled models were significant.

which the federal sentencing system evaluated the same bundle of characteristics depending upon whether they belong to whites or blacks. As a secondary objective, however, we attempted to determine the extent to which legally relevant factors other than the guidelines scores played a key role in sentencing outcomes, particularly racial differences in sentence lengths. In short, we attempted to address questions such as what nonguideline variables continue to play a role and are they legally relevant or extralegal factors?

To accomplish these objectives, we built two sets of models. In each, we estimated separate equations for blacks and whites and performed the tests outlined in equations 1 through 6. We first estimated a set of models that did not include measures for the guidelines scores—our "unconstrained" models. These models identified how variables similar to those used in constructing the guidelines were weighted in sentencing decisions, and these models were used primarily to compare with our second set of models. The second set of models—our guidelines scores models—contained measures for the final offense severity and criminal history categories prescribed by the guidelines, as well as other legally relevant variables that were statistically significant in either the black or white equations. By comparing results for these models, we were able to assess certain effects of the guidelines.

Variables

We used a number of measures of legally relevant and extralegal factors. We grouped these variables into categories of factors associated with sentencing outcomes. Our categories included those factors associated with specific offense conduct, ongoing criminal behavior, conviction information, information about sentences, relevant criminal history information, extralegal factors, statutory minimum sentences, the circuit in which defendants were sentenced, and the criminal history and offense level scores under the guidelines. The variables used in the final models follow, beginning with the dependent variable:

1. Dependent Variable

- natural log of the length of sentence imposed, in months

2. Specific Offense Conduct Variables

- a dummy variable for the major offense code, or primary offense charged

- dummy variables for the number of specific offense conduct points associated with the dollar amount stolen[9] (the omitted category included dollar values that were associated with three or more guidelines' specific offense conduct points)
- a dummy variable to indicate the presence, threat, or use of a weapon
- a dummy variable to indicate involvement, i.e., more culpability in the offense, such as leader or supervisor
- a dummy variable to indicate victim injury

3. Ongoing Criminal Behavior Variables

- a dummy variable that indicated whether the offense was one of multiple acts
- a dummy variable to indicate whether the offense was part of ongoing criminal behavior (the omitted category for these two variables was that the offense was a single act)
- a dummy variable to indicate that the defendant was under criminal justice status at the time of arrest, e.g., on bail

4. Information About the Conviction

- the number of counts of conviction
- a dummy variable to indicate that the defendant was convicted by a guilty plea

5. Information About Sentences

- a dummy variable to indicate that consecutive sentences were imposed
- a dummy variable to indicate an upward departure
- a dummy variable to indicate a downward departure (including downward departures for substantial assistance)

6. Criminal History Information

- the number of adult convictions
- the number of prior robbery convictions
- a dummy variable to indicate whether the defendant had prior revocations of supervision
- the number of times the defendant was incarcerated for 5 or more years

[9]We used dummy variables for specific offense conduct points rather than the dollar value of the amount stolen because defendants were sentenced according to two versions of the guidelines. In those versions, different dollar amounts were associated with different specific offense conduct points, and consequently, different final offense severity scores and different sentence lengths.

- the number of times the defendant was incarcerated for more than 1 and fewer than 5 years

7. Extralegal Factors

- the number of months free since the defendant's previous incarceration (up to 5 years free and including those never incarcerated)[10]
- the defendant's age in years
- a dummy variable to indicate the defendant's gender

8. Statutory Minimum

- a dummy variable to indicate that the defendant was subject to a statutory minimum sentence

9. Circuit in Which Sentenced

- a series of dummy variables to identify the circuit in which the defendant was sentenced

10. Guidelines Scores

- the defendant's final offense severity score
- a squared term on final offense severity to test for nonlinear effects
- a series of dummy variables to indicate the defendant's criminal history category

In addition to these variables, we tested for the effects of a number of other measures of both legally relevant and extralegal variables. However, these measures were not significant in the equations we estimated. If they had very low t-values, indicating that they had no unique effect on sentence length outcomes, controlling for the other variables, we dropped these variables from the remaining analysis. The dropped variables included such factors as education level, marital status, and months employed in the previous year.[11]

[10]Those free more than 5 years received a value of 61 months. Those never incarcerated received a value of 97 months. We also tested for the effects of a dummy variable that coded whether or not a person had ever been incarcerated. That variable proved not to be significant in our models.

[11]We were unable to test for the influence of some extralegal factors because of poor data. For example, defendants' income, as an indicator of class status, was missing in over 91 percent of the cases available for analysis.

Specification and Estimation Issues

A number of issues potentially biased the estimates of the racial gap in sentencing. Two important ones were selection and omitted variable bias. A third, measurement error, was a problem we could do little about in this secondary analysis.

Selection Effects

The absence of data on the presentence stages potentially introduced a selection effect. If, for example, prosecutors are less likely to pursue cases against certain groups of offenders, e.g., against blacks who commit less serious crimes, then our estimates of the racial gap in sentencing might be biased. Without observations on the presentencing stages that would determine the type of selection, it was not possible to determine the direction of the bias. If, for example, there was discrimination against blacks and against those who committed more serious offenses, then the whites who were sentenced would have committed, on average, more severe offenses than all whites; consequently, we would have underestimated the racial gap in sentencing. The reverse holds if there was discrimination against whites.

Omitted Variable Bias

If legally relevant variables were omitted, then the estimates of the effects of extralegal factors would be biased. We attempted to control for this by including—at the risk of introducing multicollinearity—as many legally relevant variables as possible.[12]

Results

As discussed above, we estimated two sets of models of the natural log of length of sentence imposed.[13] The first set of models—the unconstrained version—excluded guidelines scores; the second set contained them. The race-specific means and standard deviations of the variables used in the respective models are given in tables III.3 and III.4. Inspection of the data in table III.3 (for the unconstrained models) reveals few large differences in the characteristics of black and white robbers. Slightly lower fractions of blacks had robbery as their major offense, used a weapon, were involved in their offense of conviction as part of ongoing criminal

[12]Multicollinearity would not be a problem if we were not interested in individual parameter values. We therefore conducted a number of tests for multicollinearity. We followed the diagnostic procedures outlined in David A. Belsley, Edwin Kuh, and Roy E. Welsch, Regression Diagnostics: Identifying Influential Data and Sources of Collinearity (New York: John Wiley and Sons 1980). They suggest a diagnostic procedure consisting of a double condition: (1) a singular value judged to have a high condition index, and which is associated with (2) high variance-decomposition proportions for two or more estimated regression coefficient variances (p. 112). In general, our data were not ill conditioned, and multicollinearity was judged not to be a problem.

[13]We chose this specification because it provided a better fit than did the levels of sentences imposed. The coefficients on the variables are interpreted as the proportionate change in the dependent variable per unit change in an independent variable.

behavior, were convicted by plea, or were female. On the other hand, slightly higher fractions of blacks were under criminal justice status at the time of arrest, had consecutive sentences imposed, or were subject to a statutory minimum sentence. In addition, blacks had slightly fewer previous adult convictions or previous incarcerations of more than 5 years, but blacks had more prior robbery convictions.

Table III.3: Means and Standard Deviations in Models Without Guidelines Scores

Means and standard deviations of variables used in the regressions	White		Black	
	Means	Standard deviation	Means	Standard deviation
Natural log of sentence length—months—imposed	4.213	0.777	4.312	0.771
Major offense—robbery	0.949	0.220	0.911	0.285
Dollar value stolen, 1 specific offense conduct[a] point	0.336	0.473	0.337	0.473
Dollar value stolen, 2 specific offense conduct[a] points	0.182	0.386	0.186	0.389
Weapon use	0.423	0.494	0.415	0.493
Involvement/culpability	0.131	0.337	0.116	0.320
Multiple acts	0.294	0.456	0.300	0.459
Ongoing behavior	0.141	0.348	0.101	0.302
Criminal justice status at arrest	0.514	0.500	0.521	0.500
Number of counts of conviction	1.978	1.924	2.010	1.579
Convicted by guilty plea	0.764	0.425	0.674	0.469
Consecutive sentences imposed	0.182	0.386	0.219	0.414
Upward departure	0.043	0.203	0.031	0.173
Number of adult convictions	3.627	3.106	3.225	2.862
Number of prior robbery convictions	0.577	1.159	0.725	1.303
Prior revocations of supervision	0.432	0.496	0.411	0.493
Number of incarcerations >= 5 years	0.523	1.156	0.517	1.093
Number of incarcerations 1 - 5 years	0.870	1.440	0.952	1.497
Months free since previous incarceration	51.993	40.457	48.924	40.393
Gender	0.047	0.211	0.035	0.184
Statutory minimum	0.200	0.400	0.262	0.440
Circuit 2	0.024	0.153	0.097	0.296
Circuit 7	0.042	0.201	0.066	0.249
Circuit 9	0.416	0.493	0.209	0.407
Number of observations	880		484	

[a]For additional details on the definition of the variable, see footnote 9 in the text.

In short, there were few obvious differences between the average characteristics of blacks and whites that would lead one to predict overall longer sentences for blacks. There were, however, indications that on some important legally relevant variables, such as prior robbery convictions, statutory minimum sentences, and the smaller fraction of blacks convicted by plea, the black/white differences in characteristics provided legitimate reasons for the longer average sentences imposed on blacks.

Table III.4 provides the means and standard deviations for the variables analyzed in the regressions that included measures of guidelines scores. One difference between this table and the previous one is the different set of independent variables. Absent from the second set of means and standard deviations are measures on primary offense, dollar amount, involvement, multiple acts, ongoing behavior, criminal justice status at arrest, adult convictions, prior revocations, and previous incarcerations. Variables not in the first set of models but included in these are the guidelines measures, victim injury, downward departures, and age. The different set of variables in the respective models arose from our efforts to find the best set of predictors for each equation.[14]

[14]We developed our unconstrained and guidelines scores models independently. In both cases, we attempted to find the best set of predictor variables, regardless of which variables were included in the other models. We retained variables that were statistically significant in either the black or white equation and dropped those that were not. In this way, we did not force variables into the guidelines scores models. Rather, we permitted the scores to purge the models of redundant legally relevant variables. We also permitted all extralegal factors to play a role, dropping those that were not statistically significant.

Table III.4: Means and Standard Deviations in Models Including Guidelines Scores

Means and standard deviations in models of variables used in the regressions	White		Black	
	Mean	Standard deviation	Mean	Standard deviation
Natural log of sentence length—months—imposed	4.238	0.788	4.408	0.809
Weapon use	0.448	0.498	0.452	0.499
Victim injury	0.029	0.168	0.047	0.211
Number of counts of conviction	2.036	2.161	2.040	1.687
Convicted by guilty plea	0.769	0.422	0.668	0.472
Consecutive sentences imposed	0.190	0.393	0.252	0.435
Upward departure	0.051	0.220	0.040	0.196
Downward departure	0.073	0.260	0.047	0.211
Number of prior robbery convictions	0.593	1.223	0.781	1.349
Months free since previous incarceration	53.660	40.282	47.120	40.050
Age	34.027	10.215	30.628	7.742
Statutory minimum	0.175	0.380	0.249	0.433
Circuit 1	0.029	0.168	0.037	0.188
Circuit 2	0.019	0.135	0.103	0.304
Circuit 9	0.368	0.483	0.136	0.344
Final offense severity score	22.426	5.104	23.106	5.549
Final offense severity score squared	528.942	252.212	564.595	279.094
Criminal history III	0.161	0.368	0.163	0.370
Criminal history IV	0.104	0.305	0.136	0.344
Criminal history V	0.070	0.255	0.050	0.218
Criminal history VI	0.238	0.426	0.282	0.451
Number of observations	589		301	

The differences in average characteristics in table III.4 for the guidelines scores models are consistent with those found on the previous table. Additional information of interest included the fact that black bank robbers were younger than whites and had slightly higher average offense severity scores than whites.

Tables III.5 and III.6 report the results of our regressions of the natural log of sentence length on the independent variables for the separate black and white equations for each version of the model. Both sets of models fit the data reasonably well. As indicated by the adjusted R-square at the bottom of table III.5, the unconstrained models explained around 60 percent of the variation in the dependent variable for each of the equations. (They explained slightly more than 60 percent in the white equation and slightly

less than 60 percent in the black equation.) The guidelines scores models did even better, explaining over 80 percent of the variation in the dependent variable, as indicated by the adjusted R-squared in table III.6. Thus, in both cases, we have relatively strong sets of predictors.

Table III.5: Regression Results, Models Without Guidelines Scores, Dependent Variable: Natural Log of Sentence Length Imposed

	White		Black	
Independent variable	**Parameter estimate**	**t-statistic**	**Parameter estimate**	**t-statistic**
Intercept	3.460	33.740	3.729	30.180
Major offense—robbery	0.211	2.809	0.062	0.696
Dollar value stolen, 1 specific offense conduct[a] point	–0.008	–0.214	–0.132	–2.449
Dollar value stolen, 2 specific offense conduct[a] points	0.082	1.684	0.029	0.424
Weapon use	0.283	7.580	0.324	6.219
Involvement/culpability	0.127	2.517	0.002	0.031
Multiple acts	0.086	2.115	0.157	2.671
Ongoing behavior	0.209	3.878	0.213	2.472
Criminal justice status at arrest	0.087	2.355	0.101	1.935
Number of counts of conviction	0.085	8.703	0.062	3.351
Convicted by guilty plea	–0.101	–2.455	–0.042	–0.771
Consecutive sentences imposed	0.321	6.127	0.378	5.349
Upward departure	0.133	1.630	0.269	1.980
Number of adult convictions	0.027	3.409	0.038	2.842
Number of prior robbery convictions	0.118	6.794	0.064	2.946
Prior revocations of supervision	0.100	2.334	0.079	1.293
Number of incarcerations >= 5 years	0.113	6.362	0.115	4.460
Number of incarcerations 1 - 5 years	0.053	3.597	0.035	1.638
Months free since previous incarceration	–0.002	–4.223	–0.003	–4.032
Gender	–0.181	–2.335	–0.220	–1.710
Statutory minimum	0.137	2.637	0.226	3.222
Circuit 2	–0.051	–0.471	–0.256	–2.920
Circuit 7	–0.074	–0.871	–0.287	–2.917
Circuit 9	–0.132	–3.450	–0.093	–1.540
Mean dependent variable natural log of sentence length imposed	4.213		4.312	
F-value	64.229		29.607	
Probability>F	0.0001		0.0001	
R-square	0.633		0.597	
Adjusted R-square	0.623		0.577	
N	880		484	

[a]For details on variable definition, see footnote 9 in the text.

Table III.6: Regression Results, Models With Guidelines Scores, Dependent Variable: Natural Log of Sentence Length Imposed

	White		Black	
Independent variable	**Parameter estimate**	**t-statistic**	**Parameter estimate**	**t-statistic**
Intercept	0.442	1.611	0.470	1.041
Weapon use	0.148	5.238	0.133	3.009
Victim injury	–0.195	–2.607	0.110	1.155
Number of counts of conviction	0.049	7.995	0.053	3.995
Convicted by guilty plea	–0.093	–3.056	–0.008	–0.188
Consecutive sentences imposed	0.243	5.805	0.242	3.902
Upward departure	0.283	4.970	0.382	3.836
Downward departure	–0.290	–6.092	–0.323	–3.485
Number of prior robbery convictions	0.027	2.133	0.030	1.803
Months free since previous incarceration	–0.001	–1.795	–0.002	–2.995
Age	0.005	3.619	–0.0001	–0.044
Statutory minimum	0.344	7.691	0.378	5.757
Circuit 1	0.168	2.233	–0.080	–0.719
Circuit 2	–0.098	–1.092	–0.199	–3.031
Circuit 9	–0.062	–2.243	–0.032	–0.550
Final offense severity score	0.210	9.202	0.234	6.286
Final offense severity score squared	–0.003	–5.855	–0.003	–4.303
Criminal history III	0.197	4.819	0.021	0.312
Criminal history IV	0.310	5.964	0.173	2.091
Criminal history V	0.536	9.094	0.266	2.542
Criminal history VI	0.661	12.511	0.528	6.027
Mean dependent variable natural log of sentence length imposed	4.238		4.408	
F-value	184.741		77.907	
Probability>F	0.0001		0.0001	
R-square	0.867		0.848	
Adjusted R-square	0.862		0.837	
N	589		301	

Although these models fit the data relatively well, our interest was in whether or not there were differences in the coefficients between the black and white equations in each set of models, that is, whether or not blacks' and whites' legally relevant and extralegal characteristics were evaluated similarly by the sentencing system. To test for differences in the coefficients between the race-specific equations, we conducted Chow

tests on the unconstrained and guidelines scores models. The null hypothesis is that there is no difference between the sets of coefficients. We obtained different results for the two models. For the unconstrained models, we were not able to reject the null hypothesis of no difference in coefficients between the black and white equations. On the contrary, for the guidelines scores models, we rejected the null hypothesis of no difference in coefficients between the black and white equations. This finding of no difference between coefficients in the unconstrained models and differences between the coefficients in the guidelines scores models was somewhat surprising, given that the guidelines were intended to reduce disparity. However, from the Chow test we employed, we cannot determine what individual factor(s) may be responsible for the differences in coefficients, nor can we identify possible offsetting effects. To do that, we must proceed with the decomposition.[15]

General Effects

Given the results of the Chow tests, we report and discuss the results of our unconstrained models only to provide a comparison with the results of our guideline score models. Tables III.5 and III.6 report the regression results for each set of models. In the unconstrained models on table III.5 as expected, legally relevant factors such as weapon use, ongoing criminal behavior, the imposition of consecutive sentences, and statutory minimum prison requirements had large and roughly equivalent effects for both blacks and whites. There appeared to be some racial differences in the effects of other legally relevant variables, such as whether the major offense charged was robbery, levels of involvement in the crime, conviction by guilty plea, upward departures, and the prior revocation of supervised release. Of the extralegal factors that had any effects, the effects of gender rivaled in size the effects of some legally relevant factors. The circuit in which defendants are sentenced had an effect such that persons sentenced in the 2nd, 7th, or 9th were more likely to receive shorter sentences than those in other circuits.

By comparison, the regression results for the guidelines scores models showed meaningful differences from the unconstrained models, as reported in table III.6. There are a number of effects. First, 13 variables that appeared in the unconstrained model were dropped from the

[15]As an additional test, we estimated fully saturated models with a dummy variable for race and a complete set of interaction terms. We then tested the null hypothesis that the joint effects of race were equal to zero. We obtained results that ran parallel to those obtained from the Chow test. For the unconstrained models, we were unable to reject the null of hypothesis of no race effect. For the guidelines scores models, we rejected the null hypothesis of no race effects. Recalling our caveats about pooled regressions of this sort, we mention the results of this test here, but we do not discuss it as a means to examine the effects of specific variables on sentencing outcomes.

equations in the guidelines models because they were not statistically significant and because their t-values were so small.[16] Second, other than the guidelines scores themselves, victim injury, downward departure, age, and the 1st Circuit appeared in the guidelines scores models, but they did not appear in the unconstrained models. Third, the effects of upward and downward departures and statutory minimum sentences increased relative to the unconstrained models. Fourth, the effects of the remaining legally relevant variables diminished for both blacks and whites relative to the unconstrained models. Fifth, the magnitude of the effects of extralegal factors diminished.

In the guidelines scores models, the increase in the effects of departures and the increase in magnitude of effects on statutory minimums were not unexpected. These factors were not incorporated explicitly into the guidelines scores. The guidelines scores helped to differentiate between their effects and the effects of all other legally relevant variables. Relative to the less differentiated measures in the unconstrained models, the guidelines scores were strongly related to lengths of sentences that did not fall outside guidelines ranges.

Departure, consecutive, and statutory minimum sentences all may fall outside the prescribed ranges.

The diminution in the effects of other legally relevant variables was consistent with the increase in the magnitudes on departures, statutory minimums, and consecutive sentences. Much of the information about legally relevant factors was incorporated into the guidelines scores, so the remaining effects of the other legally relevant variables might measure aspects of sentencing decisions not yet structured into the guidelines. The differences between the information in the guidelines scores and the other legally relevant variables could identify that. The effects of other legally relevant variables could also identify where judicial discretion under the guidelines was likely to operate. We turn to that issue next, particularly as it relates to racial differences in sentencing.

Decomposition of Race Effects

To determine the extent to which racial differences in the distribution of offense and offender characteristics versus the effects of the independent

[16]The variables dropped included the measures for the legally-relevant variables of primary offense charged, dollar value of the amount stolen, involvement, multiple acts, ongoing behavior, criminal justice status at arrest, the number of adult convictions, whether a defendant had prior supervised release revoked, and both measures of prior incarcerations. The extralegal variables dropped included gender, and whether sentenced in the 7th Circuit. We dropped these variables from our final models that contained controls for legally relevant variables.

variables explain the racial sentencing gap, see tables III.7 and III.8. (Detailed tables showing the effects of the decomposition on a variable-by-variable basis are provided in tables III.9 and III.10) Table III.8 summarizes the results of the decomposition for the models with the guidelines scores by groups of factors. We will refer to the results of the decomposition for the unconstrained models in table III.7 only for comparative purposes.[17] The first column of table III.8 shows the gross or raw differential, as calculated by equation 3. The value in the first column is interpreted as the percentage difference in the raw differential. A negative sign on that value indicates that blacks were sentenced more harshly than whites on a respective factor. Conversely, a positive sign in that column indicates that blacks were sentenced less harshly than whites. In table III.8, for example, the value of the raw differential for specific offense conduct would indicate a five-tenths of a percent disadvantage in sentencing for blacks on that factor.

The values in the second column (which were calculated by the first sum in equation 6) show the portion of the raw differential that is due to the differences in average characteristics between whites and blacks. The values in the third column (which were calculated by the second term in equation 6) show the portion of the raw differential attributable to differences in coefficients. This portion represents the difference between the actual sentences of blacks and the sentences they would receive if they had the same coefficients as whites. In table III.8 the value in the third column on specific offense conduct (–0.8), for example, would indicate that blacks received eight-tenths of a percent longer sentences on this factor than they would if they had the same coefficients as whites.

Finally, before addressing the results in table III.8, consider the meaning of the signs in each column and the fact that there can be offsetting effects of legally relevant and extralegal factors.

The positive and negative signs in the second and third columns can have different interpretations depending upon whether one looks at legally relevant or extralegal variables. Consider the white/black differences in coefficients on legally relevant variables, as indicated in the third column of table III.8. In general, a positive in column three indicates a positive value on the white/black difference in coefficients on legally relevant variables, which indicates that whites received more severe sentences than blacks. A negative value on the white/black difference in coefficients

[17]Recall that the Chow test showed no difference between the black and white coefficients in these models.

on legally relevant variables in the third column indicates harsher sentences for blacks.

For extralegal variables, the interpretation of the signs in the third column of table III.8 might change. For example, the positive value on the white/black difference in the coefficients for extralegal variables (e.g., the fact that black robbers in these data were younger than their white counterparts) might constitute an advantage in characteristics for blacks on them.

Table III.7: Analysis of the Black/White Sentence Differential: Models Without Guidelines Scores

Causal factor	Raw differential	Portion attributed to differences in characteristics	Portion attributed to differences in coefficients
Intercept	-26.9		
Specific offense conduct	19.6	1.2	18.4
Ongoing criminal behavior	-2.2	0.7	-2.9
Conviction stage	-0.4	-1.2	0.8
Sentencing stage	-2.7	-1.0	-1.7
Criminal history information	2.1	-0.8	2.9
Extralegal factors	3.5	-0.9	4.4
Statutory minimums	-3.2	-0.9	-2.3
Circuit	0.4	-2.2	2.6
Offense severity scores	N/A	N/A	N/A
Criminal history category	N/A	N/A	N/A
Subtotal, without intercept	17.1		
Difference in intercepts	-26.9		
Total, i.e., raw differential	-9.8		
Amount attributable to characteristics		-5.1	
Amount attributable to coefficients			22.2
Amount due to differences in coefficients plus intercepts	-4.7		
Percentage due to treatment differences	48.0		

N/A = Not applicable.

Table III.8: Analysis of the Black/White Sentence Differential: Guidelines Scores Models

Causal factor	Raw differential	Portion attributed to differences in characteristics	Portion attributed to differences in coefficients
Intercept	-2.8		
Specific offense conduct	-0.5	0.3	-0.8
Ongoing criminal behavior	0.0	0.0	0.0
Conviction stage	-7.3	-1.0	-6.4
Sentencing stage	-2.2	-2.0	-0.2
Criminal history information	-0.7	-0.5	-0.2
Extralegal factors	22.6	1.0	21.6
Statutory minimums	-3.4	-2.6	-0.8
Circuit	0.8	-0.7	1.5
Offense severity scores	-30.4	-4.3	-26.0
Criminal history category	6.9	-2.9	9.8
Subtotal, without intercept	-14.2		
Difference in intercepts	-2.8		
Total, i.e., raw differential	-17.0		
Amount attributable to characteristics		-12.7	
Amount attributable to coefficients			-1.5
Amount due to differences in coefficients plus intercepts	-4.3		
Percentage due to treatment differences	25.3		

Because of these differences in interpretations of variables, there could be offsetting effects. Consequently, the overall results need to be interpreted with caution.

Returning to the overall results for the guidelines scores models in table III.8, we found 17 percent difference in the favor of whites in the raw differential (as calculated by equation 3). Slightly less than 3 percent of the white advantage was unexplained or due to the differences in intercepts. Thus, most of the white advantage was due to differences in the regression. In addition, a very small amount of the white advantage was due to overall racial differences in sentences. Only 1.5 percent of the white advantage came from differences in treatment. When that was added to the 2.8 percent of the amount attributable to the difference in intercepts, the total white advantage amounted to slightly more than 4 percent (4.3 percent). Of the overall 17 percent advantage, the 4.3 percent due to differences in coefficients on similar characteristics—or what has been

operationalized as discrimination in the literature—amounted to about 25 percent of the overall white sentence advantage. In other words, when we considered the overall effects, ignoring for a moment the potential offsetting effects of the coefficients on legally relevant and extralegal factors, the bulk of the relatively small racial gap in sentencing that accrued in the favor of whites was due to racial differences in offense and offender characteristics as opposed to differences in the coefficients of the separate equations by race.

By way of comparison with the results of the unconstrained model in table III.7, the overall white advantage was larger in the guidelines score models, but that arises from the much larger difference in intercepts. In other words, a much larger portion of the white advantage is unexplained in the unconstrained model (about 27 percent in the unconstrained model as compared to 3 percent in the guidelines scores models). When the guidelines scores were introduced, the residual difference diminished relative to the unconstrained model.

In addition, a sentencing advantage accrued to blacks when we looked only at the effects of the regression. That advantage was outweighed by the differences in intercepts. Finally, the overall amount of the white advantage due to differences in coefficients was about the same size as in the guidelines scores models; however, it amounted to a much larger percentage of the raw differential. This implied that, relative to the sentencing advantage accruing to whites, the guidelines scores models reduced by about one-half the relative amount of the racial disparity due to differences in sentences given to similarly situated offenders.

Returning to table III.8 (the guidelines scores models), we noted that there were offsetting effects. For example, blacks were treated less harshly than whites on extralegal factors (age and months free since previous incarcerations) even though blacks were younger and more likely to have spent less time out of prison. Blacks similarly obtained a minor advantage due to the effects of which judicial circuit they were sentenced in. On the legally relevant factors, blacks obtained an advantage on criminal history. These effects were offset by the larger differences in coefficients on offense severity scores. On this variable, the raw differential amounted to a 30.4-percent disadvantage for blacks; the amount attributable to differences in coefficients is 26 percent, which was the largest amount attributable to differences in coefficients on any variable.

The results on racial differences in coefficients on offense severity were consistent with the results we obtained in the previous section on differences in means. There, we found patterns in which race and offense severity interacted to produce differences in sentences throughout the range of offense severity in specific criminal history categories. In the regression results reported here, we also found a race by offense severity interaction, but here we found that blacks were more likely to be given longer sentences than whites throughout the range of offense severity. Clearly, the effects of offense severity warrant further investigation. It is a variable that is central to the operation of the guidelines, yet it is the variable with the largest differences in coefficients.

In sum, the guidelines can be seen as responsible for reducing overall the relative amount of the racial disparity in sentences. Nevertheless, the guidelines offense severity scores are responsible for the largest amount of the differences in coefficients. In other words, one of the factors that was intended to reduce unwarranted disparity is itself correlated with racial differences in the sentencing of similarly situated offenders as modelled by our regressions. We reiterate, however, that the overall amount of the difference due to differences in coefficients was small and that to the extent that the racial differences in sentencing arose from differences within sentencing ranges, it is not clear whether the disparities were unwarranted.

Table III.9: Details of Parameter Estimates and Decomposition: Models Without Guidelines Scores

	Ordinary least squares estimation of dependent variable = natural log of sentence imposed (excludes guideline scores)						
Dependent variable	White parameter estimates	Black parameter estimates	White means	Black means	Raw differential	Portion attributed to differences in characteristics	Portion attributed to differences in coefficients
Intercept	3.460	3.729	1.000	1.000	–0.269		
Major offense—robbery	0.211	0.062	0.949	0.911	0.143	0.008	0.135
Dollar value stolen, 1 specific offense conduct[a] point	–0.008	–0.132	0.336	0.337	0.042	0.000	0.042
Dollar value stolen, 2 specific offense conduct[a] points	0.082	0.029	0.182	0.186	0.010	0.000	0.010
Weapon use	0.283	0.324	0.423	0.415	–0.015	0.002	–0.017

(continued)

Dependent variable	Ordinary least squares estimation of dependent variable = natural log of sentence imposed (excludes guideline scores)				Raw differential	Portion attributed to differences in characteristics	Portion attributed to differences in coefficients
	White parameter estimates	Black parameter estimates	White means	Black means			
Involvement/ culpability	0.127	0.002	0.131	0.116	0.016	0.002	0.014
Multiple acts	0.086	0.157	0.294	0.300	–0.022	0.000	–0.021
Ongoing behavior	0.209	0.213	0.141	0.101	0.008	0.008	0.000
Criminal justice status at arrest	0.087	0.101	0.514	0.521	–0.008	–0.001	–0.008
Number of counts of conviction	0.085	0.062	1.978	2.010	0.045	–0.003	0.047
Convicted by guilty plea	–0.101	–0.042	0.764	0.674	–0.049	–0.009	–0.040
Consecutive sentences imposed	0.321	0.378	0.182	0.219	–0.024	–0.012	–0.012
Upward departure	0.133	0.269	0.043	0.031	–0.003	0.002	–0.004
Number of adult convictions	0.027	0.038	3.627	3.225	–0.024	0.011	–0.035
Number of prior robbery convictions	0.118	0.064	0.577	0.725	0.022	–0.017	0.039
Prior revocations of supervision	0.100	0.079	0.432	0.411	0.011	0.002	0.009
Number of incarcerations >=5 years	0.113	0.115	0.523	0.517	0.000	0.001	–0.001
Number of incarcerations 1-5 years	0.053	0.035	0.870	0.952	0.012	–0.004	0.017
Months free since previous incarceration	–0.002	–0.003	51.993	48.924	0.036	–0.007	0.043
Gender	0.181	–0.220	0.047	0.035	–0.001	–0.002	0.001
Statutory minimum	0.137	0.226	0.200	0.262	–0.032	–0.009	–0.023
Circuit 2	–0.051	–0.256	0.024	0.097	0.024	0.004	0.020
Circuit 7	–0.074	–0.287	0.042	0.066	0.016	0.002	0.014
Circuit 9	–0.132	–0.093	0.416	0.209	–0.035	–0.027	–0.008

(continued)

Dependent variable	Ordinary least squares estimation of dependent variable = natural log of sentence imposed (excludes guideline scores) White parameter estimates	Black parameter estimates	White means	Black means	Raw differential	Portion attributed to differences in characteristics	Portion attributed to differences in coefficients
Mean dependent variable natural log of sentence length imposed			4.213	4.312	–0.098	–0.051	0.221
F-value	64.229	29.607					
Probability>F	0.0001	0.0001					
R-square	0.623	0.597					
Adjusted R-square	0.633	0.577					
N	880	484					

[a]For details on variable definition, see footnote 9 in the text.

Table III.10: Details of Parameter Estimates and Decomposition: Models Guidelines Scores

Independent variable	Ordinary least squares estimation of dependent variable = natural log of sentence length imposed (excludes guidelines scores) White parameter estimates	Black parameter estimates	White means	Black means	Raw differential	Portion attributed to differences in characteristics	Portion attributed to differences in coefficients
Intercept	0.442	0.470	1.000	1.000	–0.028		
Weapon use	0.148	0.133	0.448	0.452	0.006	–0.001	0.007
Victim injury	–0.195	0.110	0.029	0.047	–0.011	0.003	–0.014
Number of counts of conviction	0.049	0.053	2.036	2.040	–0.007	0.000	–0.007
Convicted by guilty plea	–0.093	–0.008	0.769	0.668	–0.066	–0.009	–0.057
Consecutive sentences imposed	0.243	0.242	0.190	0.252	–0.015	–0.015	0.000
Upward departure	0.283	0.382	0.051	0.040	–0.001	0.003	–0.004
Downward departure	–0.290	–0.323	0.073	0.047	–0.006	–0.008	0.002
Number of prior robbery convictions	0.027	0.030	0.593	0.781	–0.007	–0.005	–0.002

(continued)

Independent variable	Ordinary least squares estimation of dependent variable = natural log of sentence length imposed (excludes guidelines scores)						
	White parameter estimates	Black parameter estimates	White means	Black means	Raw differential	Portion attributed to differences in characteristics	Portion attributed to differences in coefficients
Months free since previous incarceration	-0.001	-0.002	53.660	47.120	0.064	-0.006	0.070
Age	0.005	-0.0001	34.027	30.628	0.162	0.016	0.146
Statutory minimum statute	0.344	0.378	0.175	0.249	-0.034	-0.026	-0.008
Circuit 1	0.168	-0.080	0.029	0.037	0.008	-0.001	0.009
Circuit 2	-0.098	-0.199	0.019	0.103	0.019	0.008	0.010
Circuit 9	-0.062	-0.032	0.368	0.136	-0.018	-0.014	-0.004
Final offense severity score	0.210	0.234	22.426	23.106	-0.680	-0.143	-0.537
Final offense severity score squared	-0.003	-0.003	528.942	564.595	0.376	0.100	0.277
Criminal history III	0.197	0.021	0.161	0.163	0.028	0.000	0.029
Criminal history IV	0.310	0.173	0.104	0.136	0.009	-0.010	0.019
Criminal history V	0.536	0.266	0.070	0.050	0.024	0.011	0.013
Criminal history VI	0.661	0.528	0.238	0.282	0.008	-0.030	0.037
Mean dependent variable natural log of sentence			4.238	4.408	-0.170	-0.127	-0.015
F-value	184.741	77.907					
Probability>F	0.0001	0.0001					
R-square	0.867	0.848					
Adjusted R-square	0.862	0.837					
N	589	301					

Debate on Guidelines' Problems, Benefits, and Effects Continues

From their very inception the sentencing guidelines have been the focus of widespread debate within the courts and criminal justice system. The supporters of the guidelines we interviewed believed that the new system made sentencing more consistent and predictable. Critics we talked to, on the other hand, argued that the guidelines were too harsh and rigid; they did not allow for consideration of relevant personal characteristics of offenders, such as age or family situation, and gave too much power to prosecutors because charging and plea bargaining decisions they make can determine sentencing outcomes.

Objectives, Scope, and Methodology

As shown in table IV.1, we interviewed 53 court and criminal justice personnel in 4 federal court districts (Texas, Western; California, Northern; Maryland; and Wisconsin, Western). We asked them what benefits and problems they saw and what long-range effects they anticipated as a result of the implementation of the guidelines. We also asked them for their perceptions about how the guidelines had affected unwarranted disparity in sentencing, and we asked them for examples of instances in which they believed unwarranted sentencing disparity had occurred under the guidelines.

Table IV.1: Structured Interviews Done by Interviewee Group and Court District

	Number of interviews by court district				
Position	**TX, West**	**CA, North**	**MD**	**WI, West**	**Total**
Chief district judges	1	1	1	1	**4**
District judges	2	2	2	0	**6**
U.S. attorneys	1	1	0	1	**3**
Assistant U.S. attorneys	2	2	3	2	**9**
Federal defenders	1	1	1	[a]	**3**
Assistant federal defenders	1	1	1	[a]	**3**
Private defense attorneys[b]	3	3	3	4	**13**
Chief probation officers	1	1	1	1	**4**
Probation officers	2	2	2	2	**8**
All interviewees	**14**	**14**	**14**	**11**	**53**

[a]This district does not have a federal defender organization. Private defenders are appointed by the courts for all indigent defendants.

[b]Includes defenders appointed by the courts for indigent defendants and privately retained defense attorneys.

Before visiting the four court districts, we pretested our structured interview instruments with officials of the Administrative Office of the

Courts and personnel in the court districts of the District of Columbia and Southern Ohio. We made clarifications and refinements on the basis of comments we received.

Selection of Interviewees

We interviewed managers, supervisors, and line personnel. Where possible, we randomly selected line personnel from lists of those in the office who had done their jobs both before and after the guidelines were implemented and consequently could offer comparisons and contrasts between the two sentencing systems.

However, in several instances, we were not able to use randomly selected interviewees. For example, in the California Northern District we did not randomly select district court judges because the Chief Judge directed that we speak only with judges who volunteered to meet with us. We did not interview any line prosecutors in the Maryland District because the U.S. Attorney directed that we speak only with supervisors. Because we wanted to interview private defense attorneys who defended cases in federal court on a regular basis, we asked court clerks, public defenders, and/or chief probation officers for recommendations in each of the districts and then randomly selected attorneys from this list.

Selection of Court Districts

In selecting the four court districts to visit from the universe of 94 federal court districts, we first eliminated the 12 districts visited by the Commission for its study, the 8 districts we visited in our separate study of the use of mandatory minimum sentences, and the 2 districts we visited for our pretests. We did not consider these districts because we wanted to avoid overly burdening them with interview and documentation requests.

From our remaining universe of 72 districts, we sought 4 districts that when combined had the following characteristics:

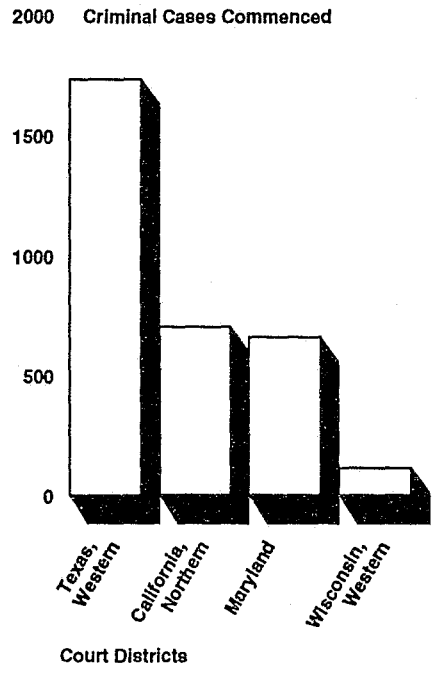
- geographic dispersion, that is, contained a representative from each of the 4 major regions of the nation (North/Northeast, West, Midwest, and South/Southwest);
- mix of urban, suburban, and rural locations;
- mix of districts that implemented the guidelines when they became effective on November 1, 1987, and districts in which full implementation was delayed as a result of local challenges to the guidelines' constitutionality;
- diversity in size of criminal caseload; and

- different patterns of change in criminal caseload over time (caseloads increasing, remaining steady, and decreasing since the guidelines were implemented).

The four major regions of the country, along with a mix of urban, suburban, and rural locations, are represented in the four districts we visited. The large Texas, Western District began using the guidelines in November 1987, while the other three districts did not fully implement them until the Supreme Court upheld their constitutionality in January 1989.

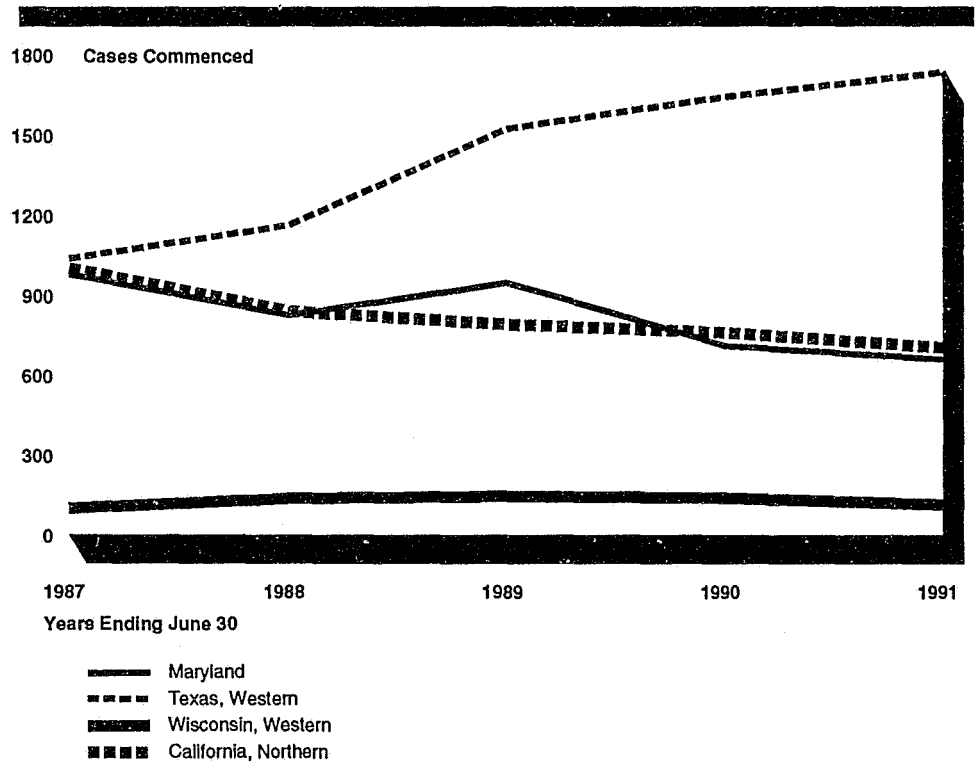
The size of the criminal caseloads varied among the four districts. During the year ending in June 1991, the average number of criminal cases commenced per district (including cases filed, reopened, and transferred from other districts) was about 500. Only 5 of the 94 district courts commenced 1,700 or more criminal cases. In the Texas, Western District, 1,732 criminal cases were commenced. The California, Northern District commenced 703 criminal cases; the Maryland District, 659; and the Wisconsin, Western District, 113 criminal cases. (See fig. IV.1.)

Figure IV.1: Criminal Cases Commenced in the Four Court Districts We Visited, Year Ending June 30, 1991



The districts also experienced different patterns of change in the number of criminal cases they commenced for the years ending June 30, 1987, through June 30, 1991. As shown in figure IV.2, criminal cases commenced increased by about 67 percent over this period in the Texas, Western District, from 1,036 cases in 1987 to 1,732 cases in 1991. Criminal cases commenced were fairly steady over the period in the Wisconsin, Western District, with an increase of about 15 percent, from 98 cases commenced in 1987 to 113 cases commenced in 1991. The number of criminal cases commenced dropped about 33 percent in Maryland, from 977 in 1987 to 659 in 1991, and the number dropped about 30 percent in the California, Northern District, from 1,000 cases in 1987 to 703 cases in 1991. We did not determine the reasons for the changes in the caseload sizes.

Figure IV.2: Criminal Cases Commenced, Years Ending June 30, 1987, Through June 30, 1991



The structured interviews provided information on various officials' views on the guidelines system, but they did not constitute a representative or statistically valid sample of opinions.

Other Studies of the Guidelines

Over the period we did our work, other studies and conferences on the guidelines also took place. We reviewed the results of the Commission's district interviews in its report. Among the other literature we reviewed were a study by Senior Eighth Circuit Judge Gerald W. Heaney[1] and preliminary results of a study by Sentencing Commissioner Ilene Nagel and University of Chicago School of Law Professor Stephen Schulhofer. Commissioner Nagel discussed this study in March 1992 at the Sentencing Institute for the Second and Eighth Judicial Circuits.[2]

[1]Judge Gerald W. Heaney, "The Reality of Guidelines Sentencing: No End to Disparity," American Criminal Law Review, Vol. 28 (1991).

[2]Ilene H. Nagle and Stephen J. Schulhofer, "A Tale of Three Cities: An Empirical Study of Charging and Bargaining Practices Under the Federal Sentencing Guidelines," Southern California Law Review (forthcoming—Nov. 1992).

Problems, Benefits, and Effects Interviewees Saw as a Result of the Guidelines

The prosecutors we interviewed generally believed that the sentencing guidelines had improved the system for sentencing federal defendants, while the majority of judges, defense attorneys, and probation officers cited more problems than benefits and/or said that they did not believe the guidelines were an improvement over the prior sentencing system.

We asked 3 open-ended questions of 53 district judges, prosecutors, defense attorneys, and probation officers on the benefits, problems, and long-range effects of the guidelines. We also recorded and analyzed general statements about the guidelines system made during the course of the interviews. The open-ended questions we asked were as follows:

- Compared to the preguidelines system, do you see any benefits in the current guidelines system?
- Do you see any problems in the current guidelines system as compared to the preguidelines system?
- Do you see any long-range effects of the guidelines?

As shown in table 3 in the letter, the most frequently mentioned benefits were less disparity or more uniformity in sentences under the guidelines and more certainty of what sentences offenders would actually serve. As shown in table 1, the most frequently mentioned problems were that the guidelines were too harsh in some cases (specifically for drug and first-time offenders and minor participants in conspiracies); too inflexible; "dehumanizing" because they reduced multifaceted human behaviors to a set of numbers; and, by limiting judges' discretion, gave prosecutors too much control over sentencing based on how they charged offenders and what pleas they accepted. Some interviewees also thought that the guidelines were too complex and difficult to use, problems some said were exacerbated by too many amendments.

Most long-term effects of the guidelines anticipated by interviewees focused on impacts on prisons, including costs; increases in populations; discipline; and readjustment of offenders back to society after long terms. (See table 2.)

Summaries of statements made by interviewees who thought the guidelines had resulted in overall improvement of the system for sentencing offenders follow:

- I like the guidelines. They put a certainty in sentencing. Also, you get consistency in sentencing across the country, which is very important. Prior to the guidelines this was not true. (A prosecutor.)
- I favor the guidelines system because I do not think discretion was used appropriately in the old system. Judges and prosecutors had the ability to be creative in fashioning appropriate punishments, but they were not responsible in using the discretion. (A prosecutor.)
- I think the guidelines are great from the perspective of the probation officer and the citizen. Guidelines are an attempt to quantify justice and do away with disparity. They provide criteria for sentencing. The defendant has a clear picture of what is going to happen to him. Everything is more out in the open. (A probation officer.)

Summaries of statements made by interviewees who did not think the guidelines had improved the system for sentencing offenders follow:

- I am very dismayed. Before the guidelines I would give my best shot at an appropriate sentence and feel that justice was done. Now I never feel that justice is done. There are subtle distinctions that make no sense, especially in terms of amounts of drugs. You could do just as well using a person's weight to make a sentencing decision. For example, a defendant I sentenced dropped a gram of cocaine on a kilo of sugar. I am supposed to sentence on the weight of the sugar. In another case, a drug courier was sent from Baltimore to Miami to get two kilos of cocaine. He came back with one, but I am still supposed to sentence him on two. What if he had been instructed to get as much cocaine as he could? (A judge.)
- The lack of discretion by the court is an enormous problem. The quality of justice has significantly decreased. The inability to give individualized sentences is a problem. Relevant factors are ignored. Prosecutors have too much control over sentences. The system is better if this discretion rests with the court, which is independent, experienced and on the record, rather than with young assistant U.S. attorneys who have a vested interest in their cases and conduct their business behind closed doors. (A probation officer.)
- The fundamental problem with the guidelines is that they are misguided in their concept. The whole idea of finding a sentence in this book (the Commission's Guidelines Manual) by going to page 33 and reading down the column is misguided. It is an artificial, abstract concept that does not have a lot to do with the human behavior it is trying to modify. I also find it repugnant that the U.S. Attorney is essentially in control of sentencing now. (A private defense attorney).

Majority of Interviewees Thought Judges Should Have More Grounds to Depart From the Guidelines

The Sentencing Reform Act of 1984 directed that federal judges have the right to depart from the guidelines when they find special circumstances in cases that the Commission did not adequately consider. However, the act also purposefully limited the use of judicial departures to those special circumstances. The Senate Judiciary Committee and the full Senate rejected amendments that would have allowed judges to depart from the guidelines whenever they thought the circumstances of the case warranted it, whether or not the Commission had considered those circumstances in the development of the guidelines.

The act also directed the Commission to be sure that the guidelines were entirely neutral on the effect race, gender, national origin, creed, and socioeconomic status of offenders have on sentencing. The Commission was also directed to take care that the guidelines reflected the general inappropriateness of considering education and vocational skills, employment record, family ties and responsibilities, and community ties of the defendant.

To specifically address concerns that the guidelines sentencing system does not allow flexibility to take special circumstances of individual cases into account, we asked 41 district judges, prosecutors, and defense attorneys the following:

- Are there factors either not included in the sentencing guidelines or included as not ordinarily relevant that you believe should be grounds for lawful departure from the guidelines?

As shown in table IV.2, the great majority of the interviewees thought that there were factors generally excluded from consideration by the guidelines that judges should be able to take into account in sentencing offenders.

Table IV.2: Interviewee Responses on Whether Factors Either Not Included in the Sentencing Guidelines or Included as Not Ordinarily Relevant Should Be Grounds for Lawful Departure From the Guidelines

	Yes	No	Did not answer	Total responses
District judges	6	3	1[a]	10
Prosecutors	9	2	1	12
Federal defense attorneys	6	0	0	6
Private defense attorneys[b]	13	0	0	13
All interviewees	**34**	**5**	**2**	**41**

[a]This judge did not answer because he said he did not want to give the guidelines system any credence by making suggestions for improvements.

[b]Includes defenders appointed by the courts for indigent defendants and privately retained defenders.

Some of the factors mentioned by judges and defense attorneys as ones that should be considered in sentencing were factors that the Sentencing Reform Act had specifically found inappropriate in all or most cases. These factors included socioeconomic status and family responsibilities. Other personal factors interviewees considered important were physical and mental health and potential for rehabilitation. Prosecutors generally mentioned more specific, narrowly defined circumstances in which they thought the guidelines should allow judicial departures.

The following were summaries of comments given by judges and defense attorneys who, contrary to the underlying premise of the Sentencing Reform Act, thought that personal and socioeconomic factors should be considered in sentencing:

- All of those factors they were so careful to exclude from consideration in the first place should be considered . . . . These are things that you would want to know about a human being before you send him away. You are talking about all of the factors that have to do with a defendant's personal background, reasons for involvement in the offense, future plans, hope for rehabilitation. The list of things about adefendant are potentially endless, but you are talking about the personal factors. (A defense attorney.)
- The range of personal or socioeconomic status factors—like age, family circumstances, and education—as well as potential for rehabilitation, should be considered. (A judge.)
- The whole life of the defendant should be considered, not just the snapshot of the offense committed and anything else the defendant has ever done bad. The white-collar defendants I represent have many good qualities such as good work ethics, charitable contributions, and

responsibility to meeting the needs of their families. The health of the defendant and any hardships that would explain the criminal behavior should also be taken into account. (A defense attorney.)

Summaries of some of the prosecutors' statements were as follows:

- The guidelines should allow downward departures for defendants who are terminally ill.
- Defendants' health, family responsibilities, and attempts at rehabilitation should be taken into account, but education and income should not be considered. Those offenders with education and opportunity are more accountable for their crimes than those who did not have the same opportunities.

Views Mixed on Whether Motions for Downward Departures Based on Substantial Assistance Should Be Made Only by Prosecutors

A guidelines policy statement directs that "upon motion of the government" [prosecutor], the court can depart from the guidelines because a defendant "has provided substantial assistance in the investigation or prosecution of another person who has committed an offense." We asked the 41 district judges, prosecutors, and defense attorneys the following:

- Should the prosecutor continue to be the only person permitted to make motions for downward departures based on substantial assistance, or should the defense attorney and/or the court also be permitted to make such motions?

As shown in table IV.3, the views of the interviewees on this subject were mixed. All of the prosecutors thought this authority should remain exclusively theirs, while all of the defense attorneys thought that defense attorneys and/or the courts should also have authority to make the motions. The 10 judges were split in their views with 4 judges thinking the authority should stay with the prosecutor and 6 thinking that they and/or defense attorneys should also be permitted to make the motions.

Table IV.3: Interviewee Responses on Who Should Be Permitted to Make Motions for Downward Departures Based on Substantial Assistance

	Prosecutor only	Defense attorney and/or the court	Total responses
District judges	4	6	**10**
Prosecutors	12	0	**12**
Federal defense attorneys	0	6	**6**
Private defense attorneys[a]	0	13	**13**
All interviewees	**16**	**25**	**41**

[a]Includes defenders appointed by the courts for indigent defendants and privately retained defenders.

Summaries of comments from interviewees who thought prosecutors should continue to have the sole authority for making motions for downward departures for substantial assistance included the following:

- The prosecutors are the only ones who can judge if the defendant really did provide substantial assistance. (A prosecutor.)
- There are potential abuses to the prosecutor's authority, but the U.S. Attorneys Offices have enough checks and reviews to assure that departures are given fairly and when warranted. All motions for substantial assistance in this district are reviewed by a supervisory assistant U.S. Attorney and the U.S. Attorney to assure that the system is not abused. (A prosecutor.)
- If the defense were allowed to make the motions, they would do so in every case with frivolous claims of substantial assistance. It would slow the process down and involve morehearings to decide how good the cooperation was. The court has no business enforcing the law or encouraging cooperation. (A judge.)

Summaries of comments from interviewees who thought defense attorneys and/or judges should also have the authority to make motions for downward departures based on cooperation included the following:

- The fox is watching the henhouse. This discretion needs to go back to the court and not to the charging authority where there is no external review. (A defense attorney.)
- I cannot tell you the number of times when I feel the defendant has cooperated and provided assistance, but we don't get it. It comes down to a trust factor between the prosecutor and the defense attorneys. What would be substantial assistance to one attorney may be a drop in the bucket to someone else. (A defense attorney.)

- The present system is a serious weakness. It gives too much power and leverage in negotiations to the prosecutor. It opens up the possibility for abuse. (A judge.)

Other Studies of the Guidelines' Impact Found Similar Benefits, Problems, and Effects

The themes expressed in our interviews were also documented in other studies and/or a conference on the impact of the guidelines during the same period.

The Commission asked some similar and some different questions from the ones we asked in interviews it conducted in 11 judicial circuits for its study of guidelines implementation. From its interviews with 245 judges, prosecutors, defense attorneys, and probation officers, the Commission did not report on whether the guidelines had improved the system for sentencing offenders. It asked interviewees whether the guidelines had been effective in meeting the congressionally established purposes of sentencing. The Commission found that defense attorneys were generally negative in their assessments, but a sizeable majority of judges, probation officers, and prosecutors interviewed gave favorable assessments.

Interviewees in the Commission and GAO studies cited some common benefits and problems in the guidelines system. Common benefits cited included decreased disparity and increased predictability or certainty. Common problems cited were reduced judicial discretion, inflexibility, and harshness.

The Commission policy decision to give the prosecutor the sole authority to determine when a motion for a downward departure for substantial assistance to the government can be made was the subject of discussion at the March 1992 Sentencing Institute for the Second and Eighth Judicial Circuits, court districts within the Northeast and Midwest. Debate centered on whether the policy statement was merely advisory or binding on judges.

Senior Eighth Circuit Judge Gerald W. Heaney, following interviews of court personnel and a review of more than 800 presentence reports for 1989 in 4 districts within the Eighth Circuit, found that plea bargaining practices vary from district to district. He also found that no uniform Department of Justice standards to guide U.S. Attorneys in making substantial assistance motions had been articulated. Thus, he concluded that policies may vary from district to district and practices may vary from defendant to defendant with no mechanism for review. Judge Heaney also

found that the length of time an offender can expect to serve in prison had increased substantially under the guidelines. He recommended, among other things, that district court judges be authorized to give a sentence shorter than that mandated by the guidelines without a motion from the prosecutor on a finding that an offender has given substantial assistance to the government.[3]

Sentencing Commissioner Ilene Nagel and Professor Stephen Schulhofer, University of Chicago School of Law, interviewed officials and reviewed a total of approximately 700 case files in 8 court districts with the objective of determining whether prosecutors' plea agreements circumvented the guidelines. In a paper on their draft results delivered at the Sentencing Institute for the Second and Eighth Circuits, Commissioner Nagel noted evidence that plea agreements circumvented the guidelines in about 20 to 35 percent of the cases they examined. She said that the issue of harshness in sentencing, particularly when mandatory minimum sentences could apply, was a factor when the guidelines were circumvented. Among the draft suggestions by Commissioner Nagel and Professor Schulhofer to limit circumvention of the guidelines through plea negotiation was one that the Commission look at the balance the guidelines establish between flexibility and structure to see if the line has been drawn at the appropriate place. They also recommended that the Commission look at several ways to reduce the harshness of sentences for some categories of offenders, including nonviolent, nondrug, first-time offenders and low-level offenders involved in drug conspiracies.

Most Interviewees Thought Unwarranted Sentencing Disparity Did Not Occur Frequently

One of Congress' main goals in authorizing the sentencing guidelines was to reduce unwarranted disparity in sentencing of similar offenders who had committed similar crimes. A number of studies and reports, including a GAO testimony before the House Judiciary Subcommittee on Criminal Justice in 1978, had documented differences among federal court districts in the treatment of similarly situated offenders.[4]

We asked district judges, prosecutors, public and private defense attorneys, and probation officers in the four districts their perceptions about how the guidelines had affected unwarranted disparity in

[3]At a November 13, 1991, symposium on the guidelines at the Georgetown University Law Center, Commission Chairman, Judge William W. Wilkins, Jr., took exception to both Judge Heaney's methodology and conclusions. We summarize Judge Heaney's findings here as an indication that differences in presentencing practices vary; and such variations can affect the sentencing imposed.

[4]Disparities in Criminal Sentencing and Prospective Practices in Federal District Courts (Apr. 24, 1978).

sentencing, and we asked them for examples of instances in which they believed unwarranted sentencing disparity had occurred under the guidelines.

As shown in table IV.4, most interviewees thought that unwarranted sentencing disparity in their districts had decreased or stayed the same under the guidelines. Of the 48 interviewees responding to this question, 21 said that unwarranted sentencing disparity had decreased, and 18 said it had stayed the same. This result was consistent with the frequent mention of less disparity or more uniformity in sentences in response to the open-ended question on benefits to the guidelines system. No prosecutors thought there was more unwarranted sentencing disparity under the guidelines than before, while responses among other groups were mixed.

Table IV.4: Interviewee Responses on Changes in Sentencing Disparity in Their Districts Under the Guidelines

	Judges	Prosecutors	Defense attorneys[a]	Probation officers
Decreased	2	10	5	4
Stayed the same	4	2	6	6
Increased	2	0	5	2
Did not answer	2	0	3	0
All interviewees	**10**	**12**	**19**	**12**

[a]Includes federal defense attorneys and privately retained and appointed defenders.

As shown in table IV.5, most interviewees did not view unwarranted disparity in sentencing as a frequently occurring problem. Only one interviewee said unwarranted disparity occurred in most or all cases. However, 36 of the 48 interviewees who responded to the question believed that unwarranted disparity occurred in some or a few cases.

Table IV.5: Interviewee Responses on How Frequently Unwarranted Sentencing Disparity Occurred in Their Districts Under the Guidelines

	Judges	Prosecutors	Defense attorneys[a]	Probation officers
In no cases	1	3	0	1
In a few cases	2	6	4	4
In some cases	1	3	11	5
In many cases	3	0	1	2
In most or all cases	0	0	1	0
Did not answer	3	0	2	0
All interviewees	**10**	**12**	**19**	**12**

(Table notes on next page)

[a]Includes federal and privately retained and appointed defenders.

Instances in Which Interviewees Thought Unwarranted Sentencing Disparity Occurred Under the Guidelines

When asked for examples of unwarranted sentencing disparity under the guidelines, interviewees in all four practitioner groups cited circumstances in which they thought unwarranted disparity occurred. We note that some interviewees considered as unwarranted disparities situations in which offenders having similar criminal histories and offense conduct were sentenced differently. Some also considered that unwarranted disparity existed when offenders with different levels of culpability were sentenced similarly.

Examples of types of unwarranted disparity interviewees thought occurred under the guidelines and summaries of statements they made in illustration of each type were as follows:

(1) The guidelines require similar sentences for defendants with different levels of culpability in drug cases.

- The biggest reason for disparity under the guidelines is that a person who is in a position to assist the government gets a downward departure—and a lesser sentence—than defendants who do not have much information or who are not the first to talk. A lot of sentences depend on the order in which the conspirators were arrested and prosecuted. (A judge.)
- A drug runner who flew large amounts of drugs in from Florida and a female addict who lent him money to buy the plane he used for drug running received the same guidelines sentences of 17-1/2 years for very different levels of culpability. (A judge.)
- The leader of a drug ring should not get the same sentence a mule gets, but I have two examples of how this occurred. In one case, a young mother allowed a drug ring to warehouse drugs in her apartment for $100 a month to supplement her welfare check. She was charged with the entire conspiracy and sentenced to 15 years in prison. In another case, a cocaine dealer with a prior record cooperated in his case and received a reduction for substantial assistance and a sentence of 45 months. A codefendant picked up money for the dealer and did not even know he was involved in a cocaine ring. He was sentenced to 21 months, and it would have been longer if the judge had not given him a downward departure. (A probation officer.)

Concerns about offenders with different levels of culpability receiving similar sentences were based on the Commission's policy decision to consider offenders' actual criminal conduct in sentencing as opposed to conduct for which the offender is charged and convicted. An offender is to be held accountable at sentencing for the conduct of the other participants in a jointly undertaken criminal activity if the defendant could reasonably foresee the behavior of his/her codefendants.

(2) Prosecutors exercise discretion in negotiating charges and plea agreements.

- The discretion that was once with probation officers and the court is now with the prosecutors. They wheel and deal and decide what to charge, and they control the cooperation issue. (A probation officer.)
- Prosecutors are as individualized as judges. Some are hard-nosed and will not make deals, but others will. (A judge.)
- In one case, a prosecutor manipulated the weight of drugs in a plea agreement for a defendant who cooperated. The probation officer never knew the amount of drugs involved. (A defense attorney.)

(3) Judges exercise discretion in use of departures and award of acceptance of responsibility.

Though the majority of interviewees thought that judicial influence over sentencing was reduced under the guidelines, some pointed out that different judges, court districts, and circuits still sometimes sentenced similarly situated offenders differently because they applied the guidelines differently.

- The two judges here have marked differences in attitudes on when to allow downward departures for substantial assistance that show up in marked differences in the ultimate sentences. (A defense attorney.)
- In our district, except in extraordinary circumstances, when we make a motion for departure based on substantial assistance, judges will give two level reductions if the assistance is within the defendant's own case and up to four levels if the assistance is provided on other cases, as well. In other districts, once the motion is made, judges may give up to a seven or eight level reduction. (A prosecutor.)
- One judge will depart under a given set of circumstances and another judge will not depart under the same set of circumstances because he wants to do the safe thing. (A probation officer.)

(4) The guidelines include irrelevant factors that affect the length of some drug offenders' sentences.

- On one LSD distribution case, the weight of the paper containing the LSD increased a defendant's sentence. Another defendant distributing equal amounts of LSD not contained on paper or contained on a lighter weight of paper would get a less severe sentence. Inclusion of the weight of the paper as part of the drug amount creates disparity. (A probation officer.)
- One defendant gets caught selling an ounce of methamphetamine and another defendant gets caught with a gallon of an acetone wash waste product. The second defendant will get a higher sentence than the first because of the gallon drum. (A prosecutor.)

Some interviewees also said that the following factors result in different sentences for similar defendants:

- errors in applying the guidelines,
- differences in the skills and knowledge of personnel (i.e., defense attorneys, probation officers, and investigative agents) assigned to cases, and
- whether a case is prosecuted in the federal system or in state or local courts.

We note that these last three circumstances would be likely to exist in any federal sentencing system. However, interviewees commented that the complexities of the guidelines system increased the possibilities for application errors and exacerbated differences in skills and knowledge that people have.

When asked specifically whether private attorneys appointed for indigent defendants in their districts had a "generally adequate" knowledge of the guidelines, 21 of the 48 judges, prosecutors, defense attorneys, and probation officers responding said no. Among the 27 interviewees who said these attorneys had "generally adequate" knowledge were those who qualified their responses with such statements as the following:

- Adequate, but not good. (A prosecutor.)
- Barely adequate. The knowledge does not compare with the expertise of the public defense attorneys who use them daily. (A judge.)

- Generally adequate is the appropriate term. Some are better than others. I've seen some defense attorneys who were "at sea" using the guidelines. (A probation officer.)

Guidelines' Impact on Staff, Workloads, and Case Processing Times of the Courts and Court-Related and Investigative Agencies

In September 1987, just before the sentencing guidelines went into effect, we reported on the potential impact of the guidelines on the criminal justice system as required by the Sentencing Reform Act.[1] We noted then that it seemed widely accepted from the interviews we conducted that the guidelines would increase workloads for virtually all criminal justice components. We said that the full impact of the guidelines would become clear only when there was empirical evidence available on how they were implemented.

Our review indicated that the guidelines probably increased the workload of appellate and district judges, prosecutors, federal and community defenders, and probation officers. According to interviewees, workloads for some pretrial service officers, investigative agents, United States Marshals Service (USMS) personnel, and district court clerks and their staffs also increased as a result of the guidelines. The magistrate judges we interviewed said their workloads were minimally affected by the guidelines because most of the cases they decided were not felonies to which the guidelines applied. With the exception of the workload analysis done by the Probation Division of the Administrative Office, the court components and agencies had not done workload studies or other empirical studies to quantify the impact of the guidelines on their operations.

The Commission was not required by the Sentencing Reform Act to estimate the potential impact of the guidelines on the workload of the federal courts and court-related agencies. Its report on the implementation and impact of the guidelines did not address this issue.

Objectives, Scope, and Methodology

To determine whether the significant increases in workload that our earlier work predicted came about in actual guidelines experiences, we did headquarters and district interviews. (See app. IV for a description of how we selected the court districts to visit and the personnel to interview in the districts.) In our visits to the 4 court districts, we asked 56 appellate and district judges, prosecutors, federal and private defenders, and

probation officers who apply the guidelines directly in their work how their workload had changed since the guidelines were implemented. We asked them whether their overall workloads had changed and whether specific components of their job took more or less time as a result of

[1]Sentencing Guidelines: Potential Impact on the Federal Criminal Justice System (GAO/GGD-87-111, Sept. 10, 1987).

implementation of the guidelines. We also asked whether other factors had increased their workload and how the guidelines ranked in significance compared to these other factors.

We asked similar questions in joint meetings in each district of managers, supervisors, and line personnel (a total of 22 interviews) at agencies as well as court components less directly involved in implementing the guidelines: the Federal Bureau of Investigation (FBI), the Drug Enforcement Administration (DEA), USMS, clerks of court, magistrate judges, and pretrial service officers. We selected officials who had both pre- and postguidelines experience.

We interviewed Administrative Office officials and reviewed budget, staffing, and workload data for criminal justice system components, including probation offices, U.S. Attorneys' offices, and federal and community defenders' offices. We analyzed Administrative Office statistics for the years ending June 30, 1986, through June 30, 1990, to identify trends in case filings, trial rates, and length of time between conviction and sentencing for federal criminal defendants. We reviewed some Administrative Office staffing data and the results of a workload study completed by its Probation Division.

Interviewees Said Some Aspects of Their Work Took Longer Under the Guidelines

Interviewees said that some aspects of their jobs were new under the guidelines and others took longer to do under the guidelines. Working through the criminal justice process from investigation of crimes to sentencing of offenders, some interviewees identified aspects of their jobs in every part of the process that were more time consuming under the guidelines. These aspects included the time to investigate cases, negotiate plea agreements, resolve disputes, participate in sentencing hearings, house and transport offenders, and process cases and sentencing appeals through the courts.

Investigation of Cases

Some FBI and DEA agents said their investigations had to be documented more carefully under the guidelines than before. In particular, some agents said they had to spend more time documenting drug quantities and criminal histories of suspects, because these factors are important determinants of sentences under the guidelines.

Pretrial Services

Two of the districts we visited had a pretrial services office separate from the probation offices. Probation officers handled pretrial interviews and services for offenders in the other two districts. In one of the districts the pretrial services officials we interviewed said that the guidelines had no impact on their workload. In the other district, officials said under the guidelines they spent more time waiting to meet with their clients because defense attorneys delay their access. Defense attorneys wanted to meet with their clients first to prevent them from giving the pretrial service officers information that might later be used against them in guidelines sentencing calculations. They also said they made more court appearances under the guidelines.

Negotiation and Review of Plea Agreements

Under the guidelines, the great majority of criminal cases continued to be resolved by plea agreement. According to 1990 Administrative Office statistics, more than 70 percent of all federal criminal cases were closed as a result of a plea agreement.

As shown in table V.1, the majority of prosecutors and defenders we interviewed said that plea negotiations were more time-consuming for them under the guidelines, and half of the district judges interviewed said that reviewing plea agreements was more time-consuming for them.

Table V.1: Impact of the Guidelines on Time Required to Negotiate/Review Plea Agreements

Official	More time-consuming	Less time-consuming	Equally time-consuming	Total
District judges	5	0	5	10
Prosecutors	10	1	1	12
Defense attorneys	14	3	2	19

Source: GAO interview results.

Administrative Office guidance on preparing presentence investigation reports under the guidelines instructs probation officers to do independent assessments of the impact any plea agreements would have on guidelines sentences. Probation officers we interviewed said that they did not do this function before the guidelines. Of the 12 officers we interviewed, 4 said that assessing the impact of plea agreements took a great deal more time under the guidelines, 5 said it took moderately more time, 2 officers said that under local policies in their district they do not generally assess the impact of plea agreements, and 1 said it was no more time consuming.

Some investigative agents also said they spent more time on plea negotiation under the guidelines. They said that meetings with prosecutors to review evidence and specific facts to include in plea agreements took longer under the guidelines.

Summaries of statements by interviewees who thought plea negotiation and review took longer under the guidelines are as follows:

- Under the old system, plea agreements contained recommendations for sentencing. Now plea agreements determine sentences. I must see if the agreements fit the guidelines, and I explain the guidelines to the defendants at the time of the proceeding. This takes about 10 more minutes in court when taking a guilty plea. (A judge.)
- Now we have to negotiate factors that the court would have dealt with in the past . . . factors like whether the government can prove drug weights and when it determines the offense behavior began. (A defense attorney.)
- There are different factors to take into account now like base offense level, role in the offense, and acceptance of responsibility. Before the guidelines we simply made our charges and sentencing recommendations. Now we must clear our pleas internally to see that they are within the policies of the Justice Department and the U.S. Attorneys Office. (A prosecutor.)
- We did not assess the impact of plea agreements before the guidelines, so it is a little more time-consuming now, but it is not something we spend a lot of time on. (A probation officer.)

Resolution of Disputes Before Sentencing

The process of informally resolving disputes among prosecutors, defenders, and probation officers over the contents of presentence investigation reports was also new under the guidelines. As one probation officer explained, in the preguidelines era, both counsels were to submit their own versions of the offense to the court, probation officers confidentially were to submit presentence reports with sentencing recommendations to judges, and judges were to announce their findings at formal sentencing hearings. Under the guidelines, probation officers are to submit drafts of presentence investigation reports to prosecutors and defenders, and the parties are to attempt to resolve any disputes before the sentencing hearing. Disputes that cannot be resolved are to be decided by district judges during the formal sentencing hearings when sentences are imposed.

As would be expected since dispute resolution was new under the guidelines, the majority of interviewees involved in the process said it was more time consuming for them than before the guidelines were implemented. (See table V.2.)

Table V.2: Impact of the Guidelines on Time Required to Resolve Disputes

Official	More time-consuming	Less time-consuming	Equally time-consuming	Total
District judges[a]	6	0	4	10
Prosecutors	11	0	1	12
Defense attorneys	17	2	0	19
Probation officers	12	0	0	12

[a]Some district judges said that they did not become involved in dispute resolution at this stage while others said that they reviewed objections to presentence investigation reports and probation officers' responses as part of their preparation for sentencing.

Source: GAO interview results.

Some investigative agents also commented that they received more calls from prosecutors and probation officers at this stage to clearly establish the facts of cases.

Summaries of examples officials gave of why they thought dispute resolution was more time-consuming under the guidelines than before follow:

- You find yourself arguing with the prosecutor or the probation officer or both about whether you are supposed to add 2 points or 4 points or deduct or whatever. It used to be that you looked in terms of the big picture. Now you find yourself arguing over small points, like whether a piece of stolen equipment was worth $9,500 or $10,001, because those are the differences that can affect the length of time a client will serve. (A defense attorney.)
- There are just more facts to be resolved than existed before the guidelines—for example, drug amounts, cooperation, role in the offense, and acceptance of responsibility. (A prosecutor.)
- Resolving disputes is one of the most time-consuming parts of the presentence investigation. Attorneys may make frivolous objections hoping that at least one will be accepted by the judge. The probation officer has to deal with each objection in writing, review his notes and the presentence investigation report, possibly contact the case agent as well as the prosecutor again, call the defense attorney and try to iron things

out, and research case law. All of that is very time-consuming. (A probation officer.)

Other Presentence Investigation Work

The probation officers we interviewed said that other aspects of their jobs related to researching and writing presentence investigations, assessing the impact of plea agreements, and resolving disputes were more time-consuming under the guidelines. The entire approach to completing presentence investigations has changed because of the guidelines. According to guidance provided by the Administrative Office, before the guidelines, probation officers were to concentrate on assessing defendants' potential for rehabilitation and understanding the circumstances that caused the defendant to commit the crime. Under the guidelines, probation officers instead are to address specific facts in cases that relate to guidelines applications.

Table V.3 lists other aspects of presentence investigations that some probation officers thought were more time-consuming for them to do under the guidelines.

Table V.3: Impact of the Guidelines on Time Required to Complete Other Presentence Investigation Work

Task	More time-consuming	Less time-consuming	Equally time-consuming
Determine criminal histories	11	0	1
Write presentence reports	10	1	1
Investigate facts	9	0	3
Complete supervisory reviews	9	0	3
Make sentencing recommendations	5	3	4

Source: GAO interview results.

Holding Sentencing Hearings

As shown in table V.4, most of the participants in sentencing hearings we interviewed (district judges, prosecutors, and defense attorneys) also said that sentencing hearings took longer under the sentencing guidelines.

Table V.4: Impact of Guidelines on Time Required to Participate in Sentencing Hearings

Official	More time-consuming	Less time-consuming	Equally time-consuming	Total
District judges	9	0	1	**10**
Prosecutors	11	1	0	**12**
Defense attorneys	14	3	2	**19**

Source: GAO interview results.

Officials said that hearings took longer because more detailed facts related specifically to guideline applications were addressed, particularly if disputes were not resolved informally by the sentencing hearing. Summaries of their comments follow:

- I now need to announce detailed findings on the record. More details are needed for appellate review. (A district judge.)
- Each guidelines factor can significantly increase or decrease a sentence, so there is now considerable factual and legal argument on each individual factor. It is like a scorecard. Lack of agreement between the parties on certain issues can greatly increase the length of the hearing. (A prosecutor.)

Housing and Transportation of Prisoners Detained During Court Proceedings

Most USMS personnel we interviewed in the four court districts said that guidelines implementation had caused an increase in the number of prisoners they housed and transported. Some interviewees said that because defendants were exposed to potentially longer sentences under the guidelines, they were less likely to be released on bond, and thus more likely to be in USMS custody. As available jail space filled, they noted, they were housing prisoners farther away from courts, increasing the time they spent transporting them to and from court appearances. They said this is not a new problem, but one exascerbated by the guidelines. Some interviewees also said that there were more evidentiary hearings under the guidelines that required them to transport prisoners to court.

In his 1990 annual report, the director of USMS predicted that the sentencing guidelines, along with implementation of the Comprehensive Crime Control Act of 1984, the Anti-Drug Abuse Act of 1988, and other recent changes in the criminal justice system, would continue to affect USMS workload. One measure of USMS workload is the average daily number of prisoners in USMS custody which nearly doubled from 7,328 in fiscal year 1986 to 13,390 in fiscal year 1990.

Processing Cases

In each district we visited, the clerks of court and their staffs said that the guidelines had increased the time it took to process cases through the system. For example, one interviewee said that the form they prepared at sentencing, the Judgment and Conviction Order, was more complicated and time-consuming to complete under the guidelines. In addition, some interviewees said that the workload of court reporters increased because hearings were longer.

Sentencing Appeals

The Sentencing Reform Act of 1984 expanded the authority of the U.S. courts of appeals to review sentences. Both defendants and the government can appeal guidelines sentences.

As shown in table V.5, the total number of criminal appeals filed increased from 6,012 in the year ending June 30, 1988, the first year for which statistics of guidelines appeals were available, to 9,949 in the year ending June 30, 1991. The percentage of appeals that included a sentencing issue increased from 3 percent in 1988 to 64.9 percent in 1991.

Table V.5: Criminal Appeals Filed, Years Ending June 30, 1988, Through June 30, 1991

Year	Total appeals of criminal cases	Total appeals of guidelines cases	Appeals with a sentencing issue	Percent of appeals with a sentencing issue
1988	6012	225	179	3.0
1989	8020	4412	3251	40.5
1990	9493	7319	5386	56.7
1991	9949	8259	6460	64.9

Source: GAO analysis of Administrative Office data.

Despite the increasing numbers of sentencing appeals filed, 26 of the 35 appellate judges, district judges, prosecutors, and defenders who had experience with cases involving sentencing appeals said they did not find the sentencing appeals system too time-consuming, and many said they handled them the same way they handled other appeals.

Several of the nine interviewees who thought that the sentencing appeals system was too time-consuming stressed that it was the volume of sentencing appeals that took time for them to handle, not the complexity of the appeals. Some also noted, however, that the sentencing appeal rate overall would decline as time passed and case law regarding sentencing issues became clearer.

Personnel in district clerks of court offices said that their workload was affected by expansion of authority to appeal sentences. Interviewees in all four districts said clerks' staff spent more time under the guidelines handling administrative aspects of preparing files to be forwarded to the appellate courts, including arranging for district court proceedings to be transcribed for appellate court review.

Supervised Release Revocation Hearings

Supervised release is the new form of postimprisonment supervision created by the Sentencing Reform Act that replaced parole. Under the guidelines, district judges gained new responsibility for considering whether offenders violated conditions of their supervised release and determining how the offenders should be penalized. The Commission issued nonbinding policy statements to assist the judges in their deliberations. Unlike parole revocation cases, prosecutors are to represent the government in supervised release revocation cases. Defense attorneys are to represent offenders in both types of proceedings.

Most district judges, prosecutors, defense attorneys, and probation officers said they had little or no experience handling supervised release revocation cases. One defense attorney said that few offenders sentenced under the guidelines had been released from prison to begin terms of supervised release. However, some interviewees expressed concern about the future impact of these proceedings on their time. For example, one district judge described them as "sleeping monsters" that would have a heavy impact on judges' time. A probation officer predicted the impact would be felt beginning in 1993, when the first offenders with more serious guidelines convictions started coming out of prison, and he did not think judges would be happy to hear from probation officers on revocation cases.

Staffs of two of the district clerks of court also said they had new duties related to supervised release revocation proceedings. As described by one chief deputy clerk, the clerks' offices are to receive petitions from probation officers indicating violations of supervised release, they are to issue bench warrants for violators' arrests on court order, and after revocation hearings are held, they are to prepare new judgment and commitment orders.

Data Did Not Show Increases in Number or Length of Trials Under the Guidelines

Although some interviewees attributed increases in their workloads under the guidelines to the fact that guidelines cases were more likely to go to trial, national trend data show that plea and trial rates remained stable.

As shown in table V.6, trial and plea rates for defendants of all crime types across all federal districts remained stable from the year ending June 30, 1986, the year before the guidelines were implemented, through June 30, 1990. Trial and plea rates in the districts we visited also remained generally stable over this same period. For example, the change in trial rates during this period ranged from a 2.2-percent increase in Western Texas to a 1.1-percent decrease in Maryland. All districts we visited had trial rates under 12 percent in the year ending June 30, 1990.

Table V.6: National Case Disposition Rates for All Defendants, Years Ending June 30, 1986, Through June 30, 1990

	Rate (in percent)					
Disposition	**1986**	**1987**	**1988**	**1989**	**1990**	**Change**
Dismissed	15.8	15.9	15.9	15.4	14.5	–1.3
Plea	70.8	71.3	71.1	70.8	71.6	+0.7
Trial	13.4	12.8	13.1	13.8	13.9	+0.5
Total	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**0.0**

Note: Totals may not add to 100 because of rounding.

Source: GAO analysis of Administrative Office data.

As shown in table V.7, trial lengths for all crime types across all federal districts have also remained stable under the guidelines. More than 70 percent of all trials continued to take 3 days or less in the year ending June 30, 1990. This information was not readily available for the same period in the four districts we visited.

Table V.7: Percentage of Trials Completed by Days for All Federal Defendants, Years Ending June 30, 1986, Through June 30, 1990

	Percent of trials completed					
Length of trial	**1986**	**1987**	**1988**	**1989**	**1990**	**Change**
1-3	73.7	73.7	75.7	76.6	76.6	+2.9
4-9	21.2	20.9	19.8	18.5	19.0	–2.2
10-19	3.8	3.7	3.2	3.5	3.2	–0.6
20+	1.4	1.7	1.3	1.4	1.2	–0.2
Total	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**0.0**

Note: Totals may not add to 100 because of rounding.

Source: GAO analysis of Administrative Office data.

Data Showed Link Between Guidelines' Implementation and Increases in Case Processing Time Frames

Consistent with interview results showing that interviewees thought they spent more time on various aspects of their jobs under the guidelines, Administrative Office case processing data showed that the median time for cases to move through parts of the criminal justice system increased after implementation of the guidelines. The median time from case filing to disposition (by plea agreement, conviction or acquittal at trial, or dismissal) increased after the guidelines were implemented, and the median time from conviction to sentencing increased as well. The data did not directly attribute the increase to the effects of the guidelines as opposed to other system interventions.

As shown in table V.8, the median number of months from case filing to disposition for all defendants nationally increased 1.3 months from 3.2 months in the year ending June 30, 1986, to 4.5 months in the year ending June 30, 1990. The increase was similar to the national data for the four districts.

Table V.8: Median Number of Months From Case Filing to Disposition for All Crime Types, Years Ending June 30, 1986, Through June 30, 1990

	Number of months					
District	1986	1987	1988	1989	1990	Change
Maryland	3.8	3.9	4.1	3.6	5.0	+1.2
Western Texas	2.8	2.9	3.5	4.0	4.5	+1.7
Western Wisconsin	4.1	3.4	3.7	3.9	4.2	+0.1
Northern California	3.0	3.7	4.0	4.9	5.1	+2.1
National	3.2	3.4	3.6	4.1	4.5	+1.3

Source: GAO analysis of Administrative Office data.

As shown in table V.9, the median number of days from conviction to sentencing increased 28 days after the guidelines were implemented from 41 days in the year ending June 30, 1986, to 69 days in the year ending June 30, 1990. The most marked increase occurred between the year ending June 30, 1988, and the year ending June 30, 1989, when nationwide implementation of the guidelines began. The data is consistent with interviewees stating that it took them longer to handle sentencing issues.

Table V.9: Median Number of Days for Cases to Move From Conviction to Sentencing, Years Ending June 30, 1986, Through June 30, 1990

	Number of days					
District	1986	1987	1988	1989	1990	Change
Maryland	43	48	52	78	81	+38
Western Texas	32	35	37	49	59	+27
Western Wisconsin	52	41	49	56	58	+6
Northern California	39	42	49	57	75	+36
National	41	42	46	61	69	+28

Source: GAO analysis of Administrative Office data.

Interviewees Said the Guidelines Had the Most Significant Effect on Workload of All Recent Interventions

During the years after the implementation of the sentencing guidelines beginning in November 1987, interventions in the criminal justice system, in addition to the guidelines, occurred that could have affected workload. For example, Congress enacted legislation establishing mandatory minimum sentences for some crimes, and, according to agency officials, investigative and prosecutorial priorities in some areas shifted to larger drug conspiracies and more complex white-collar crimes. Recognizing that these changes were taking place during the same time that the sentencing guidelines were being implemented, we attempted to distinguish their effects by asking interviewees to identify factors in addition to the guidelines that had affected their workload and then to rank the effects of these factors with the effects of the guidelines.

Most district judges, prosecutors, federal defenders, and probation officers—those interviewees most directly involved in guidelines implementation—ranked sentencing guidelines as having the most significant impact of all interventions on increases in their workload. Clerks of court and USMS personnel in the districts also generally ranked the guidelines as the most significant intervention. FBI and DEA agents said that the guidelines had increased their workloads, but they generally ranked their effects second or third behind other interventions that they thought had more significant effects. Some guidelines impacts were also reported by a pretrial services chief, and minimal impacts were cited by magistrate judges. We did not ask this question of the appellate judges and private defense attorneys we interviewed.

Specifically, 31 of the 39 district judges, prosecutors, federal defenders, and probation officers we interviewed said that the guidelines had the most significant impact on their workload of all interventions over the last several years. Legislation establishing mandatory minimum sentences and

shifts in prosecutorial priorities were ranked as having had the next greatest impacts. These results generally held true across practitioner type as well.

Table V.10 shows the distribution of frequency of responses when judges, prosecutors, probation officers, and federal defenders were asked to rank the effects of the guidelines with other changes.

Table V.10: Rank of Effects of Guidelines and Other Interventions on Workload

	Rank				
Factors	**1st**	**2nd**	**3rd**	**4th**	**Total**
Sentencing guidelines	31	7	1	0	**39**
Mandatory minimums	1	13	8	2	**24**
Prosecutorial priorities	6	10	6	2	**24**
Antidrug abuse acts	0	1	6	4	**11**
Other[a]	1	6	3	3	**13**

[a]Responses in this category included growth and increased complexity of criminal law, enhanced supervision of offenders, and bail reform.

District clerks of court staff in two of the four districts ranked the guidelines as the number one factor increasing the time to do their jobs. Clerks of court staff in the other two districts ranked the guidelines as the third and fifth factors following shifts in internal administrative procedures and mandatory minimum sentences.

USMS personnel in two of the four districts also ranked the guidelines as the number one factor increasing the time to do their jobs. In the other two districts, USMS personnel ranked the guidelines second, after implementation of other provisions of the antidrug abuse acts and shifts in prosecutorial priorities.

Special agents in one of the eight FBI and DEA offices we visited said that guidelines' implementation was the most significant factor increasing the time it took them to investigate cases. DEA and FBI agents in three other offices ranked them second or third, behind the impact of shifts in agency and/or prosecutorial priorities and mandatory minimum sentences on their time.

Responses were split in the pretrial service offices we visited. One officer ranked the implementation of the guidelines as the most significant factor

increasing the time it takes to do his job. The other officer did not think that the guidelines had an effect on his time.

The magistrate judges we interviewed in all four districts commented that they were not substantially affected by the sentencing guidelines. However, three of the magistrates did identify some guidelines' effects on their work and ranked the guidelines as first or second in significance of all factors increasing the time it to took to do their jobs.

Two Components Based Staffing and Budget Increases on the Guidelines' Impact on Workload

Two court components based requests for staff increases at least in part on the guidelines' impact on workload. The Probation and Defender Services Divisions of the Administrative Office found that the guidelines increased the time it took probation officers and federal and community defenders to do their work; and, because each staff member could handle fewer cases under the guidelines, they justified requesting new positions.

Probation Division

The Probation Division was the only component of the Administrative Office that attempted to quantify the impact of the guidelines on its workload. The Administrative Office reported the impact to Congress in budget submissions, beginning with a request for supplemental funds in fiscal year 1988, and Congress consequently authorized 596 positions through fiscal year 1991.

According to a Probation Division official, a 1981 work measurement study served as a baseline for measuring guidelines impact on workload. This study found that one full-time probation officer could complete 97 presentence (including investigations, hearings, and report preparation) and postsentence reports in a year. Officials adjusted this baseline as a result of a small work measurement study done in three districts in 1986, which found that a probation officer could complete 79 presentence and postsentence reports in a year. Just before the November 1987 implementation of the guidelines, a panel of probation office managers, using their experiences in informally implementing the guidelines the Commission proposed to Congress in May 1987, estimated the guidelines' impact on the number of presentence and postsentence reports a probation officer could complete. They determined that under the guidelines the number of presentence and postsentence reports a probation officer could do would drop to 55 in a year, a decrease of 24 investigations a year compared to the number that could be done before the guidelines. The panel did not find any guidelines' impact on other

probation officer tasks such as supervising offenders on probation, parole, or supervised release.

In January 1989, the Probation Division began a major workload study based on surveys of more than half of its probation and pretrial services personnel and visits to five districts. The study, as approved in September 1991, had results close to the 1987 estimates. It determined that a probation officer assigned to presentence investigations could complete 57 a year.

On the basis of the estimates, Division officials said Congress authorized them a total of 596 additional positions (368 probation officers and 228 clerks) for fiscal year 1991. The 368 probation officer positions represented 13.9 percent of the total 2,645 probation officer positions requested for probation services that year. In fiscal year 1992, the Administrative Office requested 27 more probation positions to implement the guidelines. It assumed that beginning in fiscal year 1991, all new presentence investigations cases would be handled under the guidelines and that no future staffing requests would be based on guidelines impacts.

Defender Services

The 1993 Administrative Office budget request to Congress estimated that the guidelines increased the time it took federal and community defenders to defend cases by 25 to 50 percent, and the Administrative Office requested and received additional positions based, in part, on the guidelines' impact. However, the Defender Services Chief said that because federal defenders do not perform discrete tasks like probation officers do, their work is not amenable to quantitative formulas, and the Division made no attempt to do workload studies to document the impact. He said that a combination of the impact of the sentencing guidelines and increased drug arrests and prosecutions accounted for almost all of the increases in Defender Services staff since fiscal year 1989 and that the guidelines were responsible for more than half of the increase. From fiscal year 1989 through fiscal year 1991, Congress authorized 461 new positions (including attorneys, investigators, paralegals, and administrative support staff). According to the Chief's estimates, more than 230 of these would have been necessary as a result of workload increases to implement the guidelines.

Conclusion

Interview results and the limited statistical data available indicated that the predictions made by interviewees in our 1987 review of the potential

guidelines' impact were correct and that the guidelines have probably increased workloads of most components of the federal criminal justice system. However, the empirical data we found lacking in our 1987 report were still not available to support this conclusion. Probation Services was the only court component to have attempted to measure the impact of the guidelines on workload. Reliable workload measures that would allow precise measurements of the impact the guidelines have had did not exist for the period before the guidelines were implemented.

Comments From the United States Sentencing Commission

UNITED STATES SENTENCING COMMISSION
1331 PENNSYLVANIA AVENUE, NW
SUITE 1400
WASHINGTON, DC 20004
(202) 626-8500
FAX (202) 662-7631

July 30, 1992

Mr. Richard L. Fogel
Assistant Comptroller General
Government Accounting Office
441 G Street, N.W., Suite 3858-C
Washington, D.C. 20548

Dear Mr. Fogel:

On behalf of the Sentencing Commission, thank you for the opportunity to review your report on the sentencing guidelines prior to its submission to Congress. We are gratified to see that through the use of a variety of different techniques the GAO arrives at the same basic findings and conclusions as the Commission: disparity has decreased under the sentencing guidelines.

We would like to make four general observations on your report and recommend revisions to clarify what we believe are errors or characterizations that could lead to serious misinterpretation.

I. Issues Related to Disparity Analyses

In what appears to be a misreading of a Commission research study (see Chapter Four, Section VI/A, "Judicial Sentencing Patterns Under the Guidelines," USSC Evaluation Report, 299-339), the GAO elaborates on analyses performed by the Commission on its fiscal 1990 monitoring data. We believe the GAO report seriously mischaracterizes the Commission's study and recommend the following clarifications:

The Commission study focuses on the within guideline-range variation of sentences by a series of sociodemographic variables. Clearly characterized as an initial, exploratory look at bivariate relationships (utilizing only a Chi-square test) (see USSC Evaluation Report, Volume II, 299), the study explicitly cautions against any claim of a causal relationship between variables (*e.g.*, gender and sentence, race and sentence, employment and sentence). At a minimum, the same caveat must be added to the

GAO's preliminary findings; *i.e.*, they are not conclusive as to causal claims.

But, perhaps more importantly, characterizing these analyses as "Unwarranted Disparity Continues to Exist Under the Guidelines" suggests a finding that clearly misstates the Commission's research as well as that of the GAO. The preliminary findings reported by the Commission do not suggest disparity under the guidelines; rather, they report bivariate relationships between variables in the exercise of judicial discretion within the appropriate guideline range. As you know, Congress provided judicial latitude in sentencing within a range of 25 percent or six months, whichever is greater, or departure for the unusual case (28 U.S.C. § 994(b)(2)). The guidelines provide individual sentencing ranges based on congressionally directed factors of offense severity and criminal history. The actual sentence is selected by the court from within the appropriate range, or, if departing, above or below it.

To elaborate, Congress did not define judicial sentencing variation within this 25-percent range as disparity. Rather, Congress intended that sentencing variation outside of the appropriate guideline range for similar offenders convicted of similar offenses would represent unwarranted disparity. The Commission's first study in the disparity section of its report examined the question of disparity outside the appropriate guideline range, and found that disparity as measured by dispersion was indeed less under the guidelines. The second and third studies in our report were preliminary observations of judicial sentencing patterns within the congressionally authorized guideline range, variations Congress chose not to define as disparity.

I emphasize that had the Commission or the GAO found a causal relationship between these variables (and we emphasize that neither study did), it would not have been a relationship that showed disparity in application of the guidelines, but rather in the manner in which judicial discretion is exercised within the statutorily provided sentencing range. Such findings, if they had occurred, might suggest a need for Congress to consider further limiting judicial discretion in order to narrow the possible introduction of unwarranted disparity within the appropriate guideline range or through departure.

Now on p. 3. Also see discussion on pp. 25-26.

A few specific examples might prove helpful. On page 4, the GAO reports that "the Commission evaluated whether unwarranted disparity -- that is, disparity resulting from offender characteristics that should normally be irrelevant to determining a sentence -- continued under the guidelines." The Commission clearly reported that it had not completed its analysis of unwarranted disparity at the time of the report's submission to Congress and that that work was continuing (see USSC Evaluation Report, 299). To characterize the descriptive profiles presented in the Commission's report as "evaluated whether unwarranted disparity... continued under the guidelines" misrepresents the Commission's work.

Now on p. 3.

Further, the GAO report on page 6 suggests that the Commission's preliminary findings point to continued disparity under the guidelines based on gender, race, employment, marital status, and age. The Commission's findings do not lead to that conclusion. Rather, the Commission reports descriptive profiles for selected offense types solely within guideline ranges and for departures (see USSC Evaluation Report, 300-339) and cautions against drawing conclusions from these preliminary analyses.

For these reasons, we believe this section should be revised and the title amended to more accurately reflect both the Commission and the GAO findings with respect to the pertinent studies.

II. Characterizations of Data Limitations

There appear to be several points of confusion in the GAO report regarding data sources in both the Commission and GAO studies. The GAO report suggests that the data limitations result from some shortcoming on the part of the Commission's efforts. This really is not the case; the data limitations are structural limitations that could not be alleviated by a more expanded data collection effort. In general, the Commission believes that while available data have limitations (as is true in virtually all research endeavors), its study and analyses are valid and provide empirically meaningful answers to the questions posed.

The GAO report implies that meaningful comparisons pre- and post-guidelines could have been made had the pre-guideline data contained additional cases. Such a characterization begs the question and fails to recognize that pre/post comparisons would be problematic if there had been as many as 50,000 pre-guideline cases. The fact is that pre-guideline and guideline cases are not directly comparable because they result from two dramatically different sentencing structures involving different decision-making processes and data that mean very different things pre- and post-guidelines. This is not a data collection issue that can be remedied with more extensive information; rather, it is a comparison problem stemming from the altered nature of the new sentencing system that restricts any conclusive comparison. (For discussion, see USSC Evaluation Report, Volume II, 273-274.) In sum, comparing pre- and post-guideline sentences is akin to comparing apples and oranges. Even if the numbers double, the comparison remains of apples and oranges.

The limited number of guideline cases available (reflecting a time span considerably shorter than the four years of operation intended by Congress) resulted from: a) the congressional directive that guidelines would apply only to offenses committed after November 1, 1987, resulting in a gradual phase-in of cases; and b) the numerous constitutional challenges to the Sentencing Reform Act and the guidelines, delaying implementation on a nationwide basis until the Supreme Court decision in

January 1989.

A specific problem with sample sizes is highlighted by the GAO analysis (see Appendix II of the GAO report). In order to arrive at larger sample (and cell) sizes, this analysis compromises the substantive question of measuring sentence variation for similar cases. Collapsing categories increases the dissimilarity of offenses and offenders, removing the possibility of definitive conclusions on unwarranted sentencing disparities.

Scientific research rarely uses population data. Rather, most research is conducted based on samples. Moreover, as you are well aware, many of the important findings and advances of science are drawn from inferences based on relatively small samples. In that sense, the Commission's study provides sufficient empirical testing for the questions it sought to address, most salient of which (highlighted in the GAO report as well) is the question of unwarranted disparity. The Commission recognizes the small sample sizes in its disparity analyses; however, it should be noted that the Commission project, while beginning with a sufficiently large number of cases, arrives at small numbers in an attempt to produce a valid pool of "similar offenders with similar offenses" in order to examine any unwarranted dissimilarities in their sentences. Taking its lead from the definition of disparity Congress provided, the opportunity to create large samples did not exist.

III. Historical Context and Legislative Intent

While some of the statements advanced by the GAO are content-correct, out of context they may be misleading to the reader. A discussion of several historical-legislative issues would enhance the contextual interpretation of the study. For example, through the enactment of the Sentencing Reform Act, Congress intended a reduction of judicial discretion in sentencing, and as such, it could have been expected that some number of judges would not be pleased with Congress's scheme of sentencing reform. It, therefore, is not surprising that GAO found that some federal judges were critical of the guideline system. Setting the context within which the GAO findings exist would add to the understanding of such findings.

In a similar vein, discussions related to workload (including increases in the number of appeals) should not be surprising given the break with past sentencing practices brought about by the Sentencing Reform Act. Congress undoubtedly recognized that a drastic change in the sentencing system would increase workload and case processing time. It was a price the legislature apparently was willing to pay for a more just, consistent, and fair system of sentencing. Further, the increased workload and time are reflective of the justified attention paid to defendants in a system that is both more objectively based on factual findings pertinent to the case and results in more severe penalties. Finally, the number of appeals has increased in response to Congress'

statutory authorization of appeals of sentences by the defendant or government. Discussion and recognition of these congressionally intended changes would provide a context for a clearer understanding of the operation of the guidelines.

IV. Title, Subtitle, and Subject Headings

If one simply glanced at the GAO report, noting the title, subtitle, and various subject headings without carefully reading associated text, the reader would come away with a very negative impression of what on closer examination proves to be an objective presentation of study results.

In general, the title, subtitle, and subject headings of the report, instead of adopting either the traditional topic headings or characterization of section content approaches, appear to represent eye-catching labels that often do not capture the essence of the section under discussion. In doing so, some headings unintentionally misrepresent the findings or cast them in a negative light, creating a more critical impression of the issue.

Now on p. 5.

For example, on page 9 of the GAO report, the heading reads "Wide Sentencing Discretion Generated Concern." While discussion related to wide sentencing ranges appears in this section, the content is much broader. The section also discusses rising crime rates and questions about the rehabilitative model during the pre-guideline period (although upon first reading, it is not clear whether the heading refers to the pre-guideline or guideline era), and an introduction of congressional intent with respect to the content of the guidelines. Someone glancing at headings would miss the important substance contained in this section. Perhaps a broader label would work better; *e.g.*, "Factors Leading to Guideline Development."

Now on p. 12.

As we argue above, the Commission believes that findings related to disparity are mischaracterized in the GAO report. A simple reading of the heading on page 21 of the GAO report furthers that mischaracterization. The Commission study reports variations in sentencing patterns within the applicable guideline range or for departures; labeling this as "Under the Guidelines" is misleading and "Unwarranted" is incorrect. Congress identified the width of the sentencing range as acceptable and therefore not unwarranted. The substance of the section might be better characterized as "Sentencing Variation Within Guideline Categories."

While we are pleased that the GAO replicated the Commission's findings of reduced sentencing disparity under the guidelines, we are disappointed that the title of the GAO report does not reflect this very important finding. Because few reform programs yield early indications of success, we would hope for acknowledgement of this significant finding. This easily could be accomplished by adding a few words to the title

of your report, *e.g.*, "Sentencing Guidelines: Disparity Reduced, But Some Questions Remain."

In conclusion, we appreciate the opportunity to comment on the GAO's draft report and are pleased that the majority of your findings support the independent findings of the Commission. While there may be any number of additional points we could make in response to the GAO report, it seems most productive at this time to limit our comments to the major concerns. We agree with your conclusions that considerable work remains in order to ensure the complete implementation of sentencing guidelines as intended by Congress.

With highest personal regards, I am

Sincerely,

William W. Wilkins, Jr.
Chairman

Comments From Judge Vincent L. Broderick, Chairman, Committee on Criminal Law of the Judicial Conference of the United States

COMMITTEE ON CRIMINAL LAW
of the
JUDICIAL CONFERENCE OF THE UNITED STATES
United States Courthouse
101 East Post Road
White Plains, New York 10601-5086

Maryanne Trump Barry
Pasco M. Bowman, II
James C. Cacheris
James G. Carr
Stanley S. Harris
George P. Kazen
Charles P. Kocoras
Richard P. Matsch
David A. Nelson
Lenore C. Nesbitt
Pamela Ann Rymer
Mark L. Wolf

Vincent L. Broderick
Chairman

[FTS] (914) 682-6139

FACSIMILE

[FTS] (914) 682-6140

July 29, 1992

Richard L. Fogel
Assistant Comptroller General
United States General Accounting Office
Washington, D.C. 20548

Dear Mr. Fogel:

Thank you for according me the opportunity to comment on the draft report, Sentencing Guidelines: Central Questions Remain Unanswered.

I am the Chair of the Judicial Conference Committee on Criminal Law. In the relatively short period available for comments, it has not been possible to develop a definitive position of the Committee on Criminal Law with respect to the report, and the comments which follow are my own. They have, however, been informed by various positions which the Judicial Conference has taken with respect to Sentencing Guideline matters, and I have received considerable assistance in formulating them from Barbara Meierhofer Vincent of the Federal Judicial Center.

I have read the draft report from the perspective of one who, throughout the life of the Sentencing Guidelines, has been involved in negotiations with the Sentencing Commission: in the formulation phases of the Sentencing Guidelines; in the development of procedures for the application of the Guidelines within the Judiciary; and in transmitting the Judicial Conference's recommendations for modification. My personal view is that the Sentencing Guidelines have become a permanent part of the criminal justice landscape, and that it behooves all of us -- Sentencing Commissioners, judges, practitioners, and members of C ngress -- to work together to render them workable.

There are obstacles to this endeavor. Before addressing the substance of the draft report, it is necessary to delineate at least some of these obstacles.

Consider first the conceptual approach which Congress took in the Sentencing Reform Act. It did not anticipate that the Sentencing Commission would, in its first endeavor, achieve the millennium. Congress expected, in fact, that in the first instance the Sentencing Commissioners would rely on historical experience in developing the Guidelines. And this is exactly what the Sentencing Commission did: it comprehensively studied what judges did in the past, and with one notable exception -- white-collar crime -- the initial Guidelines reflected an average based on past sentences for particular crimes. In the white-collar crime area the Sentencing Commission made a value judgment that historical sentences did not adequately reflect the seriousness of the criminal activity involved, and the Commission therefore deliberately set the sentencing ranges for white-collar crimes at higher than the historical levels. Congress envisioned a gradual honing process: the principal engine of Sentencing Guidelines modification was to be departures by sentencing judges, up or down, when those judges found the existence of aggravating or mitigating circumstances "not adequately taken into consideration by the Sentencing Commission." The Commission was then to consider the departures of judges in making appropriate adjustments to the Guidelines. Under the statute, in short, departures were to be the very lifeblood of the Sentencing Guidelines system: through consideration of these departures the Sentencing Commission would be in a position constantly to improve the Guidelines.

It has not worked that way, to this point, for various reasons. One reason is that for a time the Commission seemed institutionally to regard departures as an affront and as an invitation to close loopholes. This is no longer the case: the members of the Commission now, in my judgment, recognize the essential role that reasoned departures must play in the evolution of the Sentencing Guidelines. Another problem was we judges: too often we either did not recognize our duty in appropriate cases to depart, or when we did depart we did not adequately set forth our reasons, thus leaving the Sentencing Commission in the dark as to the departure rationale. Still another problem has been that the various Courts of Appeals, not apprehending the developmental nature of the Sentencing Guidelines, have too often regarded those Guidelines as cast in stone and have unduly restricted the exercise of the duty to depart by district judges. A further problem has been that litigants have been overwhelmed by the complexity of the Guidelines and many of them have been less than adequate advocates in the Sentencing arena.

Another major factor which has hindered the development and full flowering of a rationale Sentencing Guidelines system has been the spate of mandatory minimum sentencing requirements, which have been enacted in the last eight years particularly in the drug and firearms fields. Mandatory minimums are, in their very essence, inconsistent with the entire concept of Sentencing Guidelines, because they skew the entire system. The Sentencing Commission, perhaps, could have structured its Guidelines independent of the mandatory minimums, but for understandable reasons it chose not to.

As a result, a great many of the Sentencing Guidelines provide not the "heartland" which the Sentencing Commission envisioned as a preliminary step -- they provide sentences well above those which the Sentencing Commission, left to its own devices, would have prescribed. Since the Sentencing Commission has attempted to maintain some sort of ordered relationship between the sentence ranges it prescribes for various crimes, the effect has been to draw into higher ranges crimes which on a "heartland" basis would merit much lower ranges.

The judiciary and the Sentencing Commission have had their differences, but on one thing they agree: so long as mandatory minimums continue on the statute books, it will be impossible to develop a Sentencing Guidelines system that is both just and fair.

Other problems beset the road to a truly effective Guidelines Sentencing system. While judges are expected to depart from a guideline range -- in either direction -- whenever they find the existence of "an aggravating or mitigating circumstance of a kind, or to a degree, not adequately taken into consideration by the Commission...that should result in a sentence different from that described," a judge may not depart from a mandatory minimum absent a motion by the prosecutor. Thus the prosecutor in his charging decision determines whether or not a mandatory minimum is to be invoked; by his decision to move or not to move for a departure, he determines whether or not the mandatory minimum standard is to be adhered to. Thus there has developed, in the Guidelines Sentencing era, a massive transfer of sentencing discretion from the judiciary, where it belongs, to the prosecutor's office where it does not. Prosecutors have, of course, historically had almost unchanneled charging discretion. But the stakes have never previously been so high.

The synergistic relationship between mandatory minimum sentences and prosecutorial power with respect thereto has resulted in two phenomena: sentences that are unwarrantedly similar, and sentences that are shockingly disparate. The unwarrantedly similar sentences occur in cases where a mandatory minimum sentence is prescribed for all defendants charged with a given crime, even when those defendants have markedly different involvements, from minimal to positions of control. The shockingly disparate sentences occur when, in the same sort of case, an organizer cooperates against his underlings and receives a below-mandatory minimum sentence while those minimally involved receive mandatory minimums.

It has always been my understanding that a desideratum in sentencing, and a prime purpose of Congress in enacting the Sentencing Reform Act, was the reduction -- indeed the elimination -- of unwarranted disparity. Disparity, in and of itself, is not inherently wrong. Indeed a vice promoted by mandatory minimum sentence requirements is that of unwarranted uniformity: sentencing criminals with differing degrees of culpability to the same sentence. While you have paid lip service in various parts of the draft report to the reduction or elimination of unwarranted disparity as being an objective, I would suggest that the materials contained in your report (and this is, I believe, true with respect

to the report of the Sentencing Commission as well) may tell us something about the reduction of disparity, but very little or nothing about the reduction of unwarranted sentencing disparity.

I realize that the issue of unwarranted sentencing disparity is a very complicated one. It may be one that cannot be effectively analyzed statistically. Statistical assessments of the reduction of disparity will be strongly influenced by the offense characteristics and the offender characteristics which are used to define "similarly situated offenders." A comparison between a pre-sentence guideline and a post-sentence guideline dispersion of sentences for offenders who are grouped by type of offense and number of prior sentences will probably lead to quite a different conclusion with respect to dispersion than would such a comparison with respect to offenders who are grouped on the basis of type of offense, age and culpability. In the report, the emphasis seems to be on utilizing the Sentencing Guidelines offense severity level and criminal history category (and for some breakdowns, type of offense) to classify offenders as "similar." These analyses basically test whether the sentences of offenders scored into particular guideline ranges are more or less likely to cluster around the ranges than they were before those ranges existed -- that is, whether the sentencing is in accordance with the Sentencing Guidelines scheme. This is an important -- almost a baseline -- question, but a relatively narrow one within the broader area of unwarranted sentencing disparity, and it should be recognized as such. There are important questions to be asked in determining whether disparities, if they exist, are unwarranted or not. Are there variations in the investigation and reporting of offense behavior and of offender characteristics that differentially influence the amount or the type of information available to the court which is making guideline determinations, and given the same information, are the Guidelines being calculated reliably? Do the same reported circumstances lead to the same findings of fact? Dispersion of sentences within a guideline cell will tell us little about overall unwarranted disparity if there are differences in the information or in the application principles which are used to place cases within those cells.

Did Congress, and did the Commission, in determining the factors which are relevant in the Guideline scheme, adequately capture "similar offenders" for purposes of determining whether variations in sentencing are warranted? There was a time when sentencing judges considered such factors as age, drug use or the size or nature of the community in which the crime was committed. When we eliminate these and other factors from the total mix have we reduced unwarranted disparity in the sentencing of similar offenders, or have we brought about an increase in unwarranted uniformity in the sentencing of dissimilar offenders? A policy question: it is one that may be informed, but not answered, by empirical analysis. I applaud the suggestion in the report that additional research into the question of disparity is critical. I would suggest that the report should spell out how little is known; should narrow the objective to the elimination not of disparity per se, but of unwarranted disparity and I would also hope that the report would give direction to the

types of questions that should be addressed. Identifying those questions is critical if the data necessary for definitive analysis is to be collected.

Congress and the Sentencing Commission have provided that certain factors are either inappropriate (race and gender) or "not generally appropriate" or "not ordinarily relevant" ("marital status and employment status"); the Sentencing Commission deems age "not ordinarily relevant." The draft report, in characterizing variation with respect to each of these factors as "unwarranted disparity," presupposes the correctness of the determinations by Congress and the Commission. I am concerned that such a presupposition might hinder careful reassessment of the appropriate role of offender characteristics in the sentencing process. I would urge that in order to assess the impact of factors which presently are, by statute or under the Sentencing Guidelines, beyond the pale, data should be collected on those factors. I certainly approve your call for more extensive and exhaustive multi-variate empirical analysis.

The report properly notes the difficulties which would be entailed in undertaking empirical research into the extent and causes of disparity related to prosecutors' decisions. I would urge, however, that such research is critical. There are variations in prosecutorial decision-making, and such prosecutorial decision-making has received a statutory imprimatur which almost certainly differentiates it from pre-Guideline practice. The effect of this statutorily sanctioned decision-making should be explored and critically analyzed.

There is extensive debate concerning the Sentencing Guidelines: and it is reflected in an extensive literature. I suggest that it would be helpful to analyze that literature, to crystallize the issues that are being debated, and to turn them into questions for empirical analysis that could be addressed by future evaluations.

With respect to the effect of the Guidelines on the operation of the criminal justice system I believe that there is no question that the workload has been increased as a part of the impact of the Sentencing Guidelines, but I would suggest that the impact of the Sentencing Guidelines on the criminal justice system can scarcely be assessed in terms of workload alone. The draft report suggests that the Sentencing Commission be instructed to report on this matter.

A large number of entities comprise the criminal justice system and no one of them has the expertise to assess the total impact of the Sentencing Guidelines on the criminal justice system, or even to define adequately the subject areas that should be addressed. Thus operational issues for the Bureau of Prisons would include such matters as the problems entailed in managing a prison population without the incentives of parole and liberal good time, caring for an increasingly geriatric population; and responding to the different requirements of the increasing female population. The issues for the courts would include not only those with direct impact within the criminal justice area, such as expanding

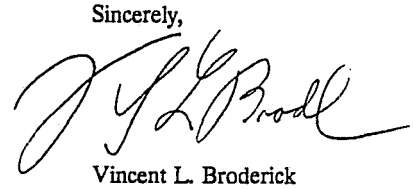
Note: we did not reproduce this appendix, but where appropriate have incorporated suggested technical changes in the text.

pretrial supervision and probation, intensive supervision with respect to persons on supervised release; the problems of coping with probation and supervised release revocation hearings; and the impact of the increased criminal workload on access to justice for civil litigants.

I suggest that the Sentencing Commission does not have any particular expertise in the areas I have delineated, and it would be more helpful if the various agencies more directly involved report on the effects which the Guidelines Sentencing system have on them.

I have attached an appendix which contains particular comments with respect to various parts of the draft report.

Sincerely,



Vincent L. Broderick

/jf

Comments From the U.S. Department of Justice

U.S. Department of Justice

Washington, D.C. 20530

JUL 29 1992

Richard L. Fogel
Assistant Comptroller General
General Government Division
U.S. General Accounting Office
Washington, D.C. 20548

Dear Mr. Fogel:

The following information is being provided in response to your request to the Attorney General, dated June 30, 1992, for comments on the General Accounting Office (GAO) draft report entitled, "Sentencing Guidelines: Central Questions Remain Unanswered." The Department would like to address two significant points. First, we found that the discussion of racial disparities in the body of the report is misleading as to GAO's actual findings. We believe GAO should provide greater explanation of these disparities in the body of the report. Second, we found that the discussion of prosecutorial discretion in Appendix IV does not fully examine the role of such discretion in the investigative and prosecutive process and, therefore, gives the impression that it is exercised in an arbitrary manner. We believe further information is necessary to dispel that impression. We discuss both of these issues below.

SENTENCING DISPARITIES

Possible Disparities Resulting Under the Sentencing Guidelines. The body of the report is misleading with regard to the report's findings on racial disparity and should contain an expanded discussion of this topic. The present discussion states that GAO has confirmed the Sentencing Commission's conclusion that racial disparity exists in sentencing. (See p. 22.) It also indicated that blacks were less likely to be convicted by plea than whites and that persons convicted by plea tended to receive shorter sentences. It suggests that this factor may be responsible for the sentencing gap between blacks and whites. Because of the controversy surrounding this topic and the suggestion in the body of the report that black defendants are punished more severely than whites, readers of the body of the report will likely conclude that GAO has confirmed the existence of this form of racial disparity as a general matter.

Now on pp. 12-13.

Mr. Fogel 2

However, the detailed work set forth in the appendices shows that in some instances where GAO found racial disparity, the difference worked in favor of black defendants. Appendix II states that blacks were more likely to receive sentences at the bottom of the guideline range, rather than in the middle of the range, by a factor of 1.3. Interestingly, blacks were also more likely than whites to receive sentences at the top of the guideline range rather in the middle of the range, by the same factor. (See p.48.) This likelihood remained very similar even after GAO controlled for racial differences in criminal history, offense seriousness, and other legally relevant factors. (See pp.67-68.) Moreover, the report indicates that while there was some racial disparity for within-range sentences, racial groups did not differ in sentences that departed from the applicable guideline range. (See p.13.) Finally, Appendix III, which considered racial differences in sentencing patterns for robbery under the guidelines, concluded that the relatively small white advantage in sentencing for this offense was due to racial differences in offense and offender characteristics as opposed to differences in treatment based on race. (See pp. 50-51.) These important conclusions should be reflected in the body of the report, which those who receive the report are more likely to read than the lengthy appendices. In addition the report and appendices should omit any language that is misleading as to the report's actual findings.

Now on p. 92.

Now on p. 109.

Now on pp.67-68.

Now on pp. 137-138.

Further, the appendices do not always fully examine the results of the statistical findings. For example, GAO notes the statistics indicate that blacks were more likely than whites to receive top-of-the-range and bottom-of-the-range sentences rather than a mid-range sentence. However, the statistics also showed that the difference between whites and blacks receiving a top-of-the-range sentence is just over one percent and the difference between whites and blacks receiving a bottom-of-the-range sentence is 4 percent. (See Table II.2) If GAO had correlated such information, it may suggest that evidence of disparities in sentencing that disfavors blacks is not conclusive. Given the public perception of widespread anti-black discrimination in the justice system, we believe that the report should clearly indicate that it did not find consistent evidence of discrimination against blacks in sentencing.

Disparities Based on Differing Levels of Culpability. GAO reports the views of some probation officers and judges that the guidelines are unfair because they cause defendants with differing levels of culpability in a criminal organization to receive similar sentences. GAO indicated that this concern is based on the Sentencing Commission's policy to consider an offender's actual criminal conduct in sentencing, as opposed to the conduct of which the offender is charged and convicted, and its policy to make an offender accountable at sentencing for the conduct of other participants in a jointly undertaken criminal activity. However, the conclusion by those interviewed that the

guidelines are responsible for treating offenders with differing levels of culpability in a like manner overlooks the fact that the guidelines provide for reductions and enhancements in sentencing based on an offender's role in the offense. Moreover, the Commission has approved and submitted to Congress guideline amendments that would narrow some offenders' "relevant conduct" for purposes of determining their offense level, particularly where an offender's part in the jointly undertaken activity is limited. If Congress takes no action to the contrary, these amendments will go into effect on November 1, 1992. Notwithstanding the opinions of the interviewees, the GAO report should reflect more accurately the manner in which the guidelines operate and the changes likely to go into effect in November.

Possible Disparities Resulting from Prosecutorial Discretion. The report further suggests that racial disparities exist in prosecutorial decisions, but does not examine possible reasons for such disparities. Although the Department does not maintain empirical data by race concerning prosecutorial decisions, it has been some prosecutors' experience that members of some racial and ethnic groups are more willing to cooperate and negotiate plea agreements than members of other racial and ethnic groups If this is true then there may be variances along racial and ethnic lines, but it would not be due to an abuse of discretion.

PROSECUTORIAL DISCRETION

The Effect of Use of Prosecutorial Discretion. The report repeats a Commission study finding that in 17 percent of all cases the decisions made by the prosecutors as to charge reductions and other plea bargains have an effect on sentencing. We do not disagree with this statement; however, the report does not distinguish the nature and validity of those decisions. Most of these prosecutorial decisions which affect the sentence are justifiable. For instance, charges may have to be reduced or dismissed because of proof problems, new evidence or unexpected events in the investigation. Further, if a defendant offers substantial assistance, a prosecutor may file a motion for departure. Such decisions made by prosecutors certainly affect the potential sentence in a case, but these decisions are entirely proper and do not represent an abuse of the guideline system.

By not distinguishing between valid decisions and those made to undermine or evade the guidelines, or even noting that often these charging decisions are entirely permissible and within the sentencing scheme, the repetition in this report of the statement that these prosecutorial decisions affect the sentence in 17 percent of all cases really yields no useful information. It is implied that all such decisions are abusive and violate the sentencing guidelines, which is clearly not the case.

The discussion indicates that plea bargaining usually results in lighter sentences, but does not consider the likelihood of conviction of the subject defendant or the possibility of conviction of other defendants if plea bargaining and motions for downward departures were not used. Without the use of plea bargaining, there may not be a conviction and, consequently no sentence. It also should be noted that prosecutorial discretion does not undermine the ability of either the defense or the prosecution to make decisions based on the possible outcomes dictated by the sentencing guidelines. Because the guidelines enable a more accurate and reliable determination of the defendant's potential sentence, they promote more effective and efficient use of plea bargaining for both the prosecution and defense, as well as more informed decisions by all parties involved.

Controls over Prosecutorial Discretion. The report notes that some judges and defense attorneys believe that too much discretion has inured to the prosecutors under guideline sentencing. Those who make these comments do so without statistical support. The report, in fairness, should note that the Department of Justice has for some time set out national policies and procedures to be followed by its prosecutors. These policies are consistent with the goals set by Congress.

The Department has taken steps to ensure that all federal prosecutors follow these uniform procedures. The Department has provided federal prosecutors with thorough and explicit instructions about charging and plea agreement decisions. Such decisions are recorded and reviewed within each office. Further, review of these procedures under the guidelines is a part of every office evaluation. The Department has actively worked to ensure that all U.S. Attorneys and Assistant U.S. Attorneys are knowledgeable about the guidelines and Department policies.

Therefore, while prosecutors, under the guideline system, continue to exercise a large measure of discretion, as in preguideline practice, the Department intends that discretion be applied consistent with the goals of the guidelines and the Sentencing Reform Act. The report would be more complete if it acknowledged that these steps have been taken.

We appreciate the opportunity to comment on the draft report. We hope that you find our comments both constructive and beneficial.

Sincerely,

Harry H. Flickinger
Assistant Attorney General
for Administration

Comments From the Administrative Office of the U.S. Courts

L. RALPH MECHAM
DIRECTOR

JAMES E. MACKLIN, JR.
DEPUTY DIRECTOR

ADMINISTRATIVE OFFICE OF THE
UNITED STATES COURTS

WASHINGTON, D.C. 20544

July 31, 1992

Mr. Richard L. Fogel
Assistant Comptroller General
U.S. General Accounting Office
Washington, D.C. 20548

Dear Mr. Fogel:

This is in response to your June 30, 1992, letter asking for comments on the GAO draft report, Sentencing Guidelines: Central Questions Remain Unanswered. Please note that my response considers the draft solely from the perspective of the Administrative Office of the United States Courts (Administrative Office). The Honorable Vincent L. Broderick, Chairman of the Judicial Conference Committee on Criminal Law and Probation Administration, is providing a separate response for the Committee. The comments that follow are not meant to take issue with the wisdom of the Congress and the Sentencing Commission in implementing the Sentence Reform Act of 1984, nor with the appropriateness of the guidelines themselves. Rather, my comments focus on the GAO study and the findings contained therein.

The Administrative Office has collected a great deal of information on the impact of the guidelines on the judiciary, some of which the GAO analyzed and delineated in Appendix V (but which, for some reason, was not included in the main body of the report). As an example, GAO analysis of Administrative Office data revealed that the median number of days from conviction to sentencing increased 68% since the guidelines were implemented (from 41 days in the year ending June 30, 1986, to 69 days in the year ending June 30, 1990). The evaluators noted that the most marked increase occurred between the year ending June 30, 1988, and the year ending June 30, 1989, when nationwide implementation of the guidelines began.

GAO evaluators also used Administrative Office data to quantify the link between guidelines implementation and increases in case processing time frames. As shown in table V.8, the median number of months from case filing to disposition for all defendants nationally increased over 40% from 3.2 months in the year ending June 30, 1986, to 4.5 months in the year ending June 30, 1990.

A TRADITION OF SERVICE TO THE FEDERAL JUDICIARY

Mr. Richard L. Fogel
Page 2

The above examples are provided to demonstrate that the Administrative Office had a plethora of information available from which the evaluators could draw to support specific findings and recommendations, and I urge the evaluators to do so prior to finalizing the report.

Now on p. 4.

Of specific concern to the Administrative Office is the effect guidelines implementation has had on the judiciary's workload requirements. The evaluators note on page seven of the draft report that "...it **appears** that the guidelines increased **to some degree** the workload of judges, prosecutors, probation officers, and defense attorneys....While the available data precluded us from quantifying the amounts of change, the greatest increase appeared to be for probation officers" (emphasis added). Once again, the statement that data limitations are the primary cause of GAO's inability to complete its mandate is troubling to us; especially because the Administrative Office recently completed work on revision to the work measurement formula which involved collecting an enormous amount of detailed data on the workload of probation and pretrial services officers. Included in the measured tasks were those required because of the sentencing guideline legislation. The evaluators once again buried quantitative analyses in Appendix V. Using the work measurement information provided by the Administrative Office, the evaluators noted that a probation officer assigned to do presentence investigations could be expected to complete 57 a year, a decrease of 28% from a pre-guidelines implementation work measurement study conducted in 1986. As noted in the report, on the basis of a remarkably accurate estimate (developed by a panel of probation office managers in 1987) that determined that the number of reports a probation officer could do would drop to 55, Congress authorized 596 additional positions (368 probation officers and 228 clerks) for fiscal year 1991.

In summary, sufficient data is and has been available which, along with the wealth of published material on the guidelines, reflecting both supportive and critical viewpoints, should have been analyzed and included in the report to show the impact of the guidelines on the criminal justice system.

Sincerely,

L. Ralph Mecham
Director

cc: Honorable Vincent L. Broderick

Major Contributors to This Report

General Government Division, Washington, D.C.

Lynda D. Willis, Assistant Director, Administration of Justice Issues
James H. Blume, Assistant Director
William O. Jenkins, Jr., Assignment Manager
Deborah A. Knorr, Senior Evaluator
Mary Beth McJunkin, Evaluator
William J. Sabol, Social Science Analyst
Douglas M. Sloane, Social Science Analyst

Cincinnati Regional Office

Daniel J. Kirwin, Senior Evaluator
Norman A. Hofmann, Senior Programmer
Julie A. Schneiberg, Evaluator
Frank T. Lawson, Evaluator
Jennifer C. Jones, Evaluator