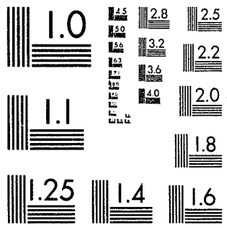
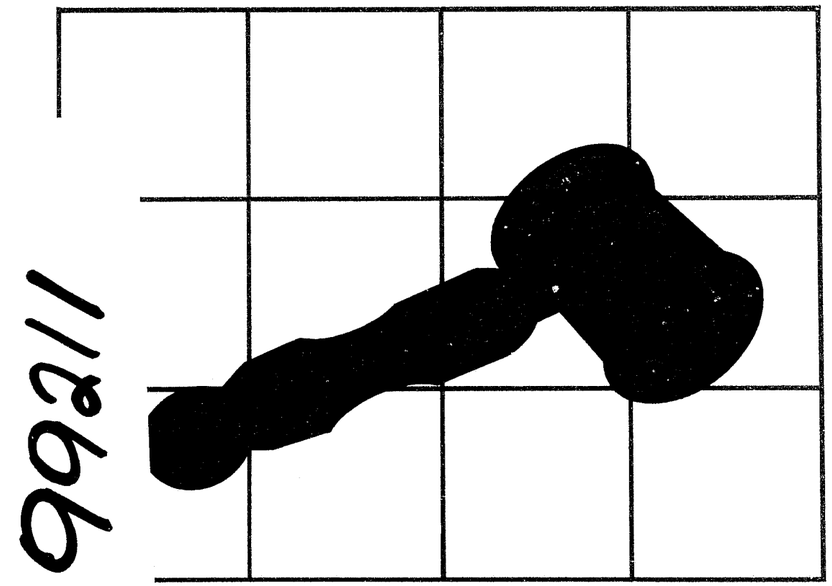


Submitted by Abt Associates Inc. to:
U.S. Department of Justice
National Institute of Justice

Guidelines Without Force:
An Evaluation of the Multijurisdictional Sentencing Guidelines Field Test



99211

GUIDELINES WITHOUT FORCE: AN EVALUATION OF THE MULTIJURISDICTIONAL SENTENCING GUIDELINES FIELD TEST

PRINCIPAL AUTHORS

Deborah M. Carrow
Judith D. Feins, Ph.D.

CONTRIBUTING AUTHORS

Beverly N. W. Lee
Lois Olinger

JULY 1985

Prepared for the National Institute of Justice, U. S. Department of Justice by Abt Associates Inc., under Grants 79-NI-AX-0118 and 82-IJ-CX-0029. Points of view or opinions stated in this document are those of the authors and do not necessarily represent the official position or policies of the U.S. Department of Justice.

ACKNOWLEDGEMENTS

This research has spanned five years, and has seen its share of changes during that time. Though both Maryland and Florida have continued their guidelines efforts, many of the people originally involved in the Multijurisdictional Sentencing Guidelines Field Test have moved on to other endeavors. The research staff of this evaluation has similarly changed with the passing of time.

Throughout the entire process, however, a few things have remained constant. One was the unfailing cooperation of the Maryland and Florida judiciary. In Florida, our special thanks are extended to Judges Marvin Mounts, Robert L. McCrary, Jr., Oliver L. Green, Jr. and the late Judge Everett Richardson. In Maryland, we would particularly like to thank Judge Marshall A. Levin. Many of his colleagues on the advisory board were also extremely helpful in our evaluation, including Judges Howard Chasanow, Audrey Melbourne, Richard Latham, John J. Mitchell, and Brodnax Cameron. Prosecutors and Public Defenders in each of the test sites were also extraordinarily cordial and cooperative, and we owe a word of appreciation to them as well.

The sentencing guidelines staff in both states undertook a difficult challenge and faced it well. We are indebted to them for their cooperation, information, and assistance throughout the evaluation and the drafting of our final report. We would particularly like to note the contributions of Kenneth Plante, Director of the Florida Multijurisdictional Sentencing Guidelines Project, and Dr. Patricia Nelson, Director of the Maryland Sentencing Guidelines Project.

Throughout this entire effort, Dr. W. Jay Merrill of the National Institute of Justice provided his support and encouragement. We are indebted to him, and to the National Institute, for the opportunity to participate in this important study.

Finally, we would like to thank our colleagues at Abt Associates for their help and support. Dr. Bradford M. Smith, the original project director, established and implemented the basic research design and provided his encouragement as we completed the project and drafted the report. Joan Mullen, Vice President of Abt Associates, provided thoughtful review and editorial assistance, and served as general sounding board for the policy conclusions of these volumes. Kenneth Carlson reviewed the manuscript and provided assistance in drafting selected portions of the report, while Dr. Jan Chaiken reviewed the impact analysis and helped us improve our report.

TABLE OF CONTENTS

TABLE OF CONTENTS
(continued)

LIST OF TABLES

LIST OF TABLES

(continued)

LIST OF FIGURES

LIST OF FIGURES

(continued)

LIST OF FIGURES

(continued)

CHAPTER 1

OVERVIEW OF THE MULTIJURISDICTIONAL SENTENCING GUIDELINES FIELD TEST EVALUATION

"The Magistrate is the guardian of justice, and, if of justice, then of equality also."

The experiment in sentencing described in this report was one of many attempts to preserve this Aristotelian precept in contemporary systems for the administration of justice. In many respects, the last generation has been a time of turmoil in the justice system. The number and seriousness of many crimes has increased. Old theories of the purposes of sentencing have been discredited. And from all sides come calls to take sentencing decisions out of the hands of judges: the public wants sentences harsher; legislators want to mandate them; and federal courts may simply prohibit them if prisons fail to meet Eighth Amendment standards.

In the face of these demands, perhaps the most remarkable fact of all is how little change has really occurred in sentencing. Despite a continuing search for alternative sentences, the basic options remain probation, incarceration and fines. Sentencing decisions are still made by judges, frequently working with sparse, unreliable information and trying to serve vague and conflicting goals. The "four horsemen" of sentencing purposes--deterrence, incapacitation, rehabilitation and retribution--have been debated at such length that no two observers agree on their meaning. We continue to try to serve all four masters despite proof of their mutual incompatibility. And we continue to rely heavily on the traditional forms of punishment despite the absence of proof that any of them accomplish any of these purposes efficiently--or at all.

Considering the ambiguity of the assignment, it is hardly surprising that judges differ in how they carry it out. Most state laws permit a wide range of latitude in sentencing and provide little or no guidance about whom to place at which end of the range. At most, the statute may specify factors which should be considered (such as the prior record of the offender or enumerated aggravating or mitigating circumstances), but since the law rarely stipulates how such consideration should affect the sentence, it is not clear that equality is increased. Many legal scholars and practitioners have expressed vigorous concern about the disparities that arise from the decentralized and ambiguous nature of sentencing. Yet non-judicial attempts to reduce disparity (in the form of legislatively fixed sentences) have often merely centralized the decision and displaced the discretion without resolving any of the basic ambiguities. In contrast, sentencing guidelines have aspired to reduce the ambiguity of sentencing--while still keeping the judge as the guardian of justice.

1.1 The Evolving Concept of Sentencing Guidelines

As envisioned by their earliest proponents, guidelines would provide a framework for structuring judicial discretion, thus introducing greater uniformity in sentence decisionmaking. The sentence range for a given offense would no longer be based solely on nominal offense categories, but would now accommodate and standardize other factors that were commonly weighed by sentencing judges. Since judges might routinely consider prior criminal record or other circumstances of the offender and offense in fashioning an appropriate sanction, there would now be a sentencing structure that explicitly recognized these variables. That structure would be reflected in a matrix that provided different sentence ranges for the different conditions that might be associated with the same type of offense.

The particular sentence ranges recommended by the guidelines would be based on an empirical analysis of prior sentencing practices. As such, the original concept was geared to address the information gap created by our highly decentralized system of judicial decisionmaking. In this model, disparate practices among sentencing judges were the result of a lack of information about one another's behavior; by sharing that information, consensual policies would emerge:[1]

> Guidelines will provide information to judges which has hitherto been unavailable to those either inside or outside the judiciary. It is, finally, our view that once judges of a given jurisdiction are accurately informed as to what they <u>have</u> been doing in the past, then they can more clearly focus on what they <u>should</u> do in the future.

In addition to the so-called "empirical" derivation of the guidelines, the initial concept also involved systems initiated and developed by the courts and administered by sentencing judges on a voluntary basis. Voluntary compliance would avoid any interference with the principle of judicial discretion and was seen as a possible counter to such infringement from determinate and mandatory sentencing proposals. Numerous statutory reform efforts had clearly demonstrated that "attempts to impose solutions by fiat rarely work" and had encouraged the architects of the guidelines concept to advocate a more benign, self-regulatory mechanism.

It was this original conception of guidelines that was implemented in the two-state field test described in this report. While both states adopted the central elements of what was then considered advanced practice, in the years following the design of the test in Florida and Maryland, the guidelines concept has expanded and matured. Alternatives to the empirical, voluntary guidelines paradigm have been created and tested. This report must now join a number of other publications in endorsing the technique in general, but seriously questioning its empirical foundation as well as its reliance on voluntary administration.

In observing guidelines development in various jurisdictions, other analysts have noted that the cachet of statistical analysis involved in empirically based guidelines may have led some participating judges to ascribe to the results a validity and precision that the data did not justify. Judges were reportedly reluctant--or unable--to intervene in the development process because it was difficult to separate technical from substantive issues. As a result, project staff began to make fundamental policy decisions which were only implicit in the analytic results. Moreover, among the single jurisdictions documented to date, it was sometimes difficult to obtain enough cases to resolve the ambiguities about how sentences were related to characteristics of the offense and offender. These problems were exacerbated because there really were large, (seemingly) random components to the sentences which eroded statistical precision. As a multijurisdictional test, the design efforts in Florida and Maryland were less hampered by the problems of sample size, but faced instead the challenge of resolving ambiguities in sentencing practice not only among individual judges but also among jurisdictions differing in their sentencing practices and norms.

Even supposing the statistical methods to be flawless--a tall order considering the novelty and complexity of the task--serious conceptual questions remain. If we suppose each judge to have followed an internally consistent rationale or sentencing objective in each decision, it does not necessarily follow that a statistical average of these rationales remains internally consistent. The same is true for jurisdictional differences. Averaging could merely result in a cancellation of competing philosophies and norms, yielding a result which no one believed. Moreover, as many have observed, it is questionable whether the old sentencing philosophies deserve to be encoded for future emulation. In Minnesota, for example, legislators were concerned that their guidelines system permit adjustments to help regulate prison populations. Others have been concerned about the need to adapt to changing public standards of conduct and to respond to general changes in the criminal environment. We heartily endorse this evolution of the guidelines concept where normative principles play a key role in determining what factors are to be considered in future sentencing and how they should be weighed. As to the voluntary application of guidelines, the obvious problem is that voluntary compliance is not guaranteed to change anything. At least two evaluations of guidelines systems developed and implemented at the same time as those in Florida and Maryland concur with the need to abandon gentle persuasion in favor of more effective compliance mechanisms:

> . . . sentencing guidelines must be given the force and effect of law. Voluntary sentencing guidelines do not work; the increased confluence of interests among judges, prosecutors, and defense attorneys that might engender the support needed to make guidelines work is inconsistent with the ideal of adversary proceedings. Accordingly, the idea of voluntary sentencing

guidelines should be abandoned, if indeed it was ever taken seriously.[2]

> It can be argued . . . that the responsibility for developing and implementing statewide guidelines properly lies with the legislature and not the judiciary . . .[3]

Ultimately, whether guidelines are accorded "the force and effect of law" through legislative mandate or court rule, an appropriate appeals or sentence review mechanism was and remains a crucial element if guidelines are to be considered a serious sentencing reform.

Finally, we would encourage those who might adopt the guidelines method to experiment with further variations on the original model. Considerable variation in sentencing practice remains uncontrolled by many guidelines systems. Prosecutors' plea bargaining strategies, the tactics of defense counsel, and varying judicial policies on the proper relationship between such factors as prison capacity and prison sentences, are all sources of potential disparity that remain to be addressed in future refinements of the concept.

While it may strike the reader that--with so much left undone--little has been accomplished, we reiterate our support for the concept of guidelines as a means for developing more principled sentencing policy. How do guidelines promise to make criminal sentencing "more principled"? The answer to this question lies in the process by which guidelines are developed. In the past, responsibility for sanctioning policy has been vested solely with the legislature. Recently, this historical approach has been called into question by researchers and practitioners alike:

> We've learned from many fields--most acutely with respect to criminal justice--that ongoing study and scholarship are not usually the strong aspects of legislative endeavor. The point is clear beyond cavil for such a topic as criminal sentencing. It implies no disrespect for our legislative leaders, but only a recognition of plain reality, to observe that sentencing, prisons, and correctional practices generally do not attract sustained and steadily coherent attention among lawmakers, state or federal. There are, of course, recurrent waves of passion about crime, in the streets or the boardrooms, and periodic convulsions of one kind or another. But the basic need to study, revise and improve the law goes largely unattended.[4]

To address the consequent need to provide the legislature with more deliberative policy support, a hallmark of the guidelines process is the establishment of an independent commission or advisory group charged by the legislature with the task of

considering the issues of sentencing policy, achieving a consensus, and forming a structure that will reflect the chosen course.

Exactly what that course should be is a matter of values for which there is no absolute answer. At one time or another, guidelines have been affiliated with a variety of sentencing rationales. Not surprisingly, these have varied according to the ideology and political motivation of the advocates. To some, greater uniformity has promised to produce a more consistent level of severity, all the better to serve the presumed punitive purpose of sentencing. Others have seen the promise of "just deserts" or strengthened ties to the rehabilitative purpose of criminal sentencing, or a means to serve the ends of "incapacitation." In fact, however, the guidelines concept--in its normative form--holds no automatic allegiance to any single side of the sentencing "purposes" debate. It is merely a technique for expressing sentencing policy--a decisionmaking aid that may serve whatever broader goals its designers choose to pursue. Depending on their construction and implementation, guidelines may increase or decrease or stabilize prison populations. They may reduce some penalties and increase others. Whatever the designated purpose, the guidelines are simply a tool for setting and implementing that standard. Setting the standard relies on a rational, consensus-building approach seldom seen in proposals for statutory sentencing reform. Implementing the standard offers opportunities for the kinds of adaptive changes in policy that more rigid revisions in sentencing structure are ill-equipped to accommodate. Thus, while this report documents problems involved in translating the concept into effective practice, to the extent that future sites involved in guidelines implementation can learn from the practical difficulties encountered by their predecessors, the efforts of Florida and Maryland will be richly rewarded.

1.2 Research Objectives of the Multijurisdictional Test

The basic design for this research began to take shape in 1978, with a grant from the National Institute of Justice to the Mitre Corporation to develop a plan for a multijurisdictional test of sentencing guidelines. Earlier feasibility research in individual jurisdictions had established the general mechanics of empirical guidelines formulation and showed that at least some courts were willing to participate in the process and use the results. That feasibility research left open, however, several major questions which this evaluation was designed to address.

The first of these was the feasibility of developing guidelines which encompassed more than one jurisdiction. A qualitative difference was perceived between purely local guidelines, which resolved only inter-judge (rather than interjurisdictional) differences in sentencing norms, and statewide guidelines, which would attempt to create a single standard for all courts in the state. Although the multijurisdictional test was not yet to encompass an entire state, it did involve urban, suburban, and rural courts collaborating to develop a single, uniform set of

guidelines. In answering this question of feasibility, two concerns were of paramount importance. First, would judges be receptive to guidelines that imposed a multijurisdictional rather than local standard of sentencing? Given the common reports of vast differences in sentencing norms among different regions in a state, this concern was indeed justified. Second, would the empirical development of such guidelines be feasible? Could the data collection, analysis, and design be carried out, and would they result in a reasonable set of guidelines for all participating jurisdictions?

The second major question to be addressed by the evaluation was the ultimate effect of guidelines on sentencing disparity. The data available to the evaluators of the single-jurisdiction tests had been rather limited, for a variety of reasons, and findings on whether sentences actually became more homogeneous under the new rules were also limited. Since the early data used in guidelines construction did suggest substantial individual variation among courts, the multijurisdictional study was conceived to diminish the risk of findings dominated by the idiosyncracies of one or two possibly atypical jurisdictions.

The implementation of guidelines in multiple jurisdictions added a further evaluation design possibility which had been unavailable to the single-jurisdictional studies. Statistical comparisons within one court were necessarily based on a contrast between data collected before and during guidelines implementation. Many aspects of the local environment were changing during these periods of observation, and it was difficult to say whether changes in sentencing patterns, where demonstrated, should be attributed to the implementation of sentencing guidelines or to some other concurrent change. In the multijurisdictional design, other courts in the same state were monitored to provide an indication of possible confounding trends that might influence sentencing patterns with or without guidelines.

Yet another question to be answered through this evaluation concerned the impact of guidelines on the broader operations of the courts and the criminal justice system. With the introduction of a new kind of sentence reform, standard operations of the courts might be influenced in unanticipated ways. Plea negotiations, filing practices, and court workload might be affected, as might correctional agency operations. The evaluation was charged with documenting the consequences--negative or positive--of implementing guidelines.

The final objective of the evaluation was to record the experience of the sites in the development and implementation process so that others might profit from their experiences. Too often the valuable lessons of experimentation are lost as time passes and memory of the early experience is altered by current events and operations. By following the guidelines experiment since its inception and documenting the processes and decisions of this effort, these lessons may be preserved and shared with others.

1.3 Methodology

The methodology employed in conducting this study involved both qualitative and quantitative research techniques. A brief overview of the basic approaches utilized is given below. Additional information regarding the design, data collection methods, and analysis procedures used is presented in later chapters of this report and in Appendices F, G and H.

Process Analysis. Without an understanding of exactly how a sentencing guidelines effort is implemented, it is impossible either to analyze data about sentencing or to interpret the results with confidence. The purpose of the process analysis, therefore, was to examine the manner in which the guidelines were developed in each jurisdiction and the way in which they were utilized by judges and other system participants. Among the topics explored in this study component were: the legal and organizational context in which the field test was conducted; the manner in which the guidelines were developed, including the participation of judges in the development process; the ways in which guidelines were implemented and their use monitored; support-building efforts on the part of project staff and advisory board members; and the response of the courts to the changes.

Four primary data collection methods were used to gather the process information:

1) Interviews with system participants. Extensive formal interviews were conducted with personnel of the court system in each test jurisdiction, including:

   - judges;
   - prosecutors;
   - defense counsel;
   - probation and parole authorities responsible for preparing presentence investigation reports; and
   - court administrators.

   Interviews were also conducted with relevant state officials, including parole board members, representatives of the state divisions of parole and probation, members of statewide study commissions on sentencing, state legislators, and representatives of the states' highest courts.

These interviews were conducted at several different points in the life of the field test: during the period of guidelines development, soon after guidelines were implemented, again after guidelines were firmly in place, and at the conclusion of the one-year test period. The interviews provided not only an operational description of the court systems before and after guidelines but also a picture of local support or resistance to the introduction of the sentencing reform. (Appendix K lists the individuals interviewed during the course of this study.)

2) Distribution of questionnaires. Several months into the guidelines test, written questionnaires were distributed to all judges, prosecutors, and defense attorneys in the test sites. Respondents were asked to supply information on guidelines operation, as well as on the perceived impact of guidelines on the court system.

3) Observation and documentation of the work of the sentencing guidelines project staff and Advisory Board in each state. Data collection activities included observation of Board meetings, observation of and interviews with project staff, interviews with Board members, and review of project documentation. On-site activities were complemented by telephone contact with Board members and project staff.

4) Examination of written resource materials, including:

- relevant statutes and case law in each state;
- available documentation on procedures and operations of the courts, prosecution, the defense, the probation and parole division, and the correctional system; and
- available aggregate statistics, including court case processing data.

Impact Analysis. The primary objective of the impact analysis was to assess the effects of guidelines implementation on sentencing uniformity and severity. The sentencing outcomes analysis relied primarily on information about actual sentencing decisions for burglary cases proceeding through the courts; data were gathered from court records. Focusing on a single offense category was intended to reduce a major source of variation in sentencing decisions--the nature of the offense charged.

The burglary data analysis made use of two different sets of contrasts to help assess whether there were observed changes in sentencing patterns and whether such changes could be attributed to the implementation of the guidelines. Time contrasts involved comparing data from before the guidelines went into effect with data from the test year. Geographical contrasts were also employed, through collection of data in court jurisdictions within each state that were not participating in the guidelines test. These comparison sites, which were selected to match as closely as possible the characteristics of the participating jurisdictions, were included so that we could distinguish more clearly between changes in sentencing behavior resulting from the introduction of guidelines and changes in the state court system generally. Simple descriptive statistics, as well as multiple regression models, were employed to estimate and statistically test for possible guidelines impact. Chapter 8 provides a detailed discussion of this approach.

In addition to the burglary analysis, the evaluation also made use of a hypothetical or simulation analysis. This component involved analysis of data gathered by asking judges and prosecutors to assign sentences to defendants in five hypothetical criminal cases. The case descriptions covered five different offenses and provided the facts upon which sentences were to be based. Like the burglary analysis, time contrasts and geographical contrasts were used to detect changes in sentencing patterns and whether these changes were due to guidelines. The results of this analysis are described in Appendix H.

Compliance Analysis. Guidelines use in the multijurisdictional field test was voluntary in two senses. Procedurally, the judges' cooperation was sought in completing paperwork and in actually consulting the scoresheets and grids when sentencing, but no sanctions for noncooperation were established. Substantively, the judges were not required to sentence within the guidelines. A certain number of extra-guidelines sentences were anticipated, with the judge asked simply to record the reason for sentencing above or below the recommended range.

The purpose of the compliance analysis was to examine the evidence on both procedural and substantive compliance in the Florida and Maryland test jurisdictions. Five basic questions were posed:

1) Did judges consult the guidelines in reaching sentencing decisions?

2) Were scoresheets completed on eligible cases?

3) Did sentences fall within the guidelines ranges?

4) Were reasons provided for extra-guidelines sentences?

5) What kinds of reasons were given for extra-guidelines sentences?

To answer these five questions, several kinds of data were collected and analyzed.

First, we gathered information on whether and how judges used the guidelines in reaching sentence decisions, through interviews with Florida and Maryland judges at the end of each state's test year. Second, we examined data sets developed by the guidelines projects from the scoresheets completed on cases sentenced in the test period. These scoresheets contained information about the offense and offender, the score elements and total, the sentence, and (if applicable) the judge's reason for going outside the guidelines. The scoresheet data formed the basis for answering questions about substantive compliance. Third, the burglary data described above also played a role in the compliance analysis. Because they represented the universe of burglaries sentenced during the test year, for that one family of offenses we were able to assess whether scoresheets were indeed completed and filed on all eligible cases. Finally, interview data from judges, prosecution and defense attorneys, and parole and probation officers allowed fuller understanding of the patterns found.

Synthesis. The findings from the process, impact, and compliance analyses together form the basis for the conclusions of this evaluation.

1.4 Overview and Summary

In this section, we provide readers with a roadmap to this report and a preview of its findings. The report as a whole is divided into two parts (and an additional set of appendices). It is structured so that Part I (Chapters 2 through 6) gives a full account of the multijurisdictional sentencing guidelines field test and of the major results of this evaluation. Part II, consisting of Chapters 7 and 8, contains a more detailed presentation of the quantitative analyses of judicial compliance with the guidelines and sentencing impacts.

Part I opens with an account of the origins of efforts to reform judicial decisionmaking. Chapter 2 reviews the debate over disparity in criminal sanctions and the genesis of guidelines as one approach to limiting parole board or sentencing discretion. It describes the several federally funded research projects about guidelines which were immediate forebears of the multijurisdictional field test. In so doing, it makes clear that the empirical, voluntary nature of the multijurisdictional test--features which are criticized in this and other reports--represented the only paradigm being implemented or analyzed at the time the Florida and Maryland projects began.

Chapter 3 provides an overview of the field test, from its origins with the National Institute of Law Enforcement and Criminal Justice (now the National Institute of Justice) to the present-day status of sentencing guidelines in Maryland and Florida. Its narrative previews certain important topics--guidelines development,

implementation and impact--to which the succeeding chapters return for in-depth analysis. The objective of Chapter 3 is to provide readers with an historical overview of guidelines activity in Florida and Maryland, establishing a context for the more detailed, evaluative discussions that follow.

In Chapter 4, the experience of Florida and Maryland in the development and design of empirical guidelines is examined for two purposes: to highlight issues relevant to the creation of other guidelines systems and to assess the advantages and drawbacks of the empirical approach. Both states' guidelines projects followed the development sequence mandated in the field test: construction of guidelines using information about past sentencing practices gathered from court data; the application of statistical techniques in order to discern patterns in sentencing practice among the participating local jurisdictions; and translation of these patterns into guidelines grids that related certain recommended sentences to specific characteristics of the offender and the crime. However, there were important differences in how the two projects carried out these tasks. The contrasts between them in the development process and in the resulting guidelines are the basis for the following conclusions:

- Policy decisions play a more critical role, even under the empirical approach, than has previously been recognized, and the ability of guidelines to reduce sentencing disparity depends on these decisions far more than on the empirical techniques used in development.

- The greater political acceptability of empirical guidelines in the eyes of state legislatures and the judiciary at large does not translate easily into support for specific guidelines, which can differ markedly from local patterns and norms because they represent an amalgam of practices from various jurisdictions.

- The development of empirical guidelines entails a sharing of judicial authority, although it is less explicit than that usually required to develop normative guidelines. Whereas the latter typically involves representatives of the defense, prosecution, corrections and the public, the former results from the critical role of a technically trained staff and the blurring of boundaries between technical and policy decisions.

- Greater emphasis on the normative process of guidelines development can help to overcome many of the methodological, logistical, and political difficulties associated with more empirically based guidelines.

Chapter 5 focuses on key issues in the implementation of sentencing guidelines. The multijurisdictional field test was an experiment in the use of judicially initiated, voluntary guidelines. These features had significant implications for the potential impact on sentencing behavior. Through analysis of three aspects of implementation--support building, accommodating local procedures and norms, and monitoring compliance--we examine the achievements and problems of the Florida and Maryland projects with respect to bringing about sentencing reform. The analysis shows that:

- The benefits of a judicial mandate were overemphasized in the areas of local support and compliance. Judicial authority was not sufficient to assure the cooperation of prosecution, defense, or all individual judges.

- The focus on judicial authority tended to exclude the prosecution and defense from the guidelines development process, yet only by their early involvement could an adequate foundation of support have been built.

- Planning for guidelines systems must take into account differences in local procedures and conditions, in such a way that implementation can be tailored to them and thus give guidelines a greater chance of local acceptance.

- Guidelines monitoring can play an important role in facilitating implementation and tracking compliance, but its potential impact is limited under voluntary guidelines. Compliance in Florida and Maryland was only modest in some respects (see Chapter 7 findings), largely as a result of implementation difficulties but ultimately as a consequence of the voluntary nature of the field test.

- Voluntary guidelines systems are unlikely to be fully successful. Stronger implementation mandates, including legislative authorization or court rule, are needed. In addition, appeals mechanisms for extra-guidelines sentences can help to ensure compliance while serving as a source of continued refinement of the guidelines policy.

In Chapter 6 we turn to an account of how the several participating jurisdictions in each state adapted to the coming of sentencing guidelines. This adaptation was interactive: there were changes in local court operations due to guidelines introduction and changes in how the guidelines operated as a result of

differences in local norms and procedures. Previous studies have observed the resistance of court systems to the organizational changes associated with various sentencing reform efforts. Our findings of limited guidelines impact on criminal sentencing (see Chapter 8 results) are explained, in part, by the following conclusions about adaptation:

- In most sites, the guidelines had very little effect on prosecutor's screening and charging behavior, due in large part to certain design features of the systems (e.g., grids and scoring).

- There were strong effects on the plea negotiation practices of the courts where judges were committed to guidelines use. The most common stemmed from the fact that the guidelines ranges functioned more as upper boundaries on the sentences imposed than as lower boundaries on negotiated terms.

- The facility with which discretion shifts between judges and prosecutors represents the most significant challenge to guidelines effectiveness, because the charging and negotiating functions strongly influence the scoring (and thus the guidelines range) for any case.

- Guidelines' intent and use can be subverted in actual practice. Guidelines designs which minimize the potential for disrupting court operations, combined with careful monitoring and training activities, are crucial if the intent of the guidelines is to be realized.

Chapters 5 and 6 drew extensively on the evaluation's quantitative findings about compliance with the guidelines and about their sentencing impacts. Details of the analyses that produced those findings are the substance of Part II of this report.

As discussed in Chapter 7 of Part II, compliance covers a range of issues, from completing necessary paperwork--i.e., filling out scoresheets for each eligible case and providing reasons for out-of-range sentences--to using guidelines in judicial decisionmaking and rendering sentences according to the recommended ranges. While there was general awareness of the guidelines, judges varied in the extent to which they used them in sentencing. Overall, we estimate that in Florida, only about 57 percent of the cases eligible for guidelines had scoresheets completed and filed; in Maryland, this percentage was approximately 70 percent (although the rate in Maryland reflects some loss from editing by the central staff). The analysis also reveals that:

- In Florida, there was considerable site variability, not only in judicial use of guidelines but also with respect to the relationship between actual and recommended sentences and the provision of reasons for extra-guidelines sentences. Rates of compliance were far more uniform in Maryland.

- While the vast majority of sentences were within the guidelines ranges in both states, the proportion of scoresheets sentences falling within the ranges in Maryland was relatively low, considering the fact that Maryland's sentencing ranges were very broad and overlapped considerably.

- Reasons given by Maryland's judges frequently challenged the guidelines by citing factors already taken into account in the offender or offense scores; this was less often the case in Florida.

Chapter 8 describes the impact of the multijurisdictional field test on sentencing disparity and severity. Examination of data on all burglary cases sentenced during the test year allowed us to compare sentencing for a single type of offense during the test period with sentencing during an earlier period and in a set of nonparticipating (comparison) sites. Summarizing the results of these analyses briefly, we found that in all but one of the eight participating jurisdictions, the introduction of sentencing guidelines had at best a very modest impact on sentencing behavior:

- In Florida, sentence severity increased slightly in the test and comparison sites. However, there was no change in the uniformity of burglary sentences in the Florida test sites, relative to prior sentencing patterns or to the decisions of judges in the comparison sites. Thus, the primary objective of guidelines--reduction of disparity within and across jurisdictions--does not appear to have been met under the guidelines test in Florida.

- In Maryland, sentencing variation for burglary cases in the urban jurisdiction was significantly lower during the guidelines test year than in the year before guidelines use. However, no changes in sentencing variation were detected in the other three Maryland test sites or in the comparison sites. When all cases from all Maryland test sites were considered, there was an overall decrease in variation which was due to the effect of the Baltimore City courts. Burglary sentences of the test site judges in Maryland increased in severity with the introduction of guidelines, though there was also a trend toward more severe

sentences in the nonparticipating sites. Taken together, the analyses suggest that although the guidelines' impact was limited, the field test has demonstrated that there is indeed a potential for sentencing change under the guidelines system.

The empirical findings on sentencing guidelines are thus somewhat mixed. Compliance fell well below expectations. With one exception, sentencing disparity was not reduced, though the problems with implementation and compliance may very well account for this finding. Still, given the positive finding in a large urban jurisdiction and the considerable accomplishments of both states in moving towards a guideline system, we do not believe that our findings invalidate the concept of guidelines as a method for improving the uniformity of criminal sanctions. Rather, they should prove useful in the further evolution and refinement of guidelines theory and practice. The Florida and Maryland projects have made a substantial contribution to the knowledge required for effective guidelines efforts in their own states and wherever guidelines will be developed and implemented.

FOOTNOTES

1. Leslie T. Wilkins, Jack M. Kress, Don M. Gottfredson, Joseph C. Calpin, and Arthur M. Gelman, Sentencing Guidelines: Structuring Judicial Discretion (Washington, D.C.: U.S. Daprtment of Justice, National Institute of Law Enforcement and Criminal Justice, 1978), p. 32.

2. William D. Rich, L. Paul Sutton, Todd N. Clear, and Michael J. Saks, Sentencing by Mathematics: An Evaluation of the Early Attempts to Develop and Implement Sentencing Guidelines (Williamsburg, VA: National Center for State Courts, 1982), p. xxix.

3. Richard F. Sparks, Bridget A. Stecher, Jay S. Albanese, and Peggy L. Shelly, with a chapter by Donald M. Barry, Stumbling Toward Justice: Some Overlooked Research and Policy Questions about Statewide Sentencing Guidelines (Newark, NJ: School of Criminal Justice, Rutgers University, 1982), p. 97.

4. Franklin E. Zimring and Richard S. Frase, The Criminal Justice System: Materials on the Administration and Reform of Criminal Law (Boston: Little, Brown and Company, 1980), p. 999.

PART I

Part I of this report provides a comprehensive account of the development and impact of sentencing guidelines under the multi-jurisdictional field test. After a review of the historical development of guidelines (Chapter 2) and the background and results of the field test (Chapter 3), it turns to a detailed analysis of three primary areas of interest. In Chapter 4 we examine guidelines development, focusing on the implications of the empirical approach. Chapter 5 explores implementation in Florida and Maryland, emphasizing the activities required to make voluntary systems effective and discussing the ultimate advisability of a stronger guidelines mandate. Finally, Chapter 6 assesses the variety of impacts that may result from guidelines introduction, including system adaptation and sentencing change.

CHAPTER 2

THE CONCEPT AND EARLY PRACTICE OF SENTENCING GUIDELINES

Dissatisfaction with the rehabilitative value of America's prisons and growing concern over what appeared to be unbridled discretionary power afforded judges and parole boards have led to numerous revisions in sentencing law and practice designed to reduce the indeterminacy of sentences and circumscribe the discretion of those who determine the length of incarceration. First implemented in 1976, sentencing guidelines is one alternative. Instrumental in the support of this concept, the National Institute of Justice (formerly the National Institute of Law Enforcement and Criminal Justice) has sponsored several sentencing guidelines research and evaluation efforts. The field test of multijurisdictional sentencing guidelines, funded by NIJ, represents its most recent endeavor.

This chapter gives an account of the historical background of sentencing guidelines. (An overview of their use around the country is provided in Appendix A.) The chapter opens with a brief review of the shifting currents of thought concerning sentencing goals and the exercise of judicial and parole discretion. It then gives a summary of the major pieces of research through which empirical, voluntary guidelines were first developed and then evaluated. Most guidelines efforts in U.S. courts today can trace their lineage to this NIJ-sponsored research and share these two primary characteristics. The chapter concludes with an introduction to the multijurisdictional field test and this evaluation, which are more fully discussed in Chapter 3.

2.1 Historical Background

2.1.1 The Rehabilitative Model and the Rise of Indeterminacy

Originally, America's penal laws, modeled after England's and composed of statutes which prescribed fixed sentences for specific offenses, were based on a theory of retribution. Although a few early attempts at reform presaged a trend toward rehabilitation, it was not until after the Civil War that rehabilitation was embraced as the goal of criminal law. At the National Congress on Penitentiary and Reformatory Discipline in 1870, reformers declared, "the surpreme aim of prison discipline is the reform of criminals, not the infliction of vindictive suffering."[1]

Parole was first adopted in 1876 at the Elmira Reformatory in New York State.[2] Central to this approach were the assumptions that reformation was the "right of the convict" and that every prisoner should be given "special treatment"; that the way to achieve this was to permit prisoner authorities "to lengthen or shorten the duration of the term of incarceration."[3]

A key element of the rehabilitative goal was the view that incarceration was imposed as a means of treating the offender, rather than a means of punishment. Prisons were seen as hospitals, of sorts, where the environment could be controlled to change the offender and enable him to resume a productive role in society. Fixed sentences of the antebellum period were obviously incompatible with an approach which required individualized treatment plans. They permitted neither the extension of a sentence for those offenders needing longer periods of treatment nor the early release of those responding favorably to rehabilitative efforts.

Influenced by penal efforts of Maconochie in Australia from whom the notion of "good time" derives, and Crofton in Ireland, who also employed positive reinforcement contingencies, American prison reformers generated a movement which led to the widespread acceptance, between 1880 and 1920,[4] of indeterminate sentencing with parole--a system which answered the needs of a rehabilitative approach. Under such a system the sentencing judge specifies the range of permissible time to be served by establishing a maximum and sometimes minimum sentence range, and the parole board adjusts the sentence according to rehabilitative progress made. By the 1930s almost every state had adopted an indeterminate sentencing structure.

2.1.2 The Failure of Indeterminacy: Concern Over Sources of Disparity

Although the new laws attempted to take individual circumstances into account and thereby, at least theoretically, be fairer, the rehabilitative ideal and the indeterminacy in which it operated came under attack by judges, prisoners and social reformers in the late 1960s. Criticism was focused on two major issues: the effectiveness of rehabilitation and the fairness with which offenders were sentenced and released.

Prison uprisings drew national attention to the grossly inadequate conditions of confinement. It became clear that prisons were providing custodial care, at best, and the notion that they could, on the medical model of treatment, "cure" criminals fell into disfavor. Empirical research lent scientific credibility to the assertion that the rehabilitative approach was ineffective. Most prominent among these studies was a review of all available reports (231 studies) on attempts at rehabilitation made between 1945 and 1967. This study concluded that rehabilitative efforts, to date, had no appreciable effect on recidivism.[5]

Largely as a result of prisoners' protests, the issue of fairness was brought to the fore. Under the rehabilitative model, making adjustments in the amount of time served based on rehabilitative progress seemed not only desirable but more reasonable. Given that the goal of rehabilitation seemed no longer feasible, judges' ability to assess treatment needs in rendering sentences and parole boards' ability to evaluate prisoners' progress in treatment programs in determining release came into question. Decisions extending or shortening length of confinement which were formerly considered individualized allowances were now perceived as being arbitrary and

unfair. The apparently dissimilar treatment of similarly situated offenders led to concern over constraining judicial and parole board decisions.

Judicial Discretion

One of the most influential critics of indeterminate sentencing practices and the disparity to which they lent themselves was Judge Marvin Frankel, who noted that "the almost wholly unchecked and sweeping powers we give to judges in the fashioning of sentences are terrifying and intolerable for a society that professes devotion to the rule of law."[6] Indeterminate sentencing practices, he argued, were arbitrary--indeed, irrational--and had "produced more cruelty and injustice than the benefits its supporters envisage[d]."[7] Other critics of the disparate treatment of similarly situated offenders included Gaylin, Morris, Orland, Churgin and Curtis.[8]

Systematic research into possible inequities in sentencing began in the 1920s with a work by Thorsten Sellin entitled "The Negro Criminal: A Statistical Note," in which he introduced the topic of judicial discrimination. Comparing the biennial reports of the attorneys general of several southern states, he found that blacks were discriminated against in criminal proceedings and that differential treatment by police and courts artificially increased apparent black criminality.[9] The first attempt to measure sentence variation empirically in terms of observable criteria was made by Guy Johnson in 1941 in his study, "The Negro and Crime." He investigated homicides in North Carolina, Georgia, and Virginia from 1920 to 1940 testing the hypothesis that sentence severity was primarily a function of race of the victim and secondly a function of race of the offender (i.e. that black assailants of white victims would be given the longest sentences). He found that of 141 black offenders with black victims, only seven were given life sentences and no death sentences were imposed. In 22 cases of black offenders with white victims, however, seven received life sentences and six death sentences. He concluded that while black offenders of black victims were treated with relative lenience, black offenders of white victims were treated harshly.[10]

Subsequent research on sentencing disparity has focused on two areas: providing empirical evidence of its existence and determining what factors account for it. The latter group of studies has attempted, with limited success, to explain disparity in terms of characteristics of the offender--race being the most frequently cited variable.[11]

Most studies on sentencing have relied on information about actual past sentencing decisions made by different judges in different cases. This makes it difficult to determine whether observed differences, if any, in sentences resulted from warranted or unwarranted sources of variation. Warranted variations are those caused by factors generally considered legitimate in sentencing, e.g., different sentences for different offenses. Unwarranted variations are those caused by factors generally considered to be improper sources of variation, e.g., different sentences for otherwise

equally situated defendants. Figure 2-1 presents the specific elements for each of these sources of sentencing variation. It also points out important areas where there is no consensus about the legitimacy of a factor's influence on sentencing.

A study conducted with fifty federal district court judges of the Second Circuit used a different research tactic to demonstrate clearly that judicial variation does exist. The judges were asked to impose sentences on the same set of twenty hypothetical cases based on actual presentence reports selected as representative of sentencing in courts. They imposed highly disparate criminal sanctions when given precisely the same information about the offenses and offender.[12]

The clarity with which the Second Circuit study demonstrated the existence of unwarranted disparity and its source is, however, rare among studies of sentencing variation. Indeed, few studies of sentencing have successfully measured individual or combined effects of specific factors or assessed the degree to which their significance may vary from court to court, across regions and over time.

Despite the lack of hard empirical evidence of sentencing disparity and its causes, criticism of the broad discretionary powers afforded judges under indeterminate sentencing structures generated a great deal of interest in controlling judicial discretion.

The Parole Board Decision

Under the indeterminate sentencing structure, the decision regarding length of incarceration was a responsibility shared by the sentencing judge and the parole board. While the judge decided on a range of time to be served, the parole board determined the actual amount of time served by setting the date of release based upon rehabilitative progress. At the same time that discontent with rehabilitation grew and judicial discretion came into question, the issue of uncertainty about when prisoners would be released emerged as a major prisoners' rights issue.

The American Friends Service Committee in Struggle for Justice called the uncertainty surrounding release "one of the more exquisite forms of torture . . . [which] contributed to the dehumanization and personal disintegration of penal servitude."[13] Frankel echoed the popular belief that "parole boards operate without orderly and uniform criteria for judgment, often moved by 'political' pressures or the winds of public opinion without the benefit of mature and organized wisdom."[14] Calling them "unpredictable, unexplained, and unexplainable," he asserted that parole boards kept the rules about what would gain release a mystery to prisoners.

Aside from being criticized on moral grounds, parole boards came under scrutiny regarding the question of disparity. "Charges of lack of procedural due process, arbitrariness, capriciousness, defensive self-protectiveness, failure to specify reasons for decisions, and working at cross purposes to rehabilitation were among the complaints."[15]

FIGURE 2-1

Sources of Sentencing Variation

Warranted Variation	Unwarranted Variation	Unresolved Value Judgments*
• Conviction offense. • Defendant's prior conviction record. • Deliberateness of the criminal act. • Resulting harm of the criminal act. • Stability of the defendant's background.	• The dissimilar treatment by <u>different</u> judges within the same jurisdiction of equally situated defendants--i.e., a defendant's sentence should not depend on which judge he gets. • The dissimilar treatment by the <u>same</u> judge of equally situated defendants--i.e., a defendant's sentence should not depend on the judge's mood or prejudices. • The dissimilar competence/experience of Assistant State Attorneys--i.e., a defendant's sentence should not depend on which Assistant State Attorney is assigned to the case. • The dissimilar competence/experience of defense attorneys--i.e., a defendant's sentence should not depend on a defendant's ability to pay for legal services.	• The dissimilar treatment between jurisdictions of equally situated defendants--e.g., lesser sentences for armed robbery in urban as compared with rural jurisdictions. • Vulnerability of the victim. • Evidence of the defendant's remorse. • Defendant's prior arrest record. • Defendant's prior juvenile record. • Offense as charged. • Actual and perceived parole eligibility. • The conditions of confinement and/or level of crowding in jails and prisons. • The availability of sentencing alternatives.

*The issue of the dissimilar treatment of unequally situated defendants--e.g., lesser sentences for income tax evasion than for stealing government checks--is beyond the scope of this review.

Most empirical research on disparity has focused on the courts. Studies substantiating parole board disparity or attempts to systematically isolate the factors which produce disparity in paroling decisions are few. The studies that have attempted to account for disparity in parole board decisions, like most studies on factors in sentencing disparity, have focused primarily on race of the offender.[16]

Criticism of the parole board decision centered around the uncertainty of release and the mental anguish it caused prisoners. Though the evidence was largely anecdotal and not conclusively validated by empirical research, parole boards were also attacked for being arbitrary. Interest grew not only in constraining judicial discretion but in limiting parole board decisions as well.

2.2 Limiting Judicial and Parole Board Discretion

In order to rectify the perceived inequities in the indeterminate sentencing structure and in response to the rejection of the rehabilitative model, reformers sought to limit the discretion of public officials responsible for the administration of criminal sanctions. Legislative reforms of both parole board and judicial decisions were one possible alternative. By introducing new criminal codes, parole release might be abolished and prescribed sentences could be specified. Other less rigid avenues of reform such as guidelines would still structure the discretion of parole boards and judges without necessarily confining them to exact terms.

Lawmakers have responded swiftly to the debate about the need for structuring parole board and judicial discretion in order to reduce disparity. Mandatory or mandatory minimum sentencing and determinate sentencing are two common legislative reforms of sentencing and release practices. Not incidentally, the popularity of these reforms may stem from their perceived potential to increase sentence severity as well as uniformity.

Mandatory sentencing constrains judicial discretion by stipulating that the offender must be incarcerated. Under this scheme the judge's choice to suspend sentences or grant probation is eliminated. A common variant of mandatory sentencing is mandatory minimum sentencing which further requires a specific minimum length of time to be served. Usually, such laws apply only to certain offenses, especially those involving firearms and narcotics. Florida, Michigan and Massachusetts, for example, have mandatory minimum sentencing laws for firearm-related offenses.

Determinate sentencing, also known as presumptive or flat-time, has several variants. Such legislation sharply curtails parole board discretion regarding release and to varying degrees, judicial discretion regarding length of time to be served. The most important feature of determinate schemes is that, generally, the date of release is known at the time a sentence is imposed. Under determinate schemes, save for good time earned, any administrative decision as to release is eliminated.

Departments of probation and parole may, however, retain their supervisory function over offenders once they are released. The extent to which judicial discretion over length of sentence is curbed under determinate sentencing codes varies widely from state to state; similarly, the degree of determinacy concerning the IN/OUT decision also varies.

In 1976, Maine became the first state to adopt a determinate sentencing code,[17] abolishing both the parole board's release authority and its post-release supervisory function. The judge, then, has sole responsibility for imposing fixed sentences up to legislatively prescribed maxima. Clearly, judges in Maine have considerable latitude in choice of sentence length.

California, the second state to implement a determinate sentencing code (in 1977), abolished the parole board release decision but retained its role as post-release supervisor. Judicial discretion regarding length of confinement in California is the most sharply curtailed among states with determinate sentencing codes. Judges must choose one of three specific sentences once the conviction offense is known. The mid-length sentence is the "presumptive" one which judges must impose in the absence of mitigating circumstances meriting the shorter sentence or aggravating circumstances warranting the longer sentence. Judges may lengthen the presumptive term by prescribed amounts under certain circumstances known as enhancements (e.g., if there was great bodily harm).[18] Other states that have adopted determinate sentencing codes are Colorado, Connecticut, Illinois, Indiana, New Mexico, and North Carolina.[19] Minnesota has also established determinate sentences; however, unlike other states, the presumptive sentence ranges were developed by a legislatively established Sentence Commission and are presented in the form of a guidelines matrix.

Appellate review of sentences has been proposed as a judicially-oriented means of controlling sentence variation and developing a "case law" of sentences which could serve as a standard for judges at the trial court level. Support for this approach was voiced in a variety of circles, including the American Bar Association[20] and Judge Marvin Frankel,[21] and appellate review systems have been instituted in a number of states.[22] As originally developed, appellate review was limited to affirming or reducing the original sentence; however, modifications in concept and practice introduced over the years have allowed for increases as well as decreases in sentences, and appeal by the prosecution as well as the defense.

Sentence review conducted at a "peer" level by a panel of three or more trial judges has also been suggested. The review may reconsider the appropriateness of the punishment and may be used to modify, mitigate, or even overturn the imposed sentence. This procedure is used in Maryland, where all defendants who have received sentences of two or more years are eligible for sentence review. Application for review, which does not necessarily stay the execution of sentence, must be filed within a certain length of time after imposition of sentence. As in trials, the offender is entitled to counsel, and the panel may request a copy of the presentence report,

review a transcript of the trial, and/or conduct a hearing on the matter. The sentence review panel may order a different length of sentence from that imposed by the trial judge.[23]

Sentencing commissions were proposed by Judge Marvin Frankel as a possible remedy for the ills of indeterminate sentencing. He envisioned a commission on sentencing as a:

> . . . permanent agency responsible for (1) the study of sentencing, corrections, and parole, (2) the formulation of laws and rules to which the studies pointed, and (3) the actual enactment of rules subject to traditional checks by Congress and the courts.[24]

The establishment of a United States Commission on Sentencing, fashioned after Frankel's proposal, is part of a bill first introduced in 1973 and recently passed by Congress.[25] Originally known as S.1, the Sentencing Reform Act allows the sentencing judge, instead of the parole board, to determine the precise sentence that an offender would serve, within a narrow range of guidelines prescribed by the Commission. These guidelines would articulate, for the first time, the general purposes and goals of sentencing to be considered by the judge prior to imposing a sentence. If the judge were to impose a sentence outside the guidelines, he would be required to provide written reasons for doing so, and appellate review to a higher court would follow. The Act was envisioned as a means to "curb judicial sentencing discretion, eliminate indeterminate sentences, phase out parole release, and make criminal sentencing fairer and more certain."[26]

Sentencing commissions have been legislatively adopted in Minnesota, Pennsylvania and Washington. In Minnesota, sentencing guidelines developed by the commission went into effect on May 1, 1980.[27] In Pennsylvania, sentencing guidelines were introduced and defeated in legislature for several years before the bill authorizing creation of the Sentencing Commission passed in 1978. Guidelines developed by the commission went into effect in July 1982. In July 1981, a sentencing commission was established by the Washington legislature to develop new guidelines which would replace those used in the state's Superior Courts. The commission submitted its report to the legislature in January 1983. Guidelines were implemented in 1984.

Guidelines assist decisionmakers in arriving at individual sentences or release dates in a framework of broader policy. The tool provided by guidelines to judges or parole boards is usually a two-way grid containing scores for the seriousness of the offense and the offender's criminal history or parole prognosis. For each combination of scores, an appropriate disposition--type of sentence, length or range of time to be served, time before release--is indicated.

Guidelines need not be mutually exclusive with other methods of limiting discretion. For example, they may be coupled with appellate review of sentences or they may be the means provided by sentencing commissions to guide judges in their decision-making. In such instances, use of the sentencing guidelines may be mandatory. However, in most cases, guidelines have been adopted voluntarily by the judiciary and are viewed as informational rather than binding. Even the common requirement of a written reason for sentences deviating from the guidelines range tends to be justified as a feedback mechanism for revision of the "typical" decisions indicated in the grids.

The largely voluntary nature of sentencing guidelines has been a matter more of historical evolution than of conscious choice, although the political rationale--self-reform to defend judicial discretion against legislative curbs--is strong. Similarly, the grounding of most guidelines systems in data about past sentencing decisions is only one option, albeit the option most thoroughly tested in research and practice. Although the examples of another approach are limited in number, they are extremely informative. They suggest that guidelines may be explicitly normative, formulated by policy-making bodies (such as sentencing commissions). The content of such guidelines may deliberately differ from the patterns of past decision-making. Similarly, guidelines may carry the force of law, with grid ranges presumptive rather than advisory in nature. As we examine the genesis of guidelines, it will become evident how empirical, voluntary guidelines came to be the dominant form. However, the reader should bear in mind the broader concept of guidelines posed here, even as we trace the origins and influence of guidelines more narrowly defined.

2.3 The Genesis of Guidelines

In this section, we briefly review the genesis of the concept of guidelines, first as applied to parole decisions and then to judicial sentencing in criminal cases. As we shall see, the earliest guidelines efforts were based on empirical analysis of past decisions, a heritage that has strongly influenced contemporary practice. Here we trace the lineage of early guidelines efforts, from the work for the U.S. Board of Parole in the early 1970s through the evaluation of the sentencing guidelines pilot implementation released in 1980. Not only the multijurisdictional field test, but also guidelines efforts throughout the country, have grown from these common roots.

2.3.1 The Development of Parole Guidelines[28]

In the early 1970s, the U.S. Board of Parole (now the U.S. Parole Commission) came under attack on two fronts: the uncertainty of prisoner release and the apparently arbitrary manner in which it made its decisions. To address such criticisms, the Board, in collaboration with a group of researchers led by Don M. Gottfredson of Rutgers University, began a study to articulate general paroling policies, in order to improve procedures for parole decision-making. The study involved three distinct pieces of research designed to:

- Determine the variables parole board members used in deciding whether or not to grant parole;
- Develop an objective measure of parole prognosis; and
- Establish a method for producing more consistent offense severity ratings, thereby making decisions more equitable.

First, in order to determine the variables used in the parole decision, researchers tested four factors, selected on the basis of prior research, as predictors of the release decision:

- Severity of the present offense,
- Institutional program participation,
- Institutional discipline, and
- Chances of favorable parole outcome, defined as: no new conviction resulting in a sentence of sixty days or more, no return to prison for technical violation, and no outstanding absconder warrant.

Using a sample of actual cases, parole board members evaluated each case on the above four dimensions. The two factors that were most highly correlated with the initial decision to release were seriousness of offense and chances of favorable parole outcome. In subsequent decisions--cases in which the parole board had previously denied a prisoner release--the rating of institutional discipline was more highly correlated with the outcome of the decision.

Since parole prognosis and severity of offense were the variables considered most important in the initial parole decision, researchers sought to develop more systematic measures of these variables. To establish an objective measure of parole prognosis, they conducteed a two-year follow-up study on 2,500 inmates released from federal prisons on parole in 1970 and 1972. Researchers gathered information items about each case which they hypothesized would predict parole board members' ratings of parole prognosis. These were distilled down to nine salient factors, which were used to construct a scale for predicting parole prognosis. Tested on other data, the salient factor scale retained its predictive power.

Finally, the researchers sought to develop consistent offense severity ratings in order to make decisions more equitable. Parole board members were asked to rank sixty-five offense descriptions by seriousness of the offense. Two ranking exercises were conducted. The first time board members were told not to deliberate for more

than a few minutes on each case and not to change their answers. This exercise was designed to produce an easily and efficiently administered--albeit simplistic--feedback measure. The second time members were instructed to deliberate carefully on each case and were encouraged to recheck their choices and make as many changes as desired. The second exercise, in contrast to the first, was designed to assist in the formulation of a carefully considered prospective policy.

Although the research was designed to articulate explicit parole policy, the " 'guideline method' formally began when the Board made a decision to use the initially descriptive data . . . as prescriptive of policy to be followed."[29] Thus, the empirical model of how parole board decisions are made was used in the actual construction of parole guidelines.

The parole guidelines as formulated are displayed in a matrix (see Figure 2-2). Along the vertical axis are six categories of offense severity; on the horizontal axis are four categories of parole prognosis. The customary or policy range for a particular case is found at the intersection of the decision-maker's rating of seriousness of offense and his rating of the offender's chances of success. By setting the exact time prisoners will serve early in their terms and providing a consistent framework within which parole members can make decisions, parole guidelines address both the question of uncertainty of release and of disparity.

The parole guidelines were adopted on a trial basis in October 1972 and were formally incorporated in regulations applicable to the entire federal prison system in 1973. The guidelines, though the subject of litigation, remained in use. During the pilot period, two policies were also adopted: members were required to provide written reasons for departures from the guidelines, and provisions were made for periodic review and revision of the guidelines. The parole guidelines not only stimulated similar work at the state level,[30] but formed the basis of subsequent research in sentencing guidelines.

2.3.2 The Development of Sentencing Guidelines

In an environment of discontent with rehabilitation and criticism of the unchecked powers judges had in sentencing, some of the same researchers who developed parole guidelines saw the potential usefulness of transferring the guidelines concept to sentencing. "The concept of guidelines for sentencing seemed a natural extension of the idea as applied in the parole timesetting decisions," according to Wilkins. In addition to the environment of dissatisfaction with rehabilitation and sentencing disparity, Wilkins cites two factors that contributed to the transition from parole to sentencing guidelines:[31]

- The strong collaborative working relationship between researchers and parole board members in the development of parole guidelines; and

Figure 2-2

Adult Guidelines for Decision Making: Customary Total Time Served before Release
(Including Jail Time)

Offense Characteristics: Severity of Offense Behavior (Examples)	*Offender Characteristics: Parole Prognosis (Salient Factor Score)*			
	Very Good (11-9)	*Good (8-6)*	*Fair (5-4)*	*Poor (3-0)*
Low				
Immigration Law Violations Minor Theft (Includes larceny and simple possession of stolen property less than $1,000) Walkaway	6-10 months	8-12 months	10-14 months	12-16 months
Low Moderate				
Alcohol Law Violations Counterfeit Currency (Passing/Possession less than $1,000) Drugs: Marijuana, Simple Possession (less than $500) Forgery/Fraud (less than $1,000) Income Tax Evasion (less than $10,000) Selective Service Act Violations Theft From Mail (less than $1,000)	8-12 months	12-16 months	16-20 months	20-25 months
Moderate				
Bribery of Public Officials Counterfeit Currency (Passing/Possession $1,000-$19,999) Drugs: Marijuana, Possession With Intent to Distribute/Sale (less than $5,000) "Soft Drugs," Possession with Intent to Distribute/Sale (less than $500) Embezzlement (less than $20,000) Explosives, Possession/Transportation Firearms Act, Possession/Purchase/Sale (single weapon–not sawed-off shotgun or machine gun) Income Tax Evasion ($10,000-$50,000) Interstate Transportation of Stolen/Forged Securities (less than $20,000) Mailing Threatening Communications Misprision of Felony Receiving Stolen Property With Intent to Resell (less than $20,000) Smuggling/Transporting of Aliens Theft/Forgery/Fraud ($1,000-$19,999) Theft of Motor Vehicle (Not Multiple Theft or for Resale)	12-16 months	16-20 months	20-24 months	24-30 months
High				
Burglary or Larceny (Other than Embezzlement) from Bank or Post Office Counterfeit Currency (Passing/Possession $20,000-$100,000) Counterfeiting (Manufacturing) Drugs: Marijuana, Possession With Intent to Distribute/Sale ($5,000 or more) "Soft Drugs," Possession with Intent to Distribute/Sale ($500-$5,000) Embezzlement ($20,000-$100,000) Firearms Act, Possession/Purchase/Sale (sawed-off shotgun[s], machine gun[s], or multiple weapons) Interstate Transportation of Stolen/Forged Securities ($20,000-$100,000) Mann Act (No Force–Commercial Purposes) Organized Vehicle Theft Receiving Stolen Property ($20,000-$100,000) Theft/Forgery/Fraud ($20,000-$100,000)	16-20 months	20-26 months	26-32 months	32-38 months
Very High				
Robbery (Weapon or Threat) Drugs: "Hard Drugs," Possession with Intent to Distribute/Sale (No Prior Conviction for Sale of "Hard Drugs") "Soft Drugs," Possession with Intent to Distribute/Sale (over $5,000) Extortion Mann Act (Force) Sexual Act (Force)	26-36 months	36-45 months	45-55 months	55-65 months
Greatest				
Aggravated Felony (e.g., Robbery, Sexual Act/Aggravated Assault)–Weapon Fired or Personal Injury Aircraft Hijacking Drugs: "Hard Drugs" (Possession with Intent to Distribute/Sale) for Profit (Prior Conviction[s] for Sale of "Hard Drugs") Espionage Explosives (Detonation) Kidnapping Willful Homicide	(Greater than above–however, specific ranges are not given due to the limited number of cases and the extreme variations in severity possible within the category)			

Notes:
1. These guidelines are predicated upon good institutional conduct and program performance.
2. If an offense behavior is not listed, the proper category may be obtained by comparing the severity of the offense behavior with those of similar offense behaviors listed.
3. If an offense behavior can be classified under more than one category, the most serious applicable category is to be used.
4. If an offense behavior involved multiple separate offenses, the severity level may be increased.
5. If a continuance is to be given, allow 30 days (1 month) for release program provision.
6. "Hard Drugs" include heroin, cocaine, morphine or opiate derivatives, and synthetic opiate substitutes.

Source: Don M. Gottfredson, Leslie T. Wilkins, and Peter B. Hoffman, Guidelines for Parole and Sentencing: A Policy Control Method (Lexington, MA: Lexington Books, 1978), pp. 24-26.

- The highly favorable response of the courts to parole guidelines.

Two important differences between parole board decisions and judicial sentencing were not remarked at the time but had a bearing on how well the parole model applied. First, criminal sentencing is a bifurcated decision. The judge must decide whether to incarcerate the offender or put him/her on probation; only then does the time-setting element enter the picture. By contrast, parole decisions focus only on time-setting. Second, the U.S. Parole Board is a single entity, meeting as a body and having its own administrative staff. Parole guidelines development, but especially guidelines implementation, monitoring, and revision, were facilitated by this structure. Criminal courts, on the other hand, are complex, decentralized organizations; it is to be expected that this might pose new challenges to all aspects of a guidelines effort.[32]

Sentencing Guidelines Feasibility Study[33]

In 1974, the National Institute of Law Enforcement and Criminal Justice funded the first of the sentencing guidelines studies by the Albany Criminal Justice Research Center (CJRC) researchers involved in the parole guidelines work. The focus was on the feasibility of guidelines development in state courts. Trial court judges were approached in several jurisdictions to determine their willingness to participate in the development process. Ultimately, four were selected, representing a mix of urban and rural sites. Two were designated active participants and two were designated observer courts.[34]

The cooperative relationship fostered between board members and researchers, instrumental to the development and acceptance of parole guidelines, was considered a key component in any sentencing research to be conducted. Therefore, a Steering and Policy Committee composed of judges from each site was involved in all stages of the project. Researchers hoped that judges would be more receptive to using guidelines if they actively participated in the process of their development.

The research conducted during the sentencing guidelines feasibility study paralleled the parole guidelines research in many ways. The sentencing guidelines took the form of matrices similar to parole guidelines, with severity of offense along the vertical axis, offender scale score categories along the horizontal axis and indicated disposition in the cells located at the intersection of the two axes.

However, parole boards deal only with those offenders who have been incarcerated, whereas judges must decide whether or not to incarcerate an individual and how long the particular sanction chosen must last. Thus, the sentencing guidelines formulated as a result of the feasibility study required much more research than was necessary for the development of parole guidelines. The research tasks in the sentencing guidelines feasibility study involved:

- Determining the variables judges used in the sentencing decision;
- Establishing consistent offense severity ratings;
- Developing objective offender characteristic scales;
- Providing several preliminary guideline models;
- Testing the accuracy of alternative models;
- Synthesizing the alternatives to form one set of experimental guidelines; and
- Testing the synthesized guideline model.

In each site, researchers decided to develop several preliminary alternative guideline models during the feasibility study, because the development of one model was seen as potentially restrictive of the expression of issues and concerns by the Steering Committee. For example, some models used the statutory classification of offenses. Some models did not include juvenile records. Other models included a number of additional items, such as substance abuse, in computing the offender scale score. One model used a single decision matrix for all offenses as opposed to a series --one for each class of offense. The alternative models were presented to the Steering Committee for review. Despite variations in design and construction, all models achieved approximately the same results when tested: 73-84 percent correct prediction of the sentencing decision.

The final demonstration model was developed only for the Denver Court. Development of a similar model in Vermont was impractical because collection of the sample cases for model testing had not yet been completed. Ultimately, the Committee chose to combine several features of the preliminary models for its final guidelines model. In addition, the Committee called for a series of matrices, one for each category of the felony/misdemeanor class system. Each matrix was two-dimensional, with an offense score on the vertical axis (combining statutory seriousness with factors like victim injury) and an offender score on the horizontal axis (combining prior record, legal status at time of offense, employment and school history).

On the basis of their results and acceptance of the model by the Steering Committee, the CJRC researchers concluded that the development of judicial sentencing guidelines was not only feasible but desirable.

> Guidelines . . . provide information to judges which has hitherto been unavailable to those either inside or outside the judiciary. It is, finally, our view that once the judges are accurately informed as to what they have been doing in the past, then they can more clearly focus on what they should do in the future. And, these changes, made by the judges themselves, are much more likely to be accepted and implemented in practice.[35]

Sentencing Guidelines: Pilot Implementation

In July 1976, the Criminal Justice Research Center (CJRC) was awarded a second grant to implement sentencing guidelines in four jurisdictions. The District Court of Denver remained an active participant. The Essex County Court of New Jersey changed its status from "observer" to active participant. Though Vermont withdrew from active participation, it maintained a role in the pilot sentencing guidelines research by becoming an "observer" site. Polk County's status as observer remained unchanged from the feasibility research.

Two new sites were added for guidelines implementation: the Cook County Circuit Court of Chicago, Illinois and the Maricopa County Superior Court of Phoenix, Arizona. Two new "observer" sites were added: the King County Superior Court (Seattle, Washington) and the Philadelphia Court of Common Pleas (Philadelphia, Pennsylvania). The processes of data collection and model construction were repeated in these new sites. Alternative models were developed and presented to judges for selection after making modifications.[36]

Two separate kinds of guideline models were eventually developed for actual implementation. One kind, known as the "general" or "class" model and adopted by Denver and Cook County, was derived from the Denver demonstration model developed during the sentencing guidelines feasibility research. It used one group of information items to calculate the offender's criminal history score, then applied this score to one of several sentencing matrices depending on the statutory class of the crime committed. The second kind, called the "generic" model and adopted by Essex County and Maricopa County, classified crimes by type of offense (e.g., violent, property, drug). For each type of offense a different set of information items was used to calculate the offender's criminal history score.

In November 1976, Denver became the first jurisdiction in the United States to implement sentencing guidelines. Judges decided among themselves, without legislative or administrative order, to adopt guidelines closely resembling those developed during the feasibility study. The two main features of the Denver demonstration model were maintained in the pilot implementation guidelines: the criteria regarding severity of offense and offender characteristics and the use of separate matrices for each felony and misdemeanor classification. Sentencing guidelines were used in Denver for two and a half years--longer than any other jurisdiction--until July 1979, when determinate sentencing legislation went into effect. Guidelines were also voluntarily implemented in Chicago in June 1977, in Newark in January 1978, and in

Phoenix in March 1978.[37] In all these cases, the local guidelines were made obsolete within a year by new sentencing legislation.

As a result of these legislative actions, the pilot study actually provided limited evidence about guidelines implementation. Instead, it represented a more extensive test of the development methodology already explored for parole guidelines and in the initial feasibility study.

Sentencing Guidelines: Evaluation of Pilot Implementation[38]

In early 1978, the CJRC researchers requested funds for full-scale development and implementation of statewide sentencing guidelines. The National Institute, however, decided that an evaluation of the implementation of sentencing guidelines and their impact on sentencing disparity and on court and prosecutorial practices should be conducted before implementing statewide sentencing guidelines.

The evaluation was begun in October 1978 by the National Center for State Courts (NCSC). It addressed three general issues: the empirical basis of sentencing guidelines, the impact of sentencing guidelines on the exercise of judicial discretion, and the impact of sentencing guidelines on the exercise of discretion by prosecution and defense. The examination of the guidelines' empirical basis focused primarily on data collected from the original implementation sites: Denver, Chicago, and Newark. The assessment of the impact of sentencing guidelines on judges, prosecutors, and defense attorneys' discretion was limited to Denver, Philadelphia, and Chicago with primary emphasis on the first two.

The findings of this evaluation were critical of both guidelines development and implementation. In the area of development, NCSC found several flaws in the empirical basis of the guidelines. These flaws concerned:

- the procedure by which variables were selected for inclusion in guidelines;
- the size of the samples used to construct the guidelines; and
- the measurement of sentencing as a single variable, instead of separating the decision to incarcerate from the question of sentence length.[39]

A major conclusion of the NCSC evaluation was that the place of empirical research in the future development of guidelines should be far more modest than that urged by the CJRC researchers. In fact, they pointed out that, "There is no reason why sentencing guidelines could not be wholly the product of conscious choices made by responsible agencies without reference to empirical evidence about past practices."[40]

With respect to guidelines implementation, the NCSC evaluators focused closely on the interplay of judicial and prosecutorial descisions, noting that a reform meant to change the exercise of discretion must be expected to bring changes at other points in the system where discretion is available (especially prosecutor charging).[41] They argue that voluntary guidelines cannot be expected to change the decision process without the support of judge, prosecution, and defense; that support was absent, for varying reasons, in each of the implementation sites.

The NCSC evaluation also found no evidence of guidelines impact on sentence disparity. Judges' sentences were no more likely to fall within the guidelines ranges than before the implementation; racial and sexual disparities, where they existed, were not reduced. Further, judges provided written reasons in very few cases where extra-guidelines sentences were given.

The NCSC conclusion that the guidelines had no appreciable effect on reducing disparity is qualified by two caveats. First, since guidelines had been in effect in Philadelphia for only six months, it is possible that any effects would be delayed. Indeed guidelines may not have their greatest impact until new judges are rotated onto the bench. Second, since guidelines were adopted voluntarily in all sites and no formal sanctions were imposed for failure to comply, they lacked the force and effect of law.[42] The NCSC evaluators believe that the success of guidelines in structuring judicial discretion depends on two factors:

- legislative enforcement of guidelines; and
- inclusion of method of conviction as a determining element.

Methods of conviction must be included in guidelines as a determinative factor, they contend, or the judge's role in sentencing will be diminished by the prosecutor's capacity to negotiate reduced charges.

It should be emphasized that the NCSC evaluation, though highly critical of the guidelines implementation research, did not reject the notion of using sentencing guidelines as a means of structuring judicial discretion. Nor did it assert that guidelines are incapable of reducing disparity. However, it did raise questions about two extremely important assumptions of the prior guidelines efforts: the value and necessity of their empirical basis and their preferred status as voluntary reforms.

Sentencing Guidelines: The Multijurisdictional Field Test

At roughly the same time that NIJ commissioned the CJRC evaluation of pilot implementation, it also initiated the multijurisdictional field test which is the subject of this report. All of the prior research in the area of empirical sentencing guidelines had been conducted in single jurisdictions. As an intermediate step in determining the applicability of guidelines on a statewide level, the Institute believed

a multijurisdictional field test of sentencing guidelines was in order. Unfortunately, the results of the NCSC evaluation were not available at the inception of the field test.

The field test was conceived as a two-part effort: one part would consist of the implementation of guidelines in several jurisdictions within a state; the second part would involve an independent evaluation of the implementation effort itself. The objectives of the multijurisdictional sentencing guidelines field test were set forth as follows:[43]

- To evaluate the effectiveness of sentencing guidelines as a mechanism for enhancing sentencing consistency both within and across different jurisdictions within the same state;

- To test the feasibility of developing and implementing sentencing guidelines in a multijurisdictional setting; and

- To provide a body of knowledge for jurisdictions looking for a means to structure judicial decision-making.

The test plans were developed by representatives of the National Institute of Justice--referred to as the Project Coordinating Team (PCT)--composed of representatives from the Law Enforcement Assistance Administration, the Department of Justice, and staff from Mitre Corporation, a research firm. Use of a Project Coordinating Team was standard procedure at NIJ in the development of test designs. The PCT was responsible for overseeing the development of the testable model, and the design document itself was subject to the Team's final approval. In addition, the design was reviewed by an advisory board served by some of the same researchers involved in the original sentencing guidelines research. At the time the test design was developed, prior research purported to have demonstrated the feasibility of structuring judicial discretion by means of sentencing guidelines. Summarizing the results of the feasibility study conducted in two participating and two "observer" courts, project staff from the Criminal Justice Research Center, State University of New York at Albany (CJRC) concluded:

> This study has clearly demonstrated the feasibility of sentencing guidelines. Such feasibility has been shown on two levels: methodological and practical. On the first level, we have designated specific, weighted and objective items of information which have been able to account for a large percentage of sentencing decisions made in a given jurisdiction . . . [O]n a practical level, [j]udges have been willing to take an active part in this study and have made many valuable contributions to it. Although, as of this date, guidelines have been implemented in only one jurisdiction, we do expect the full cooperation and willingness of other judges to use them when the results of this project become more widely known.[44]

Although these conclusions were drawn prior to the pilot implementation study, they clearly anticipated that guidelines could and would have an impact on judicial sentencing: "Judges have within their capabilities today the means by which they may sharply curtail, if not virtually eradicate, sentencing disparities in most American jurisdictions."[45]

In retrospect, the feasibility finding seems far more sweeping than the early research could support. When the multijurisdictional test design was being prepared, the body of research evidence really concerned only the cooperation of the judiciary in the development of a set of empirical guidelines. Guidelines use had not been examined, and the pilot implementation study was only just reaching the implementation stage (after guidelines development in four jurisdictions). The focus of the multijurisdictional field test design reflects this state of knowledge, in its concentration on guidelines design and development concerns and in its focus on empirical development and voluntary implementation of guidelines as standard features of the experiment. It also reflects the relative lack of evidence about implementation and impact, as a discussion of major test design choices will show.

FOOTNOTES

1. Leonard Orland, Prisons: Houses of Darkness (New York: The Free Press, 1975), p. 31.

2. Arthur M. Gelman, "The Sentencing Hearing: Forgotten Phase of Sentencing Reform," in Discretion and Control, ed. Margaret Evans (California: Sage Publications, Inc., 1978), p. 129.

3. Orland, Prisons, p. 32.

4. Leonard Orland, "From Vengeance to Vengeance: Sentencing Reform and the Demise of Rehabilitation," Hofstra Law Review 7 (Fall 1978): 31.

5. D. Lipton, R. Martinson, and J. Wilks, The Effectiveness of Correctional Treatment: A Survey of Treatment Evaluation Studies (New York: Praeger, 1975).

6. Marvin Frankel, Criminal Sentences: Law Without Order (New York: Hill and Wang, 1973), p. 9.

7. Ibid., p. 88.

8. Willard Gaylin, Partial Justice: A Study of Bias in Sentencing (New York: Vintage Books, 1974); Norval Morris, The Future of Imprisonment (Chicago: University of Chicago Press, 1974); Orland, Prisons; Michael J. Churgin and Dennis E. Curtis, Toward a Just and Effective Sentencing: Agenda for Legislative Reform (New York: Praeger Publishers, 1977).

9. Thorsten Sellin, "The Negro Criminal: A Statistical Note," The Annals of the American Academy of Political and Social Science 140 (November 1928): 52-64.

10. Guy Johnson, "The Negro and Crime," The Annals of the American Academy of Political and Social Science 23 (1941): 271.

11. Henry Allen Bullock, "Significance of the Racial Factor on the Length of Prison Sentences," Journal of Criminal Law, Criminology, and Political Science 52 (1961): 411-417; L. A. Foley and C. E. Rasche, "The Effect of Race on Sentence, Actual Time Served, and Final Disposition of Female Offenders," in Theory and Research in Criminal Justice: Current Perspectives, ed. John A. Conley (Milwaukee: Anderson Publishing Co.,

FOOTNOTES
(continued)

1979); Edwin L. Hall and Albert A. Simks, "Inequality in the Types of Sentences Received by Native Americans and Whites," Criminology 15 (August 1975): 199-221; Joseph C. Howard, "Racial Discrimination in Sentencing," Judicature 59 (October 1975): 121-125; Henry E. Kelly, "A Comparison of Defense Strategy and Race as Influences in Differential Sentencing," Criminology 14 (August 1976): 241-249; S. Mugford and M. Gronfors, "Racial and Class Factors in Sentencing of First Offenders," Australian and New Zealand Journal of Sociology 14 (February 1978): 58-61; R. I. Simon and K. J. Kaplan, "Latitude and Severity of Sentencing Options: Race of the Victim and Decisions of Simulated Jurors," Law and Society Review 7 (Fall 1972): 87-198; T. M. Uhlman, "Racial Justice--Black Judges and Defendants in the Metro City Criminal Court, 1968-1971," (Unpublished dissertation, 1975).

12. Anthony Partridge and William B. Eldridge, The Second Circuit Sentencing Study: A Report to the Judges of the Second Circuit (Washington, D.C.: Federal Judicial Center, August 1974).

13. American Friends Service Committee, Struggle for Justice: A Report on Crime and Punishment in America (New York: Hill and Wang, 1971), p. 9.

14. Frankel, Criminal Sentences, p. 94.

15. Don M. Gottfredson, Leslie T. Wilkins, and Peter B. Hoffman, Guidelines for Parole and Sentencing (Lexington, MA: Lexington Books, 1978), p. 2.

16. See, for example, Henry Allen Bullock, "Significance of the Racial Factor in the Length of Prison Sentences," Journal of Criminal Law, Criminology, and Political Science 52 (1961): 411-417; Leo Carroll and Margaret C. Mondrick, "Racial Bias in the Decision to Grant Parole," Law and Society Review 11 (1976): 93-107; Victor H. Elion and Edwin I. Megaree, "Racial Identity, Length of Incarceration, and Parole Decision Making," Journal of Research in Crime and Delinquency (July 1979): 232-245; Linda A. Foley and Christine E. Rasche, "The Effect of Race on Sentence, Actual Time Served and Final Disposition of Female Offenders," in Theory and Research in Criminal Justice: Current Perspectives ed. John A. Conley (Milwaukee: Anderson Publishing company, 1979), pp. 93-106.

17. Criminal Courts Technical Assistance Project (CCTAP), Overview of State and Local Sentencing Guidelines and Sentencing Research Activity (Washington, D.C.: The American University, 1980).

18. Richard Ku, American Prisons and Jails: Volume IV--Supplemental Report--Case Studies of New Legislation Governing Sentencing and Release (Washington, D.C.: U.S. Department of Justice, National Institute of Justice, 1980).

19. Jim Galvin, Brad Smith, Tanya Broder, Vince Valvano, Leslie Reiber, and David Schaitberger, Setting Prison Terms, Bureau of Justice Statistics Bulletin (Washington, D.C.: Bureau of Justice Statistics, August 1983); Sandra Shane-Dubow, Alice P. Brown, and Erik Olsen, "Sentencing Reform in the United States: History, Content, and Effect" (Madison, Wisconsin: Wisconsin Center for Public Policy, 1983). (Typewritten.)

20. American Bar Association Project on Minimum Standards for Criminal Justice, Standards Relating to Appellate Review of Sentences (New York: Institute of Judicial Administration, 1967-1968).

21. Frankel, Criminal Sentences, p. 84.

22. See Richard F. Sparks, Bridget A. Stecher, Jay S. Albanese, and Peggy L. Shelly, Stumbling Toward Justice: Some Overlooked Research and Policy Questions About Statewide Sentencing Guidelines (Newark: Rutgers University School of Criminal Justice, 1982), p. 41.

23. Joseph C. Howard, Sr., "Sentence Review: The Maryland Experience." (Typewritten.)

24. Frankel, Criminal Sentences, p. 119.

25. 28 U.S.C. §§991-998 (1984).

26. Edward M. Kennedy, "Symposium on Sentencing, Part I: Introduction," Hofstra Law Review 7 (Fall 1978): 5.

27. Minnesota Sentencing Guidelines Commission, Report to the Legislature (St. Paul, Minnesota: Minnesota Sentencing Guidelines Commission, 1980).

28. For further detail about the research conducted to develop parole guidelines, see Don M. Gottfredson, Leslie T. Wilkins and Peter B. Hoffman, Guidelines for Parole and Sentencing: A Policy Control Method (Lexington, MA: Lexington Books, 1978).

29. Leslie T. Wilkins, The Principles of Guidelines for Sentencing: Methodological and Philosophical Issues in Their Development (Washington, D.C.: U.S. Department of Justice, National Institute of Justice, 1981), p. 25.

30. Parole guidelines have been adopted either by law or administrative rule in Alaska, California, Florida, Georgia, Minnesota, Missouri, New Jersey, New York, Oklahoma, Oregon, Pennsylvania, Utah, and Washington. Galvin et al., Setting Prison Terms.

31. Wilkins, The Principles of Guidelines for Sentencing, p. 23.

32. Jolene Galegher and John S. Carroll, "Voluntary Sentencing Guidelines: Prescription for Justice or Patent Medicine?" Law and Human Behavior 7 (1983): 378-391.

33. For more detail on the sentencing guidelines feasibility research, see Leslie T. Wilkins, Jack M. Kress, Don M. Gottfredson, Joseph C. Calpin, and Arthur M. Gelman, Sentencing Guidelines: Structuring Judicial Discretion--Report on the Feasibility Study (Albany: Criminal Justice Research Center, 1976).

34. The District Court of Denver, Colorado and courts of Vermont were designated active participants, in which on-site data collection was conducted. The Essex County Court of New Jersey and the Polk County Court of Des Moines, Iowa were designated "observer" courts. Their role was threefold: (1) to "facilitate the extension of the project to other areas, should it prove successful in the experimental districts"; (2) to serve as arbitrators should the need arise; and (3) to "enrich the information available and serve to keep members of the committees honest by virtue of the fact that they belonged to the same profession." See Wilkins, The Principles of Guidelines for Sentencing, p. 38.

35. Leslie T. Wilkins, Jack M. Kress, Don M. Gottfredson, Joseph C. Calpin, and Arthur M. Gelman, Sentencing Guidelines: Structuring Judicial Discretion (Washington, D.C.: U.S. Department of Justice, National Institute of Law Enforcement and Criminal Justice, 1978), p. 32.

36. Joseph C. Calpin, Jack M. Kress and Arthur M. Gelman, "The Analytical Basis for the Formulation of Sentencing Policy," (unpublished).

37. Jack M. Kress, Prescription For Justice: The Theory and Practice of Sentencing Guidelines (Cambridge, MA: Ballinger Publishing Company, 1980), pp. 119-133.

FOOTNOTES
(continued)

38. For more detail on the evaluation of pilot sentencing guidelines implementation, see William D. Rich, L. Paul Sutton, Todd R. Clear, Michael J. Saks, Sentencing by Mathematics: An Evaluation of the Early Attempts to Develop and Implement Sentencing Guidelines (Williamsburg, VA: National Center for State Courts, 1982).

39. The NCSC evaluators contended that sentences cannot be captured accurately by one matrix using the same variables and weights and predicting both decisions. Indeed, the evaluators found that the determinants of the decision to incarcerate differed from the determinants of length of incarceration. They found that the best predictors of the IN/OUT decision were based on factors related to the offender's criminal history and social stability, including the offender's most serious prior offense, his liberty status, and his employment status. The length of incarceration was best predicted by severity of offense and factors related to plea bargaining, e.g., statutory class of offense, whether the victim was injured, whether a charge reduction was granted, method of conviction, and number of counts on which the offender was convicted.

40. Rich et al., Sentencing Mathematics, p. 14.

41. Ibid., pp. 15, 161-166.

42. Ibid., p. 160.

43. National Institute of Law Enforcement and Criminal Justice, Law Enforcement Assistance Administration, U.S. Department of Justice, Multijurisdictional Sentencing Guidelines Program Test Design (Washington, D.C.: Government Printing Office, 1978), preface.

44. Wilkins et al., Sentencing Guidelines: Structuring Judicial Discretion, pp. 27-28.

45. Ibid., p. xiii.

CHAPTER 3

THE MULTIJURISDICTIONAL FIELD TEST: DEVELOPMENT AND RESULTS

In Chapter 2, we traced the origins of the multijurisdictional field test, which emerged from several years of NIJ-initiated research on sentencing guidelines. In the following sections, we describe the general parameters of the test design, as well as the setting in which it was implemented. We also provide a brief account of the steps taken by project participants in developing and implementing the guidelines and highlight their impact on sentencing behavior in the participating jurisdictions. Taken together, the various sections of Chapter 3 set the stage for the remainder of Part I of this report, which describes the sentencing guidelines design, operation, and system adaptation.

3.1 Test Design Features

3.1.1 Guidelines Design

Since the field test represented an attempt to build upon the previous NIJ-sponsored sentencing studies, the methodology for the development of sentencing guidelines followed the basic approach of the CJRC researchers: that is, guidelines were based on an empirical analysis of past sentencing practices. This restriction of the test to empirical guidelines was probably the most critical factor shaping what happened in Maryland and Florida. The notion of normative or prescriptive guidelines as a distinct alternative had not yet been clearly articulated, and the NCSC study that first criticized the empirical approach had not been completed.

The role of judicial policy-making in empirical guidelines development was certainly acknowledged and even emphasized. However, the CJRC researchers described the statistical analysis and model building inherent in devising empirical guidelines as a way of expressing implicit judicial policy:

> [W]hen repeated decisions are involved, guidelines assume two levels of decisionmaking: the individual level, on which decisions are made one at a time; and the policy level, which represents an aggregation of individual decisions. Thus, while sentencing guidelines are used to structure individual decisions, they are, in fact, a reflection of the aggregate analysis of decisions at the policy level and it is assumed that "at the policy level it is possible to derive an equation to predict decisions on the basis of case information. This predictive ability may be interpreted as a description of latent or implicit policy which in turn provides the basis for the articulation of that policy."[1]

The test design adopted this view.[2]

The test design document offered extensive guidance on the topic of empirical guidelines design and development. The methodology--consisting of data collection, statistical analysis, guidelines model development, and guidelines validation--was described in considerable detail, covering not only the purpose of each task but also specific approaches to carrying it out. For example, a suggested list of core information items (variables) for data collection was provided, with discussion of alternative ways to record the information. Four types of guideline models were described, with details about how to develop them. Two of the four alternatives were recommended for adoption by participating states. The guidelines design choices left to the state projects were limited to the scope and content of data collection, model selection, and structure of the actual guidelines grids.

3.1.2 Structure and Operation of the Field Test

Two aspects of the field design were also delineated in the NIJ test document: the structure of the test (including site selection and project organization) and activities necessary for guidelines implementation. These topics were treated with rather different levels of detail, however. As appropriate to a document setting out specifications for grant eligibility and performance, substantial direction was given concerning the structure of the field test. In contrast to this and to the specificity of instructions about the design of empirical guidelines, however, only brief consideration was given to guidelines implementation issues.

As noted above, the multijurisdictional nature of the test was intrinsic to its purpose; guidelines had only been developed in single courts, and the National Institute believed it important to assess the feasibility of jurisdictional cooperation as an intermediate step to promoting statewide guidelines. Several alternatives for the development of multijurisdictional sentencing guidelines were considered, including:[3]

- The application of guidelines already developed and implemented in one of the earlier research sites to other judicial districts in the same state;

- The development of separate guidelines by several districts within a state, which would then be combined into one set of guidelines and applied to the test jurisdictions;

- The development of separate guidelines by a number of test jurisdictions, based on a common set of factors and weights derived from an initial examination of past practices; and

- The development of a single set of guidelines, based upon cases collected in several districts, by a panel of judges representative of the entire state, which would then be applied to test jurisdictions.

The design finally selected for guidelines development was a modified version of the last alternative. It involved using a combined data base representing all participating jurisdictions in a state as the empirical basis for a common set of guidelines to be applied to each participating jurisdiction. The panel of judges was to be drawn from the test sites only, however, and was not to include statewide representation.

Two to three states were to be chosen, and each state was to select one (and only one) urban jurisdiction, at least one suburban jurisdiction, and--if possible--a rural jurisdiction, as participating sites. Each state was to rely on two groups to collect data on past sentencing practices, develop and implement guidelines, and monitor guideline use: a project staff and an Advisory Board. In addition to the full-time project and research directors, data collectors, coders, computer programmers, and keypunch operators were to be hired for specific phases of the work. The Advisory Boards, composed of judges from participating jurisdictions, were to have responsibility for making policy decisions needed to develop guidelines and for overseeing project staff activities in general. According to the test design document, Advisory Boards might also include a number of ex officio members, such as judges from other jurisdictions, or representatives of prosecutorial, public defender, court administrative, and correctional agencies.

While these structural aspects of the multijurisdictional test received considerable attention in the test design document, only one aspect of guidelines implementation was delineated in this manner: the guidelines experiment was to be judicially initiated and controlled. The guidelines were to be developed and formally mandated by the judiciary through, e.g., an administrative directive of the state court administrator, a state supreme court rule, or a formal agreement among all judges from the participating sites. Commitment by individual judges to use the guidelines involved three actions: considering the recommended sentence and the factors on which it was based, recording the actual sentence, and giving written reasons when the guidelines sentence was inappropriate in a particular case.

By design, then, the multijurisdictional field test was a test of a voluntary guidelines system. Judges were asked not only to accept the recommended sentences but also to consult the guidelines and provide written reasons for any exceptions. However, there were no provisions for sentence review or appeal. While these features had been a fundamental characteristic of the guidelines efforts growing out of NIJ-sponsored research, they are by no means the only alternative. The Minnesota legislative mandate, for example, makes that state's guidelines sentences presumptive (binding except under special circumstances), with review of extra-guidelines sen-

tences required before the sentence becomes final. The consequences of the voluntary nature of the field test guidelines are detailed later in this report.

The test design document offered little other guidance on guidelines implementation. The following list of tasks was provided:

- development of documentation (manuals, scoresheets, etc.);
- assignment of personnel to prepare and collect the scoresheets and analyze the data;
- use of the guidelines;
- analysis of sentencing data with the Advisory Board; and
- periodic review of the guidelines.

Nevertheless, no detailed suggestions were provided for carrying them out. Whereas the prior feasibility study enabled the test designers to be quite explicit about guidelines development, the lack of prior analysis of implementation led the designers to provide only limited guidance in this area. With the advantage of hindsight--from this evaluation and others completed since the field test began[4]--we can identify three important areas of omission:

1. guidance on planning the implementation at the site (rather than state) level;
2. guidance on the potential impacts of differences in local legal culture; and
3. guidance on the participation of criminal justice system actors other than the judiciary.

By and large, the CJRC researchers assumed that the judiciary alone would have sufficient authority to implement sentencing reform. This assumption ran contrary, even at that time, to the growing awareness of the interdependence of the different criminal justice system components and recognition of the homeostatic nature of local court systems. But not until the initial release of a draft of the pilot implementation evaluation in 1980 was there full recognition of the critical role of implementation in sentencing guidelines efforts.[5] Certainly, this was a critical area in which the states were given little guidance.

3.1.3 Other Field Test Components

In addition to outlining certain features of the guidelines and field test design, the test designers included the following two components in their plans:

1. training and technical assistance; and
2. an independent evaluation.

Technical assistance and training to project staff in both states were provided by the University Research Corporation (URC). At the beginning of the field test, URC held a three-day field test training session. While this was intended for all project staff and members of each state's Advisory Boards, Maryland had not yet established its staff at the time the training was held. The training included: information on the theory and operation of sentencing guidelines; an overview of the project organization; the role of the Advisory Board and the National Institute; instruction on penal code analysis; developing coding manuals and coding sheets; and sampling and data analysis.

URC retained several individuals involved in the initial CJRC guidelines development effort to assist sites with this broader test of empirical guidelines. The URC staff and consultants were available to guidelines project staff on a continuing basis, to answer questions that arose and to provide additional technical guidance. URC made periodic visits to the test sites during guideline development and implementation and submitted monitoring reports to the National Institute. However, the technical assistance effort was not intrusive or extensive; sites were free to make their own development choices and in many respects modified the suggestions made by technical assistance staff to meet their own needs. It is for this reason that our discussion of guidelines development, implementation, and impact focuses on the efforts of the sites and excludes, for the most part, any discussion of the technical assistance effort.

As with other NIJ-initiated field tests, an independent evaluation of the multijurisdictional field test was considered an integral part of the test design. Indeed, willingness to support the evaluation effort was one of the criteria states had to meet in order to be eligible for field test participation.

The major objectives of the evaluation, as set forth in the field test design, were as follows:

- to provide a descriptive account of the guidelines development and implementation process for use by other jurisdictions interested in this process;

- to examine the impact of guidelines on sentencing practices and on other components of the criminal justice system;

- to assess the feasibility of developing and implementing sentencing guidelines as a policy tool in a multijurisdictional setting; and

- to test the effectiveness of guidelines as a method for increasing sentencing consistency.

Thus, the evaluation was intended to deal with both process and impact questions.

The major focus of the impact analysis was to be on measuring changes in sentencing uniformity, since the issue of disparity was the primary impetus for sentencing reform. The process analysis was to address two issues: how the multijurisdictional nature of the test affected guidelines development and implementation, and how the introduction of guidelines affected court practice. Although the test design document offered little guidance with respect to implementation issues, it is notable that the test designers specifically mandated examination of possible changes in screening and charging, plea bargaining, judge shopping, or other defense strategies.[6] Thus, they were not unaware of the possible system impacts which might result from the introduction of guidelines.

Following a competitive procurement process, Abt Associates Inc. was awarded a grant to conduct the evaluation. The work began at the same time that guidelines development was initiated in Florida and Maryland. To provide a descriptive account of guidelines development and implementation, direct observation and extensive interviews were conducted with personnel in each site, including judges, prosecutors, defenders, and probation agents responsible for presentence reports. Visits to each of the test jurisdictions were made before and during guidelines implementation, and after the year-long test of guidelines was over. Information for the impact analysis included scoresheet data gathered by the guidelines projects, as well as case records and other data gathered independently by the evaluator.

3.2 The Setting for the Multijurisdictional Field Test

3.2.1 The Site Selection Process

The site selection process involved two levels of decisionmaking: selection of participating states and selection of local jurisdictions within states. Each of these is discussed below.

In order to guard against possible legislative changes in sentencing practices, such as the introduction of determinate sentencing, which might render the guidelines test inoperative, NIJ decided to conduct the test in more than one state. Additionally,

a key criterion for participation in the field test was the absence of current or pending legislation that would limit judicial discretion or obviate the field test. Other criteria for selecting sites for participation in the field test included:

- participation of no less than three and no more than five jurisdictions within the state, only one of which would be urban;

- indication of interest, cooperation, and commitment on the part of judges and other relevant court personnel;

- commitment that the implementation of guidelines during the test be formally mandated;

- combined total annual caseflow among the participating jurisdictions of at least 4,000 cases;

- availability and retrievability of case information from each participating court recording information available to judges at sentencing;

- existence of ongoing, organized efforts in that state dealing with the analysis of sentencing practices, assessing and/or recommending sentencing reforms; and

- commitment to cooperate with evaluation of the development and implementation of guidelines in the field test.[7]

Four states--Florida, Maryland, Pennsylvania, and Wisconsin--expressed an interest in participating in the field test, and at least eight others were contacted to determine the level of interest in the project (Alaska, Arizona, Colorado, Connecticut, Georgia, Illinois, Minnesota, and Ohio). After reviewing sentencing-related legislation in these states, and, in some cases interviewing criminal justice personnel, sponsors of the field test determined that only Florida and Maryland met all selection criteria. Florida was considered a particularly strong candidate because of work completed by its Sentencing Study Committee. Founded to recommend sentencing reform strategies to the state legislature (which had been considering mandatory sentencing), the Florida Sentencing Study Committee had collected and analyzed 1,000 sentencing decisions, established a body of personnel experienced in data collection and analysis, and gathered empirical evidence on sentencing variations in the state. The Maryland judiciary had also demonstrated considerable initiative in exploring sentencing reform options. In 1978, the Chief Judge of the Maryland Court of Appeals established a Sentencing Study Committee and gave it the responsibility to investigate sentencing issues and reforms and to present recommendations for change to the Judicial Conference.

Each state chose to implement the test in four local jurisdictions. In Florida, the judicial circuit was chosen as the participating unit, and participating sites were: the 4th Circuit (urban), the 10th Circuit (rural), the 14th Circuit (rural), and the 15th Circuit (suburban). It is important to note, however, that most judicial circuits in Florida are composed of two or more counties, with a circuit court based in each. In the four Florida test sites, the characteristics of these counties vary substantially within circuits. Thus, for example, the "urban" test site includes three counties, only one of which truly constitutes an urban jurisdiction. Similarly, the two rural circuits both contain counties that might be called suburban by the informed observer. Nevertheless, on the whole, these four jurisdictions were deemed to conform to the general urban, suburban, and rural classification required in the test.[8] Table 3-1 summarizes the characteristics of the test sites along such dimensions as number of counties, per capita population, population of the largest city, and criminal caseload. Although not part of the test design requirements, Florida's choice of sites was also based on "the desire to have a geographic distribution reflective of the varying social and political attitudes within the state."[9] The location of each site is shown in Figure 3-1. Certainly the sites chosen represented a diverse mix, encompassing the north Florida region which is similar in many respects to rural Georgia or Alabama, the mid-state region typified by rural citrus farming and mining operations, the wealthy resort areas of Florida's south Atlantic coast, and the relatively industrial urban areas of the state. Nevertheless, with only four jurisdictions participating in the test, some jurisdictions with particular crime problems or geographic concerns could not be represented. Moreover, the final site selection did include an urban jurisdiction that is not representative of many others in Florida. Unlike the urban areas of southern Florida, such as Dade County, Jacksonville has a reputation (much deserved, as we shall see later) for conservative attitudes and severe sentencing of criminal defendants.

Maryland also implemented the guidelines in four test sites: Harford County (rural), Montgomery County (suburban), Prince Georges County (suburban), and Baltimore City (urban). The choice of jurisdictions differed from Florida's in two important respects. First, the participating sites were counties (or, in the case of Baltimore, a city) rather than circuits. As such, each jurisdiction included only one participating court, serving one relatively homogeneous community, rather than several courts located in distinct and dissimilar communities. Second, Maryland chose a different mix of jurisdictions: one rural and two suburban sites, in addition to the requisite urban site. The characteristics of the Maryland sites are presented in Table 3-2; their location is shown in Figure 3-2.

As in Florida, the choice of the urban jurisdiction raises some interesting concerns. Baltimore is the seventh largest city in the United States and the only major metropolitan area in Maryland. The pressures faced in this court--especially with regard to case volume--are unmatched anywhere in the state. This contrast is particularly evident when one compares the caseload of Baltimore with that of the

Table 3-1

Florida Test Site Characteristics

	4th Circuit	15th Circuit	14th Circuit	10th Circuit
Guidelines Project Designation	Urban	Suburban	Rural	Rural
Total population[a]	670,949	576,863	186,078	388,557
Total felony dispositions[b]	3,095	3,059	1,489	439
Population of largest municipality[a]	Jacksonville 540,920	West Palm Beach 63,305	Panama City 82,200	Lakeland 47,406
Number of counties	3	1	6	3
Number of municipalities over 50,000[a]	1	1	1	0

[a] Based on data presented in U.S. Department of Commerce, Bureau of the Census, 1980 Census of Population--General Population Characteristics: Florida, Vol. PC80-1-B11 (Washington, D.C.: U.S. Government Printing Office, 1982).

[b] Based on 1981 court data presented in the Florida Judicial System Statistical and Program Activity Report, 1980 and 1981 (Tallahassee, FL: The Office of the State Courts Administrator, 1982). We present 1981 data since 1980 data were not available for the 4th Judicial Circuit. These figures include pleas and convictions, and exclude adjudication withheld, dismissals, transfers, and deferred prosecutions.

Figure 3–1
Florida Circuits Participating in the Multijurisdictional Sentencing Guidelines Field Test

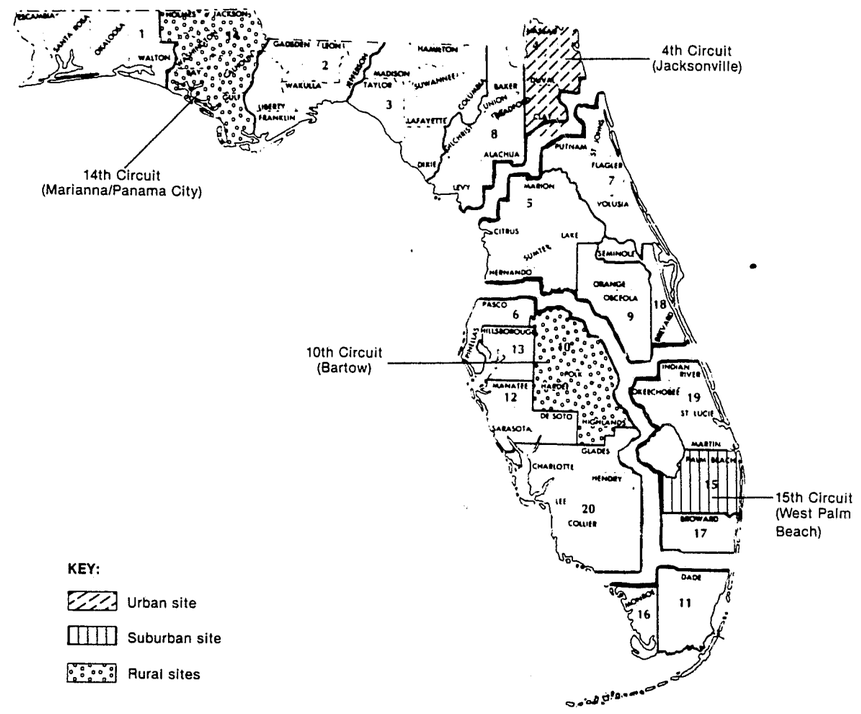


Table 3-2

Maryland Test Site Characteristics

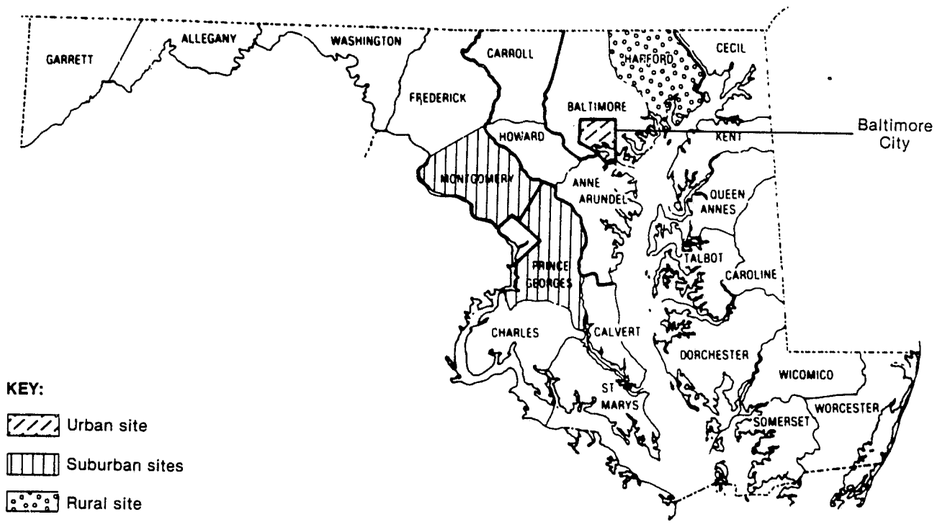
	Baltimore City	Prince George's County	Montgomery County	Harford County
Guidelines Project Designation	Urban	Suburban	Suburban	Rural
Total population[a]	786,775	665,071	579,053	145,930
Total circuit court filings[b]	22,200[c]	3,303	1,974	811
Population of largest municipality[a]	Balt. City 786,775	Oxon Hill 36,267	Bethesda 62,736	Edgewood 19,455
Number of municipalities over 50,000[a]	1	0	1	0

a Based on data presented in U.S. Department of Commerce, Bureau of the Census, 1980 Census of Population--General Population Characteristics: Maryland, Vol. PC80-1-B22 (Washington, D.C.: U.S. Government Printing Office, 1982).

b Data is for the period July 1, 1980 to June 30, 1981. Figures obtained from the Annual Report of the Maryland Judiciary: Statistical Abstract 1980-1981 (Annapolis, MD: Administrative Office of the Courts).

c The Annual Report notes that in Baltimore City, "one case represents one charge (count) rather than one incident. An audit conducted in 1980 found that by using charge statistics Baltimore City reported 2.19 times the number of filings . . . as would have been reported under a system comparable to other counties." Annual Report 1980-1981, p. 48.

Figure 3—2
Maryland Counties Participating in the Multijurisdictional Sentencing Guidelines Field Test



KEY:

Urban site

Suburban sites

Rural site

participating rural jurisdiction. Baltimore felony cases account for over 33 percent of the total state felonies, while the comparable figure for Harford County is only 3 percent.[10] As described below, this diversity posed some difficulties during the guidelines development process.

One additional consideration noted as a site selection criterion in the test design document was the commitment of each site to participate in the test. In both states it seems likely that the activities of the sentencing study committees planted the seeds of cooperation. In Florida, three of the four participating circuits (the Fourth, Tenth, and Fifteenth) were represented on the Sentencing Study Committee sponsored by the Florida Supreme Court. Similarly, judges from Baltimore City and Prince George's County served on the Maryland Study Committee, and their interest in and support of the Committee's recommendations were no doubt a force in enlisting cooperation. In addition, some degree of commitment to the test was evidenced in the grant application from the states. For example, the Florida grant application included a letter of commitment from the chief judge of each participating circuit, while in Maryland the administrative judge in each county submitted a letter indicating that judges in that county had reviewed the guidelines test design and voted to participate in the test.

In spite of this early support, it is clear that some commitments to participate in the test were hard won and that many judges understandably retained some reservations about the guidelines project. For example, in interviews conducted by our staff before the guidelines implementation, many members of the judiciary voiced little or no support for sentencing guidelines as a reform measure, and they questioned the feasibility and even the desirability of eliminating regional differences in sentencing. In addition, it was clear from interviews with public defenders and prosecutors that little information about the test had been provided to these groups, and that their support had <u>not</u> generally been sought when selecting sites for participation. It is perhaps not surprising, therefore, that defense and prosecution resisted the guidelines in some of the test jurisdictions, as will be discussed in Chapters 5 and 6.

3.2.2 The Statutory Environment

The two states participating in the multijurisdictional field test offer some remarkable contrasts in terms of the statutory environment for guidelines. Florida's criminal code is based largely on the American Bar Association's Minimum Standards relating to Sentencing Alternatives and Procedures and the Model Penal Code sentencing provisions. It uses a highly structured degree system to delineate offenses by maximum punishments. At the time of the study, Maryland criminal law included a mix of statutorily defined crimes, common law crimes, and statutorily prescribed penalties for common law and statutory crimes. These different systems had substantially different implications for the development of sentencing guidelines.

Florida statutes establish two major categories of criminal behavior: felonies, which are punishable by death or imprisonment in a state penitentiary (i.e., the sentence exceeds one year) and misdemeanors, which are punishable by sentences of less than one year in a county correctional facility. Within these broad categories, Florida statutes classify both felonies and misdemeanors by degree and specify maximum penalties for each. (See Table 3-3.)

With a constitutional amendment effective in 1973, Florida authorized three distinct court levels: the appellate courts, the circuit courts, and the county courts. Exclusive jurisdiction over felony cases was granted to the circuit courts, while county courts were granted original jurisdiction for all misdemeanor offenses. Given the test design's mandate that participating courts have original jurisdiction over felony offenses, the Florida guidelines project focused exclusively on the circuit courts, as noted above.

Within the general penalty guidelines established by statute, Florida's circuit court judges have extremely broad discretion. Except in capital cases and certain specific offenses, the judge may impose any sentence within the prescribed statutory limits. However, mandatory minimum sentences have been established for three kinds of offenses: capital felonies carry a mandatory minimum 25-year term, possession of a firearm while committing or attempting to commit selected felonies[11] carries a three-year minimum sentence, and minimum sentences and mandatory fines have also been established for drug trafficking. In addition, Florida has an habitual offender law that mandates enhanced penalties for certain repeat offenders. (See Appendix B for more detail.)

Finally, Florida statutes and rules of criminal procedure establish several other provisions which may affect felony sentencing. For example, in certain serious offenses, judges are allowed to retain jurisdiction over the offender for the first third of the maximum sentence imposed. This allows judges some measure of control over sentenced offenders, in that they may review[12] (and if necessary, vacate[13]) Parole Commission release orders occurring during that period. Florida statutes also provide some guidance concerning sentencing for multiple offenses. If the defendant is convicted of two or more distinct offenses, the court may order that the sentences for each be served consecutively or concurrently.[14] However, if the charging document contains several counts, each of which is part of the same transaction, only one sentence can be imposed.[15]

Maryland criminal law is contained primarily in Article 27 of the Maryland Code--an Article which has not been subjected to general revisions since its enactment in 1809, although various legislative modifications have been made over the years.[16] Article 27 provides no systematic distinction between felony and misdemeanor offenses in Maryland. Rather, Article 27 defines the elements of each crime, sets out the penalty, and generally (although not always) indicates the status of the crime as a misdemeanor or felony. Maximum penalties for most crimes are set forth

Table 3-3

Florida Statutory Classification

	Maximum Prison	Minimum Prison	Maximum Fines	Subsequent Felonies
FELONIES				
Capital	Death/ Life	25 years before parole eligibility		
Life	Life	30 years	$15,000	
First degree	30 years[a]		$10,000	Life
Second degree	15 years		$10,000	max. 30 yrs.
Third degree	5 years[b]		$ 5,000	max. 10 yrs.
MISDEMEANORS				
First degree	1 year		$ 1,000	max. 3 yrs.
Second degree	60 days		$ 500	max. 1 yr.

a Where specified by statute, imprisonment for a term of years not exceeding life.

b Includes felonies where degree is not specified, unless punishable by life imprisonment for the first offense.

in the statute, and minimum as well as maximum penalties have been established for over two dozen felony offenses. In addition to the specific criminal offenses defined in the statute, Maryland law provides for mandatory minimum sentences without parole for certain types of repeat offenders.

Although most criminal acts are defined by statute, criminal behavior in Maryland can also be charged under common law. Maximum penalties for some, but not all, of these offenses are included in Article 27. In 1981, the Maryland Rules were amended to prohibit charging under common law crimes; however, use of common law charges persisted throughout the period of guidelines implementation. (See Appendix C for further information.)

Maryland judges also have broad discretion in determining the actual sentence to be imposed. Generally, the judge may impose any sentence authorized by statute. Sentences for common law offenses with no punishment prescribed by statute are within the discretion of the court, subject to the limitation of "cruel and unusual punishment" in the Maryland Declaration of Rights.[17] However, there are some limitations in sentencing. For example, judges of the circuit court may not generally impose probation terms in excess of five years. In addition, incarcerative sentences of three years or more must be explained in writing by the judge (though this requirement was later waived for jurisdictions participating in the guidelines test). Finally, Maryland case law indicates that it is improper to impose separate sentences when the offenses charged are derived from the same transaction; however, separate and distinct crimes charged on the same indictment or on separate indictments may be sentenced separately.[18]

Maryland has two levels of appellate courts (the Court of Appeals and the Court of Special Appeals) and two levels of trial courts (circuit and district). Unlike Florida, jurisdiction of the Maryland circuit and district courts overlaps in some instances. While the circuit courts have jurisdiction over most criminal cases, the district courts have concurrent jurisdiction to try felony cases involving breaking and entering, and larceny where the value of the stolen property does not exceed $500.[19] Since the preponderance of felony-level cases is heard in the circuit courts, the Maryland Sentencing Guidelines Project chose to develop guidelines only for cases originating in the circuit court during the test period.

Maryland statutes, like those in Florida, establish other procedures and requirements that may affect sentencing. For example:

- Maryland judges retain jurisdiction over any sentence for a period of 90 days, and during that period they may reduce, modify, or strike the sentence. This authority has been used, in some instances, to deliver "shock sentences" in which the judge would impose a severe penalty, knowing that he would later be able to modify the sentence.

- Maryland statutes provide that the defense may request that any sentence over two years imposed by the circuit courts be reviewed by a panel of three judges appointed from the circuit in which the case was originally tried. The general mandate of the panel is to re-evaluate relevant information and to assess the appropriateness of the punishment imposed. The panel may decrease, increase, or affirm the sentence originally imposed. In practice, there is some evidence that modifications of sentence by review panels occur rarely, and that when such modifications are granted, reductions in sentence far outnumber sentence increases.[20]

3.2.3 Project Structure

Advisory Board Membership

The test design stipulated that an Advisory Board representing judges from the participating districts be established for each test site. The Board was to oversee the project staff and to make policy decisions related to the development of the guidelines. Both Florida and Maryland complied with this requirement; however, the composition of the two Boards differed, as discussed below.

In Florida, the Chief Judges of the four participating circuits were chosen as the four voting members of the Board.[21] The Board was chaired by a Justice of the Florida Supreme Court, who voted only in the case of a tie. In addition, eight ex officio members were appointed, including a state attorney, a public defender, a parole commissioner, a representative of the circuit judges' conference, a representative of the private bar, a state senator, a state representative, and a criminologist. Five of the original 12 Florida Board members had served on the Sentencing Study Commission; the Board also included a member knowledgeable about research design and methodology.

Maryland followed a different approach in appointing its Advisory Board, choosing a total of 10 judges to represent the four test jurisdictions. Representation was based on the size of the jurisdiction. Thus, Harford County had one judicial representative, Baltimore had five, and the two suburban jurisdictions had two representatives each. As in Florida, the judges chose to invite ex officio members representing the Department of Public Safety and Correctional Services, the Division of Parole and Probation, the State's Attorney Coordinator, the Public Defender, the Legislative Officer of the Governor's office, the Maryland House of Delegates, and the Maryland State Senate. In addition, the Chairman of the Maryland Parole Commission was later asked to join, and attended his first meeting in October 1980.

As will be discussed in subsequent chapters, the size and structure of these advisory groups had an important impact on their effectiveness in overseeing staff, developing guidelines policy, and building support. The Boards offer some interesting contrasts in their design and implementation roles.

Staff

The second major component of the multijurisdictional sentencing guidelines project organization was a central staff responsible for the technical development of the guidelines and general project administration. In Florida, key staff of the Sentencing Guidelines Project were drawn from staff of the Sentencing Study Committee. Clearly, this facilitated a quick project start-up and ensured the availability of experienced personnel who were familiar with data collection and empirical analysis, as well as the Florida court system.

Staff for the Florida project began work in the fall of 1979 with the award of the grant to the Florida State Court Administrator's Office. Activities of the guidelines staff were divided among the following positions:

- Project Director - Responsible for the organization and administration of the guidelines project; work directly with the sentencing commission/advisory committee in making policy decisions; and serve as liaison with the state attorneys and public defenders in each jurisdiction as well as the executive and legislative staffs interested in the progress of the study.

- Project Coordinator - Responsible for developing the methodology necessary to support the committee's recommendations and policies; oversee the data collectors in the field; conduct any ancillary research (technical or theoretical) required by the committee; serve as liaison with the Department of Corrections; and conduct the training required of and supervise the personnel assigned to complete the guidelines worksheets.

- Research Director - Responsible for developing the methodology necessary for the technical development of guidelines; perform the statistical analysis and model-building activities required; and supervise the data collectors, coders, key punchers and computer programmer.[22]

In addition to these central positions, the project hired other staff to supervise data collection, collect data in the participating jurisdictions, and code and enter data for computer analysis.

Since the Maryland project staff were recruited and hired after the initial grant award, this project experienced delays during its early months. By January 1980, however, both the project director and research director had been hired; a research assistant was added to the staff later that year. In general, staff responsibilities were similar to those of the Florida project.

In both states, the highest court bears the responsibility for policy and administration of the judicial system, and a central administrative office has been established by those courts to assist in these duties.[23] The project staffs' placement in the central administrative office in both states complied with the test design's recommendation that the project director be "available and, if possible, assigned from the agency responsible for the superintendence and/or administration of sentencing policy in the state."[24] On the one hand, this placement certainly lent the projects a degree of legitimacy that might not otherwise have been available. On the other hand, the emphasis of these central offices was clearly administrative, not judicial. Combined with the fact that the test design stressed the analytic component of guidelines development, it was only natural that the staff backgrounds would reflect project management and research rather than the law and the politics of the courtroom. This staffing pattern did little to enhance judicial perceptions that they were actively involved in the guidelines development process, and it placed an even greater burden on the Advisory Boards to enlist judges' support.

3.3 Guidelines Development and Implementation

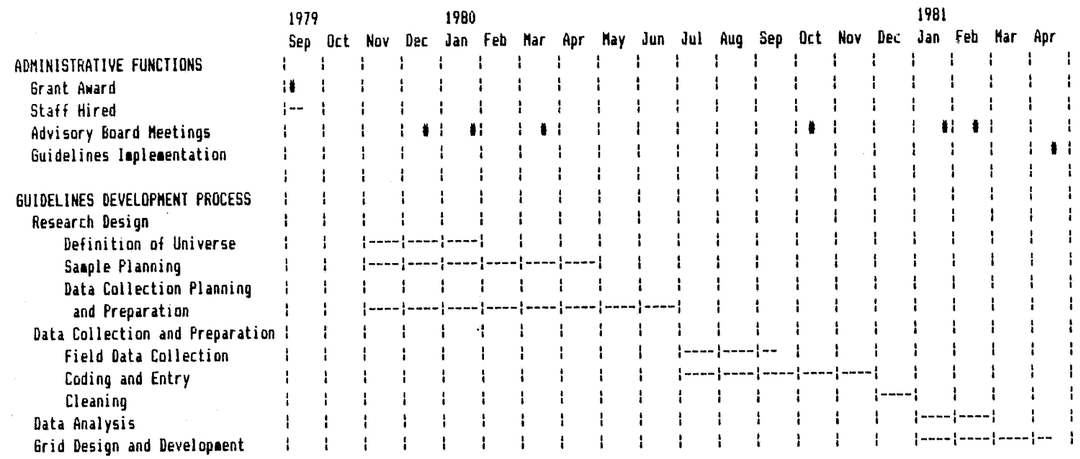
In both Florida and Maryland, the development of the sentencing guidelines was a process of interaction between the Advisory Boards and project staffs. Below, we summarize the steps that were taken to develop guidelines in each state and highlight significant facets of the implementation process. In subsequent chapters, we amplify our discussion of the development process, pointing out the many choices it entailed. We also elaborate on the process of implementation and describe its impact on the court system.

3.3.1 Florida

As we have noted, the project staff was already in place and functioning when Florida made the transition from its earlier Sentencing Study Committee grant to the multijurisdictional guidelines test in September 1979. Figure 3-3 shows how project activities concerning administration and guidelines development unfolded in Florida, from that point in time through the beginning of guidelines use. Even before the Advisory Board was formed, staff began many of the tasks involved in designing the empirical research upon which the guidelines were to be based. In particular, they were able to begin defining the sampling frame (i.e. the set of cases from which a sample would be drawn) and drafting the codebook and data collection instrument (which established the information items to be gathered and the form in which they would be collected).

Figure 3-3

Time Line for the Development of Multijurisdictional Sentencing Guidelines in Florida



Due to this early start, the staff were ready to present several items to the Advisory Board at its first working meeting in January 1980. The staff recommended and the board approved an initial strategy of offense-based sampling, to allow crime-specific analysis of prior sentencing patterns. (This strategy was meant to give the project maximum flexibility, since the data could also be aggregated into groups of offenses for purposes of analysis.) The Board also reviewed and recommended small changes in the draft codebook and instrument. Work proceeded on identification of the sampling frame and sample design; the Advisory Board reviewed and approved the plan at its second meeting in March. By the end of April, the sample was drawn. After pilot-testing and final revision of the codebook and instruments, field data collection began on July 1, 1980.

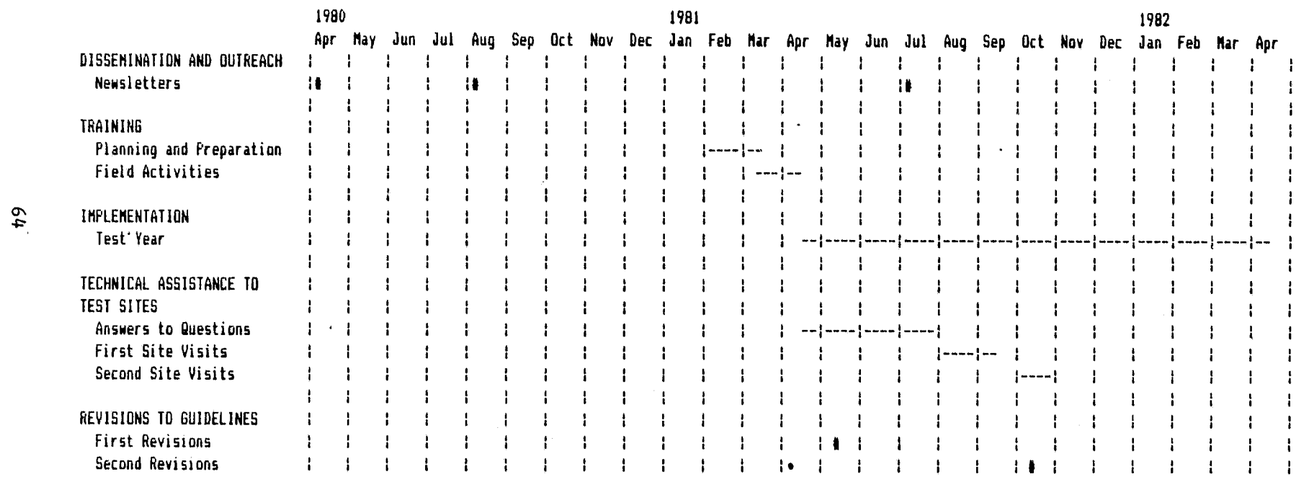
By the time the Florida Advisory Board held its third working meeting in October 1980, a great deal of preliminary work had been completed. Field work on the construction sample (i.e., the sample used to construct the guidelines) was over and coding nearly done. The staff were able to present some hypothetical grids and to begin the discussion of how to shape the actual guidelines. The construction sample data did not become available for analysis, however, until the end of December, three weeks before the next scheduled Board meeting (January 20-21, 1981). At that time, staff presented a preliminary analysis of one group of offenses; less than a month thereafter, they presented a full set of draft guidelines to the Board. Based in part on Board decisions in that February 1981 meeting, the guidelines were finalized and readied for training in late March and implementation in mid-April. Thus, the development process took 20 months in Florida, with the empirical analysis and grid development occupying the final three.

Figure 3-4 gives an overview of the Florida project's schedule for activities related to guidelines implementation. Little was done by way of dissemination or outreach prior to the beginning of guidelines use. Although staff made some attempt to inform site participants about the guidelines project during the development phase, extensive on-site efforts were not made prior to guidelines training. The staff did make use of the Judicial Forum, a general newsletter published by the State Court Administrator's Office, to publicize the project. The earliest article appeared in April 1980, announcing receipt of the grant. In the summer of 1981, preliminary results of the data collection process were featured. Staff and Advisory Board members also gave interviews to the press and published articles describing the project in professional journals. In spite of these efforts, however, participants were generally uninformed prior to training.

Preparation for training began after the guidelines models were finalized and roughly one month prior to the implementation date. At that time, guidelines documentation was prepared and a user's manual developed for distribution to judges, attorneys, and probation personnel. Drafts of the manual were used during the training to give potential users a chance to review the procedures and identify ambiguous instructions for clarification.

Figure 3-4

Time Line for the Implementation of Multijurisdictional Sentencing Guidelines in Florida



During March and early April 1981, twelve training seminars were conducted, three per jurisdiction. Separate sessions were held for three major groups--judges, attorneys, and probation personnel--in order to facilitate free exchange of information and comments and to allow project staff to review the comments and concerns of each group.

On April 1, 1981, the Florida Supreme Court signed an administrative order formally implementing the guidelines. The order set forth April 15th as the effective implementation date, with the duration of the test to be one year. The order also delineated which individuals should be sentenced under the guidelines and included a directive that judges in the participating circuits consult the guidelines in making sentencing decisions.

Implementation began on schedule in two of the circuits but was delayed in two others. The major cause of the delay in these latter sites was the unavailability of presentence investigation reports for offenders convicted as of the April 15th start-up date. However, by the summer of 1981, a majority of the offenders were being processed under the guidelines in all four sites.

Project staff were barraged with technical questions during the first few weeks of implementation, after which the questions tapered off. No on-site visits were scheduled during the period April 1 through July 1981, to allow each site to adapt to the guidelines in its own manner. In August and September, project staff conducted visits to all four test jurisdictions to speak with state attorneys, defense counsel, and probation personnel and solicit suggestions regarding guidelines modification. In October 1981, the Chief Justice accompanied project staff on a more formal visit to each site. This round of visits was viewed as "missionary work," and was intended to encourage guidelines use. Comments and criticisms were elicited, and the suggestive/non-prescriptive nature of the guidelines was stressed.

Two sets of revisions were made to the guidelines during the test year, based primarily on comments and concerns expressed by judges and other members of the criminal justice community. The first occurred in May 1981, approximately one month after the formal implementation of guidelines. At this time, judges on the Advisory Board modified the guidelines grid by collapsing the two lowest sentencing ranges. They also made a major procedural change, allowing attorneys to complete the scoresheets without having probation personnel verify guidelines scores in plea-negotiated cases. In October 1981, the full Advisory Board met to discuss major substantive revisions to the guidelines. Two significant changes were made by the Board: greater emphasis on the current offense in guidelines scoring, and the addition of extent of victim injury as a factor in one category of offenses.

The Florida guidelines continued in operation for the entire test year. In January 1982, the Florida Supreme Court submitted its report to the legislature on the experience with the multijurisdictional test. In this report, it was recommended that

the state proceed with the development of guidelines on a statewide basis.[25] Legislation was introduced later that year authorizing the judiciary to proceed with the development of guidelines, using an approach very similar to that employed during the multijurisdictional field test. Passed by the legislature, the bill was signed into law by the Governor on April 7, 1982. Under that mandate, Florida established a new statewide Sentencing Guidelines Commission and developed guidelines for the entire state of Florida. These were approved by the legislature and were implemented on October 1, 1983. Though empirically based, these guidelines incorporated other features very different from the original multijurisdictional guidelines: legislative enactment, appellate review of extra-guidelines sentences; and abolition of parole release for offenders sentenced under the guidelines.

3.3.2 Maryland

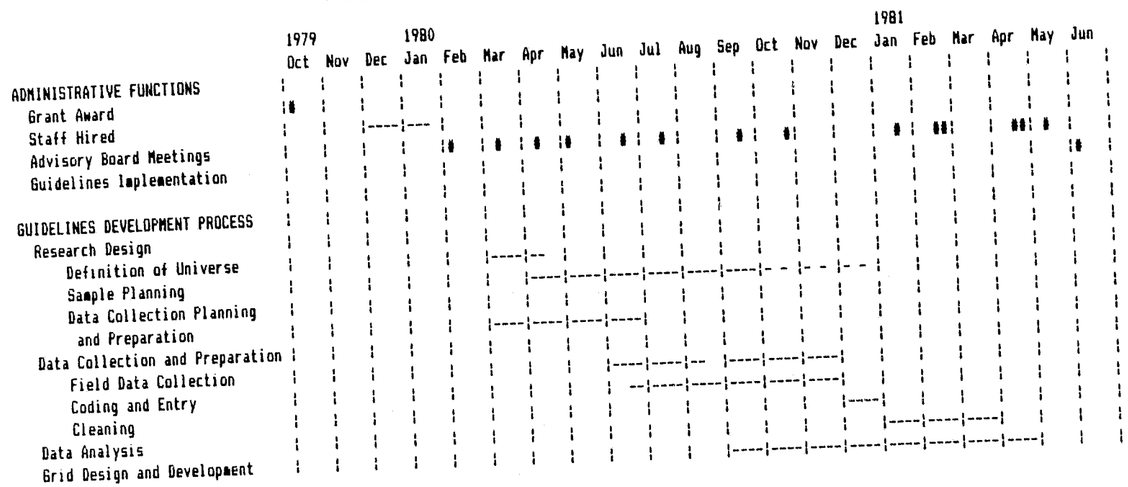
The delays in staff hiring for the Maryland project meant a slower start, except for the Advisory Board formation that built upon the earlier Committee on Sentencing. (Figure 3-5 shows the project's administrative and development activities in Maryland from grant award through guidelines implementation on June 1, 1981.) In January and February, 1980, analysis of the Maryland criminal code revealed a notable lack of structure; thus, from its first meeting in February 1980, the Board's attention was focused on defining the universe of guidelines offenses and creating a framework for classifying offenses. Several important decisions were made by the Board in April 1980: to include only cases originating in the circuit courts,[26] to sample by count for calendar year 1979, and to try to develop crime-specific guidelines models (if there were sufficient data), but to use generic ones otherwise. The three-way grouping of offenses that was later used in grid construction--property, drugs, and violent offenses--was first mentioned in this context.

In the meantime, the Board reviewed a first draft of the data collection instrument in March, made further revisions in April, and finalized the instrument in May (after pilot-testing by the judges). Data collection began in late June, slightly in advance of Florida's schedule. However, substantial problems had arisen in developing a sampling frame. Data collection began before the sample was defined, and it was rearranged, halted for a period, and generally delayed by lack of closure on sampling issues. As a result of this and other difficulties, the final data were not available until January, and analysis really began late that month.[27]

The Maryland staff first presented the Board with sample guidelines grids and scoring in September 1980, and a series of decisions about format and offender factors was made in advance of the analytic results. When these results proved "inconclusive" and "disappointing," the Board agreed to a staff recommendation that two other sources of data be brought to bear on the problem of defining the sentence ranges. On the basis of the 1979 construction sample, two sets of sentencing simulations, and a good deal of collective judgment, the Advisory Board made its final decisions on the

Figure 3-5

Time Line for the Development of Multijurisdictional Sentencing Guidelines in Maryland



grids at the end of April. The development process thus spanned 18 months in Maryland, with seven devoted to grid development and four to empirical analysis.

The time line for implementation activities in Maryland is shown in Figure 3-6. At the same time guidelines development was being carried out, Maryland initiated dissemination and outreach activities. At the very first Advisory Board meeting, in February 1980, judges discussed ways of involving other key actors in the system, including the legislature, the Parole Commission, the Governor's office, and other members of the judiciary. During the period of June to December 1980, the project newsletter provided information on the project and its Advisory Board, described ongoing data collection and analysis activities, and discussed training and implementation plans. Throughout the life of the project, project staff and Board members made presentations to state and community groups and published articles in relevant professional journals.

Training for guidelines implementation began early in Maryland. In September 1980, staff presented the Advisory Board with preliminary plans for training judges and other personnel, such as defense counsel, prosecutors, and probation staff. Early in 1981, these plans were revised, and a subcommittee of the Advisory Board was appointed to develop a strategy for judicial training.

Training materials were completed in early May 1981. Judicial training for circuit court judges in all four jurisdictions was conducted in a single session on May 21, 1981. Training for other court personnel occurred in May, June, and July of that year. While Board members helped conduct the judges' training program, project staff conducted the other training sessions which were held at each site.

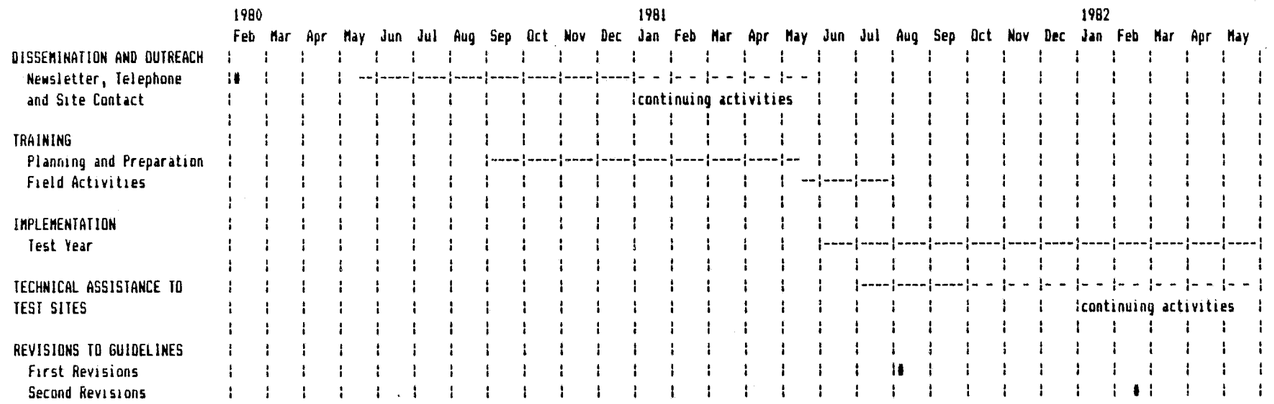
Formal implementation of the guidelines began on June 1, 1981, although actual application of the guidelines was uneven during the first several weeks of operation. Unlike Florida, where the guidelines were implemented by order of the highest court, the Administrative Judge in each jurisdiction was given the responsibility to take any measures deemed necessary to implement the guidelines. The period of the test was one year.

Staff expected to provide continuing technical assistance to each jurisdiction as needed. During the first few months of implementation, errors were found in a large number of worksheets, and follow-up phone calls were initiated to resolve identified problems. During July through September 1981, the "unacceptably high" error rate in completing scoresheets led to additional presentations to many criminal justice groups, including state attorneys, public defenders, members of the private bar, members of the Parole Commission, county probation agents, secretaries, and law clerks.

During the test year, the Advisory Board continued to meet regularly to monitor implementation and discuss possible modifications to the guidelines. The first

Figure 3-6

Time Line for the Implementation of Multijurisdictional Sentencing Guidelines in Maryland



modification occurred in August 1981, with the elimination of employment record as a factor in scoring and one procedural change--the addition of space on the scoresheet for certain judicial recommendations and signatures. These changes went into effect for convictions after September 1, 1981.

The second set of revisions occurred in February 1982. The severity of the offense score for rape offenses was increased and the consecutive/concurrent sentencing policy was clarified. These revisions became effective as of March 17, 1982.

At the conclusion of the test year, the Judicial Conference of Maryland was to decide on the issue of statewide implementation. Although the experience of the four test sites was generally deemed to be favorable, it was decided to continue for one more year on a test basis in the same four jurisdictions, to increase the knowledge and experience with guidelines before statewide implementation. Accordingly, during June and July of 1982, the Advisory Board held three day-long workshops to modify the original guidelines. Additional goals were to assess the results of the project in terms of judicial compliance, jurisdictional differences, and explained sentencing variation during the first 11 months of guidelines implementation. During the rest of the summer, the guidelines manual was completely revised and distributed with the modified guidelines. The modified guidelines went into effect in the four test sites for convictions on or after November 14, 1982. In its 1983 session, the Maryland General Assembly passed a statute which permitted the use of judicial guidelines in setting sentencing, so long as the guidelines do not prescribe sentences which exceed the statutory maximum or violate mandatory minimum sentences. Finally, in May 1983, the Judicial Conference voted to introduce statewide sentencing guidelines. These were implemented in Maryland on July 1, 1983.

Unlike Florida, the statewide guidelines in Maryland retained many of the primary elements of the multijurisdictional test, in that they were judicially developed, implemented through judicial concensus rather than a judicial or legislative mandate, and contained no provisions for appellate review. However, as in the multijurisdictional experiment, Maryland's statewide guidelines were not developed primarily through empirical analysis; normative considerations were a major factor in their construction.

3.4 Guidelines Impact

As discussed in Chapter 7 of Part II, compliance covers a range of issues, from completing necessary paperwork--i.e., filling out scoresheets for each eligible case and providing reasons for out-of-range sentences--to using guidelines in judicial decisionmaking and rendering sentences according to the recommended ranges.

Table 3-4 summarizes the analyses of Chapter 7 and reveals that compliance in both states fell short on a number of these dimensions. For example, while judges were generally aware of the guidelines, they varied in the extent to which they used

TABLE 3-4

SUMMARY OF COMPLIANCE FINDINGS
FOR FLORIDA AND MARYLAND

Compliance Issue	Florida Findings	Maryland Findings
Did judges consult the guidelines in sentencing?[a]	General awareness of guidelines but sharp differences in use. Strong jurisdictional variation; minimal use in urban court, related to high volume of plea negotiated cases.	General awareness and use of guidelines, though variations in their role in the sentencing decision. Reports of over-compliance with recommended ranges.
Were scoresheets completed on eligible cases?[b]	Scoresheets were filed on only about 60 percent of all burglary cases; there were striking jurisdictional differences.	Scoresheets were available for just under 70 percent of all burglary cases, with some variation among circuits.
Did sentences fall within the guidelines ranges?	Scoresheet sentences were within guidelines for 78 percent of all cases (from 76 to 84 percent by circuit). There were significant differences in the direction of extra-guidelines sentences, by jurisdiction and crime category. Incarceration rates also varied widely.	Scoresheet sentences were within guidelines for 68-72 percent of all cases, with virtually no site variations in overall agreement or in the direction of extra-guidelines sentences.
Were reasons provided for extra-guidelines sentences?	Nearly 80 percent of scoresheets with extra-guidelines sentences gave reasons. There were significant site differences (with a range of 74 to 95 percent).	Just over 50 percent of the cases requiring reasons actually had reasons on the scoresheet; there was some variation by site and type of crime.
What kinds of reasons were given for extra-guidelines sentences?	Reasons given for above-guidelines sentences often challenged the guidelines (by citing factors already in the scoring), but mitigating circumstances were provided for most below-guidelines sentences. Reasons were frequently lacking in specificity.	A substantial proportion of reasons for both aggravated and mitigated sentences challenged by guidelines (by citing factors already counted in the scoring). Reasons were often too vague to provide useful feedback for guidelines revision.

[a] Primary discussion in Chapter 5.

[b] Findings subject to the limitations of the scoresheet data sets collected by the Florida and Maryland projects. See discussion of data quality in Chapter 7 and of monitoring in Chapter 5.

them. Overall, the scoresheet filing rate in Florida was only about 57 percent, while the Maryland filing rate was approximately 70 percent. While the vast majority of sentences were within the guidelines ranges, judges tended to give very vague reasons for extra-guidelines sentences. This served to reduce the usefulness of the written reasons as a feedback mechanism for guidelines revision.

Table 3-4 also reveals that:

- In Florida, there was considerable site variability, not only in judicial use of guidelines but also with respect to the relationship between actual and recommended sentences and the provision of reasons for extra-guidelines sentences.

- The proportion of scoresheet sentences falling within the guidelines ranges in Maryland was relatively low, considering the fact that Maryland's sentencing ranges were very broad and overlapped considerably.

- Reasons given by Maryland's judges frequently challenged the guidelines by citing factors already taken into account in the offender or offense scores.

Chapter 8 of Part II of this report describes in detail the impact of the multijurisdictional field test on sentencing disparity and severity. Examination of data on all burglary cases sentenced during the test year allowed us to compare sentencing for a single type of offense during the test period with sentencing during an earlier period and in a set of non-participating sites. Summarizing the results of this analysis briefly, we found that the introduction of sentencing guidelines had at best a modest impact on sentencing behavior in the participating sites.

In Florida, the burglary sentences in both test and comparison sites were not more uniform during the test period, once adjustment was made for case characteristics. The use of guidelines did not reduce disparity within or across jurisdictions. However, sentence severity in the test and comparison sites appeared to increase somewhat during the guidelines test year. Although the introduction of guidelines may have had an impact on severity, it seems more likely that increased severity stemmed from changes in the local climate regarding crime. Moreover, while of interest, an increase in sentence severity was certainly not the objective of this sentencing reform.

Our analysis of Maryland burglary data revealed that the average net sentences increased in both the test and comparison sites during the guidelines test period, though the increase was greater in the comparison sites than in the test sites. Case characteristics remained fairly stable during the pre-guidelines and guidelines

test periods; thus, it is likely that similar crimes--at least for burglary offenses--carried more severe sentences during the period when guidelines were in effect.

Using the burglary data, models were constructed to enable us to estimate sentences as a function of case characteristics before and during guidelines. This analysis showed that offenders with serious prior convictions received significantly more severe sentences under guidelines than they had before. This pattern was confirmed in all four guidelines sites. Other factors (offense seriousness, restricted status, and lesser conviction offenses) played a smaller role in overall sentencing under the guidelines, increasing in some sites and decreasing in others. Convictions for multiple offenses were treated about the same before and during guidelines.

When we controlled for case characteristics, the burglary data analysis also showed potentially important findings concerning sentencing disparity. A modest but significant decrease in sentencing variation was found in the Maryland test sites; no decrease was found in the comparison sites. Further examination showed that this reduction in disparity was due to one site only--Baltimore City. None of the other test sites showed any decrease in variability, either when considered individually, or when data from the three smaller test sites (omitting Baltimore) were pooled. The implementation process in Baltimore may account for the positive finding in that city. As described in Chapter 8, Baltimore judges had many reasons for increased support of the guidelines effort, which may have influenced the sentencing outcomes in that city. In addition, the guidelines project offered effective training and support to all the Maryland test sites.

Considering the Florida and Maryland findings together, it is apparent that the impact of guidelines on sentencing is both limited and conditional. There seems to be the potential for a positive impact. Not surprisingly, it appears that this potential is strongly influenced by the ways in which guidelines are developed, implemented, and used by judges. In the remainder of Part I, we view these findings in the light of these mediating factors--the technical properties of the guidelines as constructed by the sites (Chapter 4), how they were implemented (Chapter 5), and the ways in which the system adapted to the change (Chapter 6).

CONTINUED

1 OF 4

FOOTNOTES

1. Joseph C. Calpin, Jack M. Kress, Marilyn A. Chandler, Mona Margarita, Susan Mitchell-Herzfeld, Arthur M. Gelman, and Barbara A. Broderick, "The Analytical Basis for the Formulation of Sentencing Policy," January 1978, p. 3. (Unpublished).

2. National Institute of Law Enforcement and Criminal Justice, Law Enforcement Assistance Administration, U.S. Department of Justice, Multijurisdictional Sentencing Guidelines Program Test Design (Washington, D.C.: Government Printing Office, 1978), p. 4.

3. Memorandum in Program Coordinating Team files, "Field Test of Multi-County Sentencing Guidelines in Two States."

4. William D. Rich, L. Paul Sutton, Todd R. Clear, and Michael J. Saks, Sentencing by Mathematics: An Evaluation of the Early Attempts to Develop and Implement Sentencing Guidelines (Williamsburg, VA: National Center for State Courts, 1982); Richard F. Sparks, Bridget A. Stecher, Jay S. Albanese, and Peggy L. Shelly, with a chapter by Donald M. Barry, Stumbling Toward Justice: Some Overlooked Research and Policy Questions about Statewide Sentencing Guidelines (Newark, NJ: School of Criminal Justice, Rutgers University, 1982).

5. It has been speculated that critical differences between the Parole Commission's situation (where there was successful guidelines implementation) and that of local courts were not recognized by the researchers. See Jolene Galegher and John S. Carroll, "Voluntary Sentencing Guidelines: Prescription for Justice or Patent Medicine?" Law and Human Behavior 7 (1983): 361-400.

6. NILECJ, Multijurisdictional Sentencing Guidelines Program Test Design, p. 45-46.

7. Memo to Eleanor Chelimsky from Michael B. Fishe, Metrek Division of Mitre Corporation, 9/28/78 on Sentencing Guidelines Site Visit to Florida.

FOOTNOTES
(continued)

8. The Florida grant application states: "The urban-suburban-rural criteria outlined in the Multi-jurisdictional Sentencing Guidelines Field Test Site Assessment Form . . . is not directly applicable to the population distribution of Florida. However, the recommended jurisdictions do conform with the state's perceptions of rural, urban and suburban areas and have been approved as such by a representative of the National Institute/University Research Corporation site selection team." Florida Supreme Court, Application for Federal Assistance, submitted to the National Institute of Law Enforcement and Criminal Justice, May 23, 1979, p. 9.

9. Alan C. Sundberg, Statewide Sentencing Guidelines Implementation and Review: A Report to the Legislature (Tallahassee: Supreme Court of Florida, 1982), p. 5.

10. Based on criminal cases filed between July 1, 1981 and June 30, 1982. Annual Report of the Maryland Judiciary: Statistical Abstract, 1981-1982 (Annapolis, MD: Administrative Office of the Courts).

11. These crimes include murder, sexual battery, robbery, burglary, arson, aggravated assault, aggravated battery, kidnapping, escape, breaking and entering with intent to commit a felony, and aircraft piracy. In addition, battery upon a law enforcement officer or firefighter who is on duty carries a similar penalty.

12. FLA. STAT. §947.16(3).

13. FLA STAT. §947.16(3)(f).

14. FLA. STAT. §921.16.

15. FLA. JUR. 2d Criminal Law §887.

16. State of Maryland, Commission on Criminal Law, Report and Part I of the Proposed Criminal Code, Baltimore, Maryland, June 1, 1972, p. xiii.

17. See, e.g. Lynch v. State, 2 Md.App. 546, 236 A.2d 45 (1967); Kirkorian v. State, 233 Md.324, 196 A.2d 666 (1964); McNeil v. Warden of Md. House of Correction, 233 Md. 602, 195 A.2d 612 (1963); Tillet v. Warden of Md. House of Correction, 215 Md. 596, 135 A.2d 629 (1957).

FOOTNOTES
(continued)

18. Young v. State, 151 A.2d 140 (1959); Crowe v. State, 240 Md. 144, 213 A.2d 558 (1965); Tucker v. State, 237 Md. 422, 206 A.2d 691 (1965); Berger v. State, 179 Md. 410, 20 A.2d 146 (1941); Manly v. State, 7 Md. 135 (1854); Cothorn v. Warden of Md. Penitentiary, 155 A.2d 652 (1959); Williams v. State, 205 Md. 470, 109 A.2d 80 (1954); Eyer v. Warden, 197 Md. 690, 80 A.2d 19 (1951).

19. MD. CTS. & JUD. PROC. CODE ANN., §4-301.

20. Joseph C. Howard, Sr., "Sentence Review: The Maryland Experience." (Typewritten), pp. 8-10.

21. In the case of the 15th Circuit, the Chief Judge's designate.

22. Florida Supreme Court, Application for Federal Assistance, Part III, Budget Exhibit I, budget narrative.

23. In Florida, this is the Office of the State Courts Administrator. In Maryland, it is the Administrative Office of the Courts.

24. NILECJ, Multijurisdictional Sentencing Guidelines Program Test Design, p. 53.

25. Sundberg, A Report to the Legislature.

26. Appeals, escapes, and probation revocations were further excluded.

27. The problems mentioned here are discussed more fully in Chapter 4.

CHAPTER 4

DEVELOPMENT AND DESIGN ISSUES CONCERNING EMPIRICAL SENTENCING GUIDELINES

This chapter focuses on the meaning and consequences of the empirical basis of the guidelines in the multijurisdictional field test. As we have seen, the test design--which mandated empirically based guidelines--grew out of a line of research that originated with parole guidelines and then transferred the methodology to the area of criminal sentencing. We have also seen that, subsequent to the start-up of the projects in Florida and Maryland (in fact, at about the time their guidelines development was completed), draft results of an evaluation of the prior generation of NIJ-sponsored research raised the first critical questions concerning the value of empirical guidelines per se.[1]

In this chapter, we analyze the development of the Florida and Maryland guidelines for two purposes: to highlight development and design issues relevant to the creation of other guidelines systems (whether empirical or not), and to assess the advantages and drawbacks of the empirical approach. The chapter thus seeks to offer some practical guidance to other jurisdictions, but also to pose a very basic question about whether the benefits of the empirical approach are worth the costs. We believe this question should be examined carefully before states embark on new guidelines efforts.

With respect to development and design issues, empirically based sentencing guidelines have several important characteristics:

- they are constructed using information about sentencing practices gathered from court records on criminal case dispositions in the participating jurisdictions;
- statistical techniques are used to discern patterns in these data on past sentencing practices, such as how offender characteristics or the facts of the crime are related to the sentence given;
- these patterns are interpreted as the "latent" or "implicit" sentencing policy of the judges and jurisdictions from which the data were drawn.

These statistical analyses are then translated into guidelines grids with recommended sentence ranges corresponding to case characteristics.

To a considerable degree, the issues concerning empirical guidelines are framed in reference to the alternative concept of normative or prescriptive guidelines. Normative guidelines have two defining characteristics:

- they are policy-based (developed out of direct debate about appropriate criminal sanctions); and
- they embody explicit sentencing goals (e.g., rehabilitation, retribution) in the way that they relate particular crime characteristics to the recommended punishment.

Normative guidelines also typically take the form of grids, but the sentence ranges and case characteristics are chosen through policy debate.

Two other features frequently associated with normative guidelines are also important to this discussion. Such guidelines may be developed with explicit recognition of criminal justice resource constraints (particularly those posed by prison and jail capacity). Normative guidelines are also likely to be mandatory rather than voluntary--a feature extremely important to implementation and impact issues. This follows from the need to successfully constrain decisions to resource limitations.[2]

Normative sentencing guidelines have been developed in two states: Minnesota[3] and Washington.[4] We would not argue that this limited experience supports firm conclusions about the relative virtue or viability of the normative approach. However, the contrasts between empirical and normative guidelines offer a useful framework for posing important issues about guidelines development and design in the multijurisdictional field test.

The chapter is organized in four sections. The first concerns the process of guidelines development in Florida and Maryland. The second focuses on the guidelines grids and supporting materials used during the test year. The third examines revisions to the guidelines during that period. Each raises questions about the effects of the empirical approach followed in both states. A concluding section discusses these questions further.

4.1 Guidelines Development

Because of the empirical mandate in the test design, the development process consisted of research design, data collection and data analysis phases. However, any guidelines system that involves recommended penalties related to offense or offender characteristics must accomplish a number of similar development tasks.

The multijurisdictional guidelines development entailed several important choices--key decisions for the ultimate scope and shape of the guidelines in each

state. These choices are summarized and contrasted in Table 4-1. First, offense classification sets the basic framework of recommended penalties under any guidelines system. The Florida Board's initial decision concerning offense classification had favored crime-specific sampling and analysis. However, for the purpose of sampling, the staff developed and used a six-way categorization of offenses.[5] Each category was meant to be fairly homogeneous with respect to the offense and offender characteristics influencing the sentence decision-making process.[6]

The types of cases to be covered by the guidelines was a second issue for the guidelines developers. Obviously, offense coverage determines the potential scope of impact of this sentencing reform. Florida's six categories covered 65 separate statutory offenses, accounting for 85 percent of the criminal cases they examined while preparing to gather data on past sentencing practices. These 65 offenses define the ultimate coverage of the Florida guidelines in the multijurisdictional test.[7] In addition, it was apparently assumed by most Board members that the guidelines would apply to dispositions by plea as well as by trial; this assumption became an issue within the Board late in the development period, but it did prevail. The judges also agreed to include cases where adjudication was withheld. In developing its crime categories, Maryland adopted a tripartite division of offenses, separating crimes against persons from drug offenses and crimes against property.[8] Early data collection made it clear that crime-specific analyses would not be possible using the construction sample (the data collected as the basis for empirical analysis to develop the guidelines). The three-way classification seems to have been the only one considered by the Board. Guidelines coverage was defined largely by court of origin rather than statute, due to the lack of criminal code structure and the continued charging of common law offenses. The Board discussed the coverage of all disposition types, and indeed made provision (both in construction sample collection and guidelines usage) for distinguishing sentence bargains in which the judge explicitly binds himself or herself to the agreement (so-called "ABA pleas") from all other pleas.[9]

The unit of analysis is an important consideration in developing empirically based sentencing guidelines, both because it influences the outcomes of the statistical analysis and because it may determine how the guidelines are to be applied. (This choice is related to the choice that must be made about how to handle linked convicted charges under any guidelines system.) In planning for collection of the construction sample, the Florida staff chose as its unit the "sentencing event" (defined in Table 4-1), and this became the basis of guidelines implementation as well. By contrast, the Maryland project sampled by convicted count; project staff chose this approach since Maryland judges must impose a sentence for each convicted count. However, as suggested in the test design, staff collected some case-level information. In actual guidelines use during the test year, a separate scoresheet was to be filled out on each convicted count being sentenced at one time, but rules linked the sentences on the various counts relative to the guidelines ranges.

Table 4-1

Main Choices in Guidelines Development:
Florida and Maryland

Main Choices	Florida	Maryland
Offense Classification	Six offense groups: 1) Murder, manslaughter (etc.) 2) Aggravated battery or assault 3) Burglary (occupied dwelling), robbery, etc. 4) Burglary (unoccupied dwelling, etc.) 5) Larceny, theft, stolen property, etc. 6) Controlled substances	Tripartite: Offenses against persons Offenses against property Drug offenses
Guidelines Coverage	65 offenses (accounting for 85% of criminal dispositions) All dispositions (pleas, trials, adjudications withheld)	Cases originating in the circuit courts (excluding appeals, escapes, probation and parole revocations, prayers for jury trial from District Court, arson of a dwelling, and any other offense not falling within the three offense classifications of person, property and drug)
Unit of Analysis	Sentencing event (all sanctions imposed one one offender on a given date of sentence)	Count (single allegation of single offense); some case level information connected with sample count
Nature and Scale of Data Collection	Sample size: very large Number of variables: very large Validation sample: none	Sample size: large Number of variables: large Validation sample: none

Both states chose to collect large construction samples and to gather information on (code) a large number of questions (variables) concerning each case. These choices arose in part from concern that a sufficient number of cases be drawn to avoid challenges to the constitutionality of the resulting guidelines, especially challenges alleging lack of representativeness of the sentencing practices in the participating jurisdictions. (This had been the basis for a challenge against the New Jersey sentencing guidelines in that state's Supreme Court.) In addition, neither the staffs nor the Boards wished to limit their variable lists (despite the test design's urging), because this might limit later analyses in unanticipated ways. On the other hand, neither project planned for collection or use of a validation sample, which would allow a separate check on the statistical power of the final guidelines scoring. While the basis of these latter decisions is not clear, it is evident in hindsight that the necessary time and budget would not have been available to carry out validation in either state.

We have highlighted these four choices because they were important to the way that guidelines were developed in Maryland and Florida and the way they were implemented in the multijurisdictional test. More generally, these choices must be made whenever guidelines are empirically based, and some apply even to development of wholly prescriptive guidelines (e.g. offense classification, guidelines coverage). Thus, the discussion of their evolution and implications in the rest of this chapter should prove useful to future guidelines efforts.

4.1.1 Research Design and Data Collection

Guidelines development begins with an effort to define the scope of guidelines coverage relative to the criminal code. Under the empirical approach, historical data on criminal sentencing are then collected, to form the basis for analysis and design of the guidelines grids.

The test design's guidance in the areas of research design and data collection concerned four main points:

- the need for preliminary analyses of the state's criminal code and the information available to judges when sentencing;
- a requirement that the data base be drawn from all the participating jurisdictions in the state;
- a suggested restriction of the data collection to 30-50 variables (core list provided); and
- direction to collect a validation sample after initial grid development, for purposes of examining model fit.

Supplementary references to Albany Criminal Justice Research Center (CJRC) documents and technical assistance from one of the Albany researchers contituted the additional guidance available.

The Florida project staff had gone through a somewhat similar data collection exercise for the Sentencing Study Committee, and had explored the state's criminal statutes and information sources for that work. For guidelines development, it was decided to use a three-year time frame in sampling cases from the four participating circuits; because the policy for rotating judges onto the criminal bench varies among circuits, this was considered the minimum amount of time needed to ensure that a representative caseload for the full set of judges would be included.[10] Staff then compiled a list of all criminal dispositions for that period in the four sites, using written docket books, minute books and (when available) computerized dockets from each county. Significant problems were encountered with the availability and condition or quality of docket information; in particular, methods of recording cases and docket numbers varied greatly, offense was often noted by generic title (e.g. burglary, assault) rather than statute citation, and one circuit did not have a docket per se. However, a small number of information items was collected for over 16,000 cases, enabling staff to identify the 65 most frequent statutes for convictions and to consolidate cases into sentencing events. This adjusted list of cases under the 65 statutes became the sampling frame for the construction sample.

The Maryland staff recommended, and the Board approved, a data collection effort that would take the instant count as the unit of analysis and would span calendar year 1979. It is unclear whether estimates were developed of how many cases this might yield; such figures might have been difficult to derive, since court statistics or docket entries may not correspond to single counts. For sampling, efforts were made to gather disposition data from the State's Administrative Office of the Courts for Harford and Prince George's Counties, and from the circuit court data systems in Baltimore City and Montgomery County. However, the Maryland data collection began without either a sample design or a sampling frame for any jurisdiction.

In developing the codebook and data collection instrument, the Florida staff again drew heavily on the work done for the preliminary sentencing study. That project's list of 194 variables was supplemented by an examination of the items collected for the original CJRC feasibility studies and for sentencing guidelines studies in New Jersey, Michigan, and Minnesota, as well as suggestions of the Advisory Board. The final decision to collect 220 variables was considerably at odds with the test design's target of 30 to 50. While reluctance to prejudge the importance of particular factors as influences on sentencing is understandable, the Florida staff had the advantage of an earlier analysis which could have guided selection of a subset of information items. In the end, only about half the variables were used, and staff members recognized that time and resources had been wasted by the scale of the data

collection. Yet, given the perceived importance of this step in developing empirical guidelines, it is not surprising that this approach was taken.

Maryland's codebook had its origins in a survey of Advisory Board members regarding the information sources that were available to judges and the factors they viewed as influential in the sentencing decision. The Board reviewed a first draft of the codebook and data collection instrument in March. By the time the final version was approved, in May, it had grown from 90 variables to 132. Two judges had tried the instrument out on cases coming before them for sentencing; in addition, the staff collected some cases for pretesting and training.

The NIJ test design notes that the types of guidelines models to be developed will be a critical consideration in design of the construction sample; however, the discussion in that document focuses on whether the analysis will concern the decision to incarcerate (IN/OUT) or sentence length. In a different respect, grid design influenced Maryland and Florida sampling. By deciding to develop separate crime categories, and sampling cases for each crime category rather than all crimes, both sites essentially drew stratified (rather than simple) random samples; thus, the required sample sizes were reduced and statistical efficiency was increased. These categories were the framework for later grid development. Florida's sample design involved 6 crime categories by urban, suburban, and rural[11] jurisdictions (18 strata), while Maryland's ultimately involved 3 offense groups by 4 sites (12 strata).

The primary problem encountered by the Florida project during the actual data collection concerned information sources. In an initial assessment of sources available to judges for sentencing, the Presentence Investigation Report (PSI) was identified as the most important single vehicle for presenting offender and offense information to the judge. However, it was recognized that PSIs would not exist for all cases, and that other sources would need to be used for perhaps 20 percent of the sample. The earlier sentencing study had been limited to cases with PSIs, so staff had no benchmark for the effort required to use alternative records. However, had the staff included only cases with PSIs in their construction sample, a substantial number of cases disposed by plea would have been omitted, giving an incomplete and possibly biased picture of current sentencing practice. (In some jurisdictions, where plea negotiations settle the majority of cases and where PSIs are virtually never prepared for them, the construction sample would have represented only a trivial portion of the full caseload.)

For cases without PSIs, the Advisory Board approved the use of postsentence reports.[12] However, a number of additional steps became necessary when it was found that both pre- and post-sentence reports were often unavailable or lacked necessary information:

- court case files were used because PSIs were frequently missing details on arraignment and sentencing;

- prison admission summaries were requested from the Department of Corrections for about 500 cases with neither pre- nor post-sentence reports;

- Florida Department of Law Enforcement rap sheets were used to fill in facts about criminal history;

- court clerks were consulted for information on conviction charge, as well as consecutive (or concurrent) sentencing to enable staff to determine when separate offenses should be combined into a single "sentencing event."

Despite these efforts, many Florida construction sample variables were missing data for a substantial portion of the cases.

Maryland's data collection problems began with the lack of a sample design or frame and were compounded by difficulties in locating and coding the data. It was recognized in advance that Baltimore City would have by far the largest caseload and greatest need for sampling, and it was hoped that considerable time could be saved by obtaining computerized records of criminal court cases. But the efforts and time expended in obtaining a computer tape of Baltimore City criminal dispositions proved fruitless, when a cross-check with dockets as the data collection started showed the computerized list to be very incomplete. Thus, some coders were assigned to manual listing of cases from the dockets at the same time that others began selecting and coding files.

In the other participating Maryland jurisdictions, the sampling frame was also manually compiled while coding of case files proceeded. The staff, supervising coders specially hired for the summer, did not feel the project could afford to halt the coding while the sampling frame was completed and a sample drawn. Some guesses were made as to the types of cases that would occur less frequently; since these would have the highest likelihood of falling in the sample, they were coded first. However, extra cases were certainly coded in this period and discarded later. By mid-August, this strategy was abandoned; coding was halted to permit completion of the sampling frame and development of a sample design. Sampling fractions (the proportion of cases to be drawn in each site and crime category) were set in September, using a design stratified by site and offense type, and data collection then resumed.

The second data collection problem encountered in Maryland concerned information sources and access. PSIs and case files were available for most of the sample in Prince George's, Montgomery and Harford Counties, although they were

found in varying locations. Other sources were not used if a PSI had not been prepared. In Baltimore City, however, the staff found that separate case files were created for each criminal count charged out of a single criminal incident. Thus, from 1 to 20 case files per incident had to be retrieved before coding, because there was no way of knowing in advance which file contained the PSI. Clerks placed limitations on what materials could be examined and how long they could be held, although these issues were later resolved. Further, it turned out that about half the first 1,000 Baltimore cases had no PSI, transcript or police report in the files, resulting in information gaps for two major groups of variables: "real offense" characteristics and adult criminal history. Even after prosecutor files were accessed to supplement the court files, the Maryland construction sample--like that of Florida's project--contained substantial areas of missing data.

Of course, missing data can present difficulties in creating valid empirical sentencing guidelines. If some information used by judges in setting sentences is not reflected in the court records analyzed by researchers, the resulting empirical guidelines may be unnecessarily imprecise, or they may incorrectly attribute undue importance to other correlated factors. This can happen if relevant information about the defendant is presented in oral arguments or in judge's chambers, or the judge personally observes defendants' characteristics that he or she considers relevant for sentencing.

A somewhat different kind of missing-value problem arises when data items are unavailable to both the court and the researcher for some (but not all) defendants. The empirical guidelines would typically be developed using only the nonmissing data, after which the results may incorrectly describe sentencing for the group as a whole because they are not applicable to cases with missing data. Technically, a preferable approach is to fit equations to all cases, using special coefficients to signify missing data. But no ethical justification can be given for applying the information gained from such equations in a guideline system, since some defendants would then receive longer or shorter sentences solely because data were missing in their official records.

One other problem that arose in Maryland's data collection is worth noting. The Advisory Board judges made revisions that affected the coding after it was already begun, necessitating an effort to backtrack, retrieve files a second time, and recode. As we shall see, the Board later chose variables for scoring on which information either had not been collected or was missing for substantial numbers of cases. As a result, the construction sample was more expensive and far less useful than it could have been. In this area, more decisive guidance by the staff might have been appropriate.

Reviewing the research design and data collection experience of the two projects together suggests these general lessons for empirical guidelines efforts:

- early policy decisions (e.g., about the appropriateness of considering "real offense" behavior) can save considerable data collection resources;

- because of the nature of court records, construction of sampling frames will consume substantial time and effort, but these are well-spent to avoid later field problems;

- thorough advance reconnaissance on information source locations and gaps is essential for planning data collection, both in terms of arrangements that can be negotiated to facilitate access and in terms of gauging coder work volume and work flow.

4.1.2 Data Analysis

In this section, we review the data analysis conducted in developing the Florida and Maryland guidelines. The paradigm for developing empirical guidelines involves applying statistical techniques to the data on past cases in order to identify the prevailing patterns of sentencing practice. Our purpose is not to criticize in detail the analytic approach embodied in the Albany model of empirical guidelines development; this has been done thoroughly elsewhere,[13] and it is the broader issues with respect to empirical guidelines that are the focus of this chapter. Nor do we propose to second-guess particular analytic choices made or techniques used by the project staffs. However, we can identify problems in this phase of development that are likely to be encountered by other guidelines efforts, and the consequences of these problems for the final shape of the grids.

The NIJ test design offered guidance on three primary points with respect to the empirical analysis:

- it was recommended that the grids be shaped on the basis of analyzing the decision of incarcerate (IN/OUT) rather than sentence length;

- bivariate relationships could be used to shorten the list of factors being tested; and

- a simplified weighting scheme would be the best choice for final grid definition.

In a meeting with both project staffs just before the data analysis stage, the technical advisor who had been involved in the Albany work emphasized that the point of the analysis was to reduce the number of factors to be considered in the guidelines, not to produce the final grids or factor weights.

Initial efforts of the Florida staff went into a detailed analysis of sentencing on only one category of offenses (Category 2--aggravated battery and aggravated assault). Separate analyses for uban, rural and suburban cases did reveal significant differences by type of jurisdiction in the set of factors influencing sentence. While a great deal of concern had been expressed earlier about the potential divisiveness of such findings, the Advisory Board judges did not react by defending their circuits' own practices. Instead, they discussed the meaning and appropriateness of each factor and, by a series of votes, reduced the three sets of factors to a single common set.[14]

Based on the relative ease with which the Florida Advisory Board arrived at a uniform set of factors for all jurisdictions, the urban/suburban/rural analysis was not conducted for the remaining categories.[15] Pooled data for all four sites were used to examine the factors influencing sentences for other types of offenses. Four weeks later, the staff presented to the Board a draft set of six guidelines grids, based on the factors found statistically significant and their weights in multivariate equations.[16] Again, the judges voted to eliminate some factors and add others. Some point scores (weights) were also adjusted, where the Board agreed that the corresponding recommended sentences were not in line with their views and experience. After this final pre-implementation Board meeting, the staff completed the analysis, creating point scores for the final set of variables where the judges had not set them.

When the Maryland construction sample data were finally ready for analysis, in January 1981, a March 1 start-up of guidelines implementation was still being planned. During the January Board meeting, the staff discussed plans for analyzing the data by the property/violent/drug groupings and raised the issue of sample weighting. Weighting determines how much influence a case or group of cases can have on the analytic results.[17] Because the judges were concerned that weighting by true population proportions might give too much influence to the Baltimore data (expected to make up about two-thirds of the cases), they requested that the analysis also be conducted with Baltimore balanced equally against the other three sites.

As noted in the project's progress report for this period, the 1979 data were analyzed using cross-tabulation, multiple regression and discriminant analysis technqiues. Preliminary analyses indicated that jurisdictional differences would not play a significant role. In addition, staff found that the weighting alternative discussed at the January meeting was rendered moot. Due to the large proportion of Baltimore cases missing PSIs (50.8 percent), many critical variables available only from PSIs were missing. Any analyses using these variables would therefore be based on about the same number of cases from Baltimore as from the other jurisdictions.

Analyses were conducted both on the decision to incarcerate (IN/OUT) and on length of sentence. In an early February meeting, staff reported to the Board that the results of these analyses were weak and inconclusive. While the most important independent variables could be identified,[18] they accounted for only 50 percent of the variance in either incarceration or length of sentence. The staff advised that, because these results were not more definitive, the Board would have to exercise policy judgment in selecting and scoring the factors for the guidelines grids.

4.1.3 Development Issues

Three general issues may be raised in light of the development process just described. Each concerns a consequence of choosing the empirical approach, as compared to the work of creating normative or prescriptive guidelines.

First, the research design, data collection and analysis required to develop empirical guidelines can be a substantial drain on time and resources. If the magnitude of the data collection task alone were not enough, the state of criminal case records in most jurisdictions would guarantee the significant problems would absorb extra time and resources beyond the most conservative budget and schedule. Even the past experience of the Florida staff did not protect them from this, once the need to include cases without PSIs (largely plea-negotiated cases) was established. Similarly, early efforts to obtain case listings still did not avert major delays for the Maryland staff. As a consequence of data collection problems, both projects had much less time for the analysis phase than originally planned, and less time than required to thoroughly examine sentencing patterns. Thus, the degree to which the construction sample data were used for analysis hardly justified the effort put into building the two construction samples.

The second issue concerns the relationship between staff and Advisory Board under the empirical approach. In the development process, there is an enormous amount of work in which an advisory board can have no part. Staff must exercise considerable control over the details of data collection and analysis, within broad parameters defined as much by the CJRC methodology as by the policy decisions of the Board.

Another aspect of this issue concerns the emphasis on technical language and statistical methods inherent in the empirical approach. Policy-making boards are not likely to have many members conversant with empirical research, and a substantial communications gap can result when the staff is technically trained but unable to translate analytic findings into lay terms. It is not clear that the judges of the Florida Advisory Board (the voting members) really understood the results of the analysis, although they did exercise some authority to modify them for policy reasons. The Maryland Board was left baffled and concerned by the inconclusive nature of their staff's empirical results. It seems inevitable that the choice of the empirical approach

shifts authority and control away from the policy body and toward the staff, even when problems arise in the data collection or analysis.

The third issue concerns the actual payoff from the time and resources devoted to the empirical development process. With regard to Maryland, it is evident that--for reasons of technical difficulties, policy issues or both--the analyses did not produce the definitive results expected by both the staff and Board. Much of the CJRC literature and the way empirical guidelines were sold to the judiciary implied that these analyses would translate past sentencing practice into future sentencing guidance, without the need for active policy-making. Even in Florida, where the analyses provided substantial basis for grid design, many policy decisions still had to be made. Thus, shaping the grids was a more ambiguous and more normative exercise than either the project staffs or the Advisory Boards were led to expect.

4.2 Designing the Guidelines Grids

How were the Florida and Maryland guideline grids shaped out of the research and the staff-board interaction we have described? What were the notable characteristics of the sentencing matrices first implemented for the multijurisdictional field test? This section addresses both the process question and the product question.

The NIJ test design, developed from the experience of the feasibility tests conducted by CJRC, made several assumptions about the form of the sentencing guidelines that would be developed:

- the grids would be two-dimensional, one axis scoring characteristics of the offense and the other characteristics of the offender;
- the units on the two axes would be set by smoothing the results of analyses focusing on the decision to incarcerate (IN/OUT);
- the grids would present some combination of information items about each cell (e.g. an IN/OUT label, the ratio of historical or predicted IN to OUT decisions, a sentencing range, an average sentence or other measure of central tendency).

While empirically based, it was emphasized that final selection and adjustment of the guidelines matrices would be made by the Advisory Boards, as their representation of sentencing policy.

4.2.1 The Process

As noted above, both the Maryland and Florida projects made initial choices about the ultimate shape of the grids when they decided how to categorize offenses for sampling and data collection. The Florida staff's first presentation of hypothetical grid structures to the Board (in October 1980) offered two alternatives: a set based on the felony degree structure established by Florida statute (life, first, second and third degree felonies) and a set using the six crime categories established by the staff. In both sets, the grid structures were two-dimensional; an offender score was used for one axis, while an offense score was developed for the other axis from the results of an offense severity scaling exercise that had gathered judges' perceptions of relative seriousness among offenses of the same statutory degree. Within each crime category or statutory degree, the offenses were ordered from most to least serious. Examples were shown with a matrix row for each offense and with a row for a group of offenses representing similar scores on the severity index.

At the next Florida Advisory Board meeting, staff presented the initial results of the construction sample analysis of Category 2 crimes, by type of jurisdiction (urban, suburban, rural). The packet of materials included a sample grid for Category 2 based on the urban model. Unlike the ones presented earlier, this sample grid combined the offense and offender scores into a single score, creating a one-dimensional grid. The grid contained two columns; the first showed a series of score ranges based on the combined offense and offender characteristics, and the second showed a recommended sentence range for each score range. The combined scoring of offense and offender factors seems to have evolved from using a multiple regression equation, which included both sets of factors, as the basis of scoring. The project director pointed out to the Board members that this one-dimensional format was different from other grids they had seen, but the Board members had no objection to it.[19]

At this same meeting, as described above, the judges made a number of important decisions about the uniform set of factors to be used. They implicitly decided to eliminate the offense severity index as a factor, instead scoring current offense on the basis of statutory degree. Ultimately, prior criminal record was also scored by statutory degree.

When the Florida staff presented all six draft grids to the Board, at its last meeting before guidelines implementation, two formats were again offered. Both formats used the same composite offense and offender score for one dimension. However, the two-dimensional version used statutory degree for the second axis, so that the recommended sentence would be based on both the composite score and statutory degree. The sample cases used by the Board to try out each category's guidelines were sentenced from these two-dimensional grids. Yet the final product was one-dimensional. The Board's discussion focused on adjustments to the proposed sentence ranges and on the final choice of factors for each of the six crime

- half the Baltimore cases were unusable because of missing data due to lack of PSIs.

Two steps were therefore taken to supplement the 1979 data for purposes of setting sentence ranges. First, sentencing simulations completed by judges at the annual Judicial Education Seminar were analyzed. Second, staff constructed 458 hypothetical cases, to match the cells of the now-complete grids, and asked the 10 Advisory Board judges to sentence them. Based on all three data sources, sentence rages were presented to the Board in mid-April. The staff emphasized the great dispersion and disparity in the data, proposing to center each cell on a median sentence and to specify a range of 40 percent on either side of the median. After considerable debate, the judges widened the ranges further. Cells were labelled "Probation" or given an incarcerative range. Three subcommittees were formed, one for each grid, to refine and smooth the sentences. This task was completed and final approval given at the end of April, and implementation was set for June 1.

4.2.2 The Product

The sentencing guidelines resulting from these two development and design histories offer some interesting contrasts, with important implications for their use and potential impact. Figures 4-1 and 4-2 show the products--with both scoresheets and grids--as they were implemented.[21] Table 4-2 summarizes key features, with their implications discussed further below.

Grid configuration. The Florida guidelines made use of six distinct sentencing grids, while Maryland's used only three. Determining which grid to use for a particular offense was relatively simple in Maryland, since grids were labelled by the ordinary terminology of violent, property and drug offenses. In Florida, however, distinctions among the six grids were not so apparent, being based on similarities in crime characteristics and offense seriousness. During the test year it was common for Florida participants to score cases in the wrong grid.

The contrast of one- versus two-dimensional grids is even more significant. In Florida, a single factor scale fixed the relative importance of offense and offender characteristics for all the crimes covered by a particular grid. For example, the maximum offense score for Category 2 crimes (see Figure 4-1) is 184 points, and the maximum for offender characteristics is determined by the number and degree of prior adult and juvenile convictions. However, grids with separate offense and offender scores for the axes allow changes in the relative weight of these sets of factors, in different parts of the grids. Thus, the Maryland grid for violent crimes gives the offense score increasing weight as the offense gets more serious (i.e. sentences tend to rise faster from row to row than from column to column). Put more generally, the balance of different sentencing philosophies that shape a guidelines grid is formed by the relationship of offense and offender weights.[22]

categories. Factor weight changes were considered but rejected,[20] and the Board's role in shaping the guidelines ended with approval of the grids as amended.

The process of shaping the Maryland guidelines grids began in September 1980, several months before collection of the 1979 data for the construction sample was completed. The staff distributed a set of hypothetical scoresheets and grids to the Advisory Board judges, asking them to try them in the courtroom and provide feedback. Several features of these earliest samples proved important to the final product: three grids were presented corresponding to property, drug, and violent crimes; offender ("criminal history") scores formed one axis of the two-dimensional grids; scoring was in simple, integer units. Alternatives for the second dimension were (a) groups of offenses by seriousness, and (b) offense scores relating to seriousness and other characteristics of the crime. Each cell was labelled IN or OUT, with a range of months indicating the length of probation or incarceration.

Discussion of factors continued in both the October and January Maryland Board meetings, with particular concern about information availability (e.g. for scoring juvenile record). The issue of how to handle sentencing for several convicted counts at once was raised, and the Board members noted the potential for greater prosecutor leverage if extra counts carried added points.

During the early February meeting at which the disappointing analytic results were presented, a number of important decisions were made. A common set of offender factors was chosen for all types of crimes. An offense score was also discussed, but there was agreement only on scoring the seriousness category of the instant offense, plus victim injury in violent crimes. Grid alternatives included a set combining all offense types (grouped by seriousness category) and a set with separate grids for the primary and secondary counts. A final Board decision on scoresheet and grid formats, as well as score elements and weights, was made one week later.

The Advisory Board's choice of factors was little, if at all, influenced by the analytic results, which had been presented in such a way as to downplay their value. Score weights were also set independent of any analysis or modelling. Judges also chose factors for which information was not available, either because the variable was not included in the 1979 data set or because data were missing for numerous cases. When attention turned to using the construction sample data for filling in the guidelines cells with sentences, it was found that:

- no data had been collected on one aspect of scoring the special victim vulnerability factor (presence of a handicap);
- data on juvenile record were missing for virtually all Montgomery County cases; and

Figure 4-1
EXAMPLE OF ORIGINAL FLORIDA SCORESHEET
AND GUIDELINES GRID

Docket No.	Circuit: ☐ 4th ☐ 14th ☐ 10th ☐ 15th	Judge			Date of Sentence / /
Name	Date of Birth / /		Sex: ☐ Male ☐ Female		Race: ☐ White ☐ Other ☐ Black
Primary Offense	Statute	ID Code	Degree	Counts	Date of Conv/Plea / /
Second Offense	Statute	ID Code	Degree	Offense Modifier: ☐ Attempt ☐ Conspiracy	☐ Consecutive ☐ Concurrent
Third Offense	Statute	ID Code	Degree	Offense Modifier: ☐ Attempt ☐ Conspiracy	☐ Consecutive ☐ Concurrent
Assistant State Attorney	Defense Counsel:			PSI Investigator	

		SCORE
1. Primary offense at conviction		
2nd degree felony	45 points	
3rd degree felony	15 points	_____
2. Second offense at conviction		
2nd degree felony	45 points	
3rd degree felony	15 points	
1st degree misdemeanor	3 points	
2nd degree misdemeanor	1 point	_____
3. Third offense at conviction		
2nd degree felony	45 points	
3rd degree felony	15 points	
1st degree misdemeanor	3 points	
2nd degree misdemeanor	1 point	_____
4. Number of counts of primary offense		
One	0 points	
Two or more	21 points	_____
5. Prior Adult Convictions		
Each prior capital felony	100 points	
Each prior life felony	100 points	
Each prior 1st degree felony	60 points	
Each prior 2nd degree felony	30 points	
Each prior 3rd degree felony	10 points	
Each prior 1st degree misdemeanor	2 points	
Every five 2nd degree misdemeanors	2 points	_____
6. Prior juvenile felony convictions		
Each prior life felony	50 points	
Each prior 1st degree felony	30 points	
Each prior 2nd degree felony	15 points	
Each prior 3rd degree felony	5 points	_____
7. Type of weapon used		
None	0 points	
Weapon other than firearm	6 points	
Firearm	12 points	_____
8. Victim precipitation		
Precipitation verified	0 points	
None	16 points	_____
9. Legal status at time of offense		
Free, no restrictions	0 points	
Under some form of restriction	34 points	_____
10. Role of the offender		
Accessory	-24 points	
Alone or equal involvement	0 points	
Leader	24 points	_____
Work Habits: Stable Unstable	TOTAL:	_____

Recommended guideline sentence _____

Sentence imposed _____

Figure 4-1
(continued)

Category 2

(Aggravated Battery, Aggravated Assault)

I. Statutes Included in the Guidelines

Statute	Degree	Description
784.021-1a	3	Aggravated Assault, Deadly Weapon
784.021-1b	3	Aggravated Assault with Intent to Commit Felony
784.045-1a	2	Aggravated Battery Causing Bodily Harm
784.045-1b	2	Aggravated Battery, Deadly Weapon
784.07-2b	3	Battery of Law Enforcement Officer or Firefighter

II. Variables

1. Primary offense at conviction

2nd degree felony	45 points
3rd degree felony	15 points

2. Second offense at conviction

2nd degree felony	45 points
3rd degree felony	15 points
1st degree misdemeanor	3 points
2nd degree misdemeanor	1 point

3. Third offense at conviction

2nd degree felony	45 points
3rd degree felony	15 points
1st degree misdemeanor	3 points
2nd degree misdemeanor	1 point

4. Number of counts of primary offense

One	0 points
Two or more	21 points

5. Prior adult convictions

Each prior capital felony	100 points
Each prior life felony	100 points
Each prior 1st degree felony	60 points
Each prior 2nd degree felony	30 points
Each prior 3rd degree felony	10 points
Each prior 1st degree misdemeanor	2 points
Every five 2nd degree misdemeanors	2 points

Figure 4-1
(continued)

6. Prior juvenile felony convictions

Each prior life felony	50 points
Each prior 1st degree felony	30 points
Each prior 2nd degree felony	15 points
Each prior 3rd degree felony	5 points

7. Type of weapon

None	0 points
Weapon other than firearm	6 points
Firearm	12 points

8. Victim Precipitation

Precipitation verified	0 points
None	16 points

9. Legal status at time of offense

Free, no restrictions	0 points
Under some form of restriction	34 points

10. Role of the offender

Accessory	-24 points
Alone or equal involvement	0 points
Leader	24 points

Figure 4-1
(continued)

III. Sentencing Grid

Category 2	
COMPOSITE SCORE	SENTENCE
0-85	Out
86-95	1 year (1 day-18 mos.)
96-110	2 years (1½-2½ years)
111-125	3 (2½-3½)
126-140	4 (3½-4½)
141-155	5 (4½-5½)
156-175	6 (5½-7)
176-200	8 (7-9)
201-225	10 (9-11)
226-250	12 (11-13)
251-300	15 (13-17)
301-350	20 (17-22)
351-400	25 (22-27)
401-450	30 (27-30)
451+	Life

Figure 4-2

ORIGINAL MARYLAND SCORESHEETS AND
GUIDELINES GRIDS

MARYLAND SENTENCING GUIDELINES PROJECT — Sentencing Worksheet: Offense Against A Person

Offender Name (Last, First, Middle) | Date of Offense / / | Docket Number

Birthdate / / | ☐ Male ☐ Female | Date of Sentencing / / | Sentencing Judge

☐ White ☐ Hispanic ☐ Asian | Instant Count (Title, Md. Code Article and Section)

☐ Black ☐ Amer. Ind. ☐ _____ | Disposition Type ☐ ABA Plea ☐ Non-ABA Plea ☐ Court Trial ☐ Jury Trial

Highest Education: ☐ Less Than High School ☐ High School/GED ☐ More Than High School

Date of Plea/Verdict / /

Jurisdiction: ☐ Balto. City ☐ Harford ☐ Mont. ☐ P.G.

Number of Convicted Counts At This Sentencing Event ☐

PSI ☐ Yes ☐ No

Circle appropriate number in each item below; add circled numbers for offense and offender scores.

OFFENSE SCORE

A. Seriousness of Instant Count
   1 = IV - VI
   3 = III
   5 = II
   8 = I

B. Victim Injury
   0 = No Injury
   1 = Injury, Non-Permanent
   2 = Permanent Injury or Death

C. Weapon Usage
   0 = No Weapon Used
   1 = Weapon Other Than Firearm Used
   2 = Firearm Used

D. Special Vulnerability of Victim
   0 = No
   1 = Yes

OFFENDER SCORE

A. Relationship to CJS When Instant Count Occurred
   0 = None or Pending Cases
   1 = Court or Other Criminal Justice Supervision

B. Juvenile Delinquency
   0 = Not More Than One Finding of Delinquency
   1 = Two or More Findings Without Commitment or One Commitment
   2 = Two or More Commitments

C. Prior Adult Criminal Record
   0 = None
   1 = Minor
   2 = Moderate
   3 = Major

D. Prior Conviction Offense Against A Person
   0 = No
   1 = Yes

E. Prior Adult Parole/Probation Violations
   0 = No
   1 = Yes

F. Employment Record
   -1 = Favorable
   0 = Unknown or Not Applicable
   1 = Unfavorable

☐ OFFENSE SCORE ☐ OFFENDER SCORE

GUIDELINE SENTENCE

ACTUAL SENTENCE

REASONS (IF ACTUAL SENTENCE DIFFERS FROM GUIDELINE SENTENCE)

JUDGE (white); AOC (blue); PROBATION (green); FILE (yellow); PROSECUTION (pink); DEFENSE (gold)

Figure 4-2
(continued)

MARYLAND SENTENCING GUIDELINES PROJECT — Sentencing Worksheet: Property or Drug Offenses

Offender Name (Last, First, Middle)	Date of Offense / /	Docket Number
Birthdate / / ☐ Male ☐ Female	Date of Sentencing / /	Sentencing Judge
☐ White ☐ Hispanic ☐ Asian ☐ Black ☐ Amer. Ind. ☐ _____	Instant Count (Title, Md. Code Article and Section)	
	Disposition Type ☐ ABA Plea ☐ Non-ABA Plea ☐ Court Trial ☐ Jury Trial	
Highest Education ☐ Less Than High School ☐ High School/GED ☐ More Than High School	Date of Plea/Verdict / /	Jurisdiction ☐ Balto. City ☐ Harford ☐ Mont. ☐ P.G.
	Number of Convicted Counts At This Sentencing Event ☐	PSI ☐ Yes ☐ No

Circle appropriate number in each item below; total circled numbers.

OFFENDER SCORE

A. Relationship to CJS When Instant Count Occurred
- 0 = None or Pending Cases
- 1 = Court or Criminal Justice Supervision

B. Juvenile Delinquency
- 0 = Not More Than One Finding of Delinquency
- 1 = Two or More Findings Without Commitment or One Commitment
- 2 = Two or More Commitments

C. Adult Criminal Record
- 0 = None
- 1 = Minor
- 2 = Moderate
- 3 = Major

D. Prior Conviction for Same Type Offense
- 0 = No
- 1 = Yes

E. Prior Adult Parole/Probation Violations
- 0 = No
- 1 = Yes

F. Employment Record
- -1 = Favorable
- 0 = Unknown or Not Applicable
- 1 = Unfavorable

☐ TOTAL OFFENDER POINTS

GUIDELINES SENTENCE

ACTUAL SENTENCE

REASONS (IF ACTUAL SENTENCE DIFFERS FROM GUIDELINE SENTENCE)

JUDGE (white); AOC (blue); PROBATION (green); FILE (yellow); PROSECUTION (pink); DEFENSE (gold)

Figure 4-2
(continued)

SENTENCING MATRIX
OFFENSES AGAINST PERSONS

Offender Score

Offense Score	-1	0	1	2	3	4	5	6	7	8	9
1	P	P	P	3M-2Y	3M-2Y	3M-2Y	3M-2Y	6M-3Y	1Y-5Y	1Y-5Y	1Y-5Y
2	P-1Y	P-1Y	3M-2Y	3M-2Y	3M-2Y	3M-2Y	1Y-4Y	3Y-8Y	4Y-8Y	4Y-8Y	4Y-8Y
3	P-1Y	P-2Y	1Y-5Y	3Y-8Y	3Y-8Y	3Y-8Y	3Y-8Y	5Y-10Y	5Y-10Y	5Y-10Y	5Y-10Y
4	P-2Y	P-3Y	3Y-8Y	3Y-8Y	4Y-10Y	4Y-10Y	4Y-10Y	5Y-10Y	5Y-10Y	5Y-10Y	5Y-10Y
5	P-4Y	P-4Y	3Y-9Y	4Y-9Y	4Y-10Y	4Y-10Y	6Y-12Y	6Y-12Y	6Y-14Y	6Y-14Y	6Y-14Y
6	1Y-4Y	3Y-6Y	3Y-10Y	4Y-10Y	5Y-10Y	5Y-10Y	8Y-15Y	8Y-15Y	10Y-20Y	10Y-20Y	10Y-20Y
7	2Y-6Y	3Y-7Y	4Y-10Y	5Y-10Y	5Y-10Y	5Y-10Y	9Y-15Y	9Y-15Y	12Y-20Y	12Y-20Y	12Y-20Y
8	3Y-7Y	4Y-8Y	5Y-10Y	6Y-12Y	6Y-12Y	6Y-12Y	10Y-15Y	10Y-15Y	12Y-25Y	12Y-25Y	12Y-25Y
9	3Y-9Y	4Y-10Y	6Y-12Y	8Y-15Y	8Y-15Y	8Y-16Y	15Y-30Y	15Y-30Y	25Y-L	25Y-L	25Y-L
10	8Y-15Y	8Y-15Y	8Y-15Y	8Y-16Y	8Y-16Y	10Y-25Y	15Y-30Y	25Y-L	25Y-L	25Y-L	25Y-L
11	9Y-16Y	9Y-16Y	9Y-16Y	9Y-16Y	15Y-30Y	17Y-30Y	17Y-30Y	25Y-L	25Y-L	25Y-L	25Y-L
12	10Y-17Y	12Y-20Y	12Y-20Y	15Y-30Y	18Y-35Y	18Y-35Y	25Y-L	25Y-L	25Y-L	25Y-L	25Y-L
13	12Y-20Y	14Y-22Y	14Y-22Y	18Y-35Y	20Y-40Y	20Y-40Y	25Y-L	25Y-L	30Y-L	30Y-L	30Y-L

P = Probation
M = Months
Y = Years
L = Life

Figure 4-2
(continued)

SENTENCING MATRIX
PROPERTY OFFENSES

Offense	Offender Score										
	-1	0	1	2	3	4	5	6	7	8	9
Arson, Barrack (27-8) Breaking & Entering (27-31A) Theft Less Than $300. (27-342) Other Misdemeanors	P-1M	P-3M	P-3M	1M-1Y	1M-1Y	6M-3Y	6M-3Y	6M-3Y	2Y-5Y	2Y-5Y	2Y-5Y
Attempted Arson, Dwelling or Bldg. (27-10) Bribery (27-23) Daytime Housebreaking (27-306) Forgery & Uttering (27-44) Storehousebreaking $5 or More (27-33) Storehousebreaking Day/Night (27-32) Theft Greater Than $300. (27-342)	P-3M	P-3M	P-6M	3M-3Y	2Y-5Y	2Y-6Y	3Y-7Y	3Y-7Y	3Y-7Y	5Y-10Y	5Y-12Y
Arson, Building (27-7) Burglary (27-29 & 30a)	P-6M	3M-2Y	6M-3Y	1Y-5Y	2Y-5Y	4Y-10Y	4Y-10Y	5Y-10Y	6Y-14Y	6Y-15Y	10Y-20Y

P = Probation

Figure 4-2
(continued)

SENTENCING MATRIX
DRUG OFFENSES

Offense	Offender Score										
	-1	0	1	2	3	4	5	6	7	8	9
Possession of Marijuana (27-287)	P	P	P	P	0-3M	3M-6M	3M-6M	6M-9M	9M-12M	9M-12M	9M-12M
CDS Possession, Except Marijuana (27-287)	P	0-6M	0-12M	6M-18M	1Y-2Y	2Y-2½Y	2Y-2½Y	2½Y-3½Y	2½Y-3½Y	3½Y-4Y	3½Y-4Y
CDS Distribution Schedule I-V Not PCP or Schedule I-II Narcotic (27-286)	P-12M	P-12M	6M-18M	6M-18M	1Y-2 Y	1Y-2½Y	2Y-3Y	2½Y-4Y	3Y-4Y	4Y-5Y	4Y-5Y
Distribution PCP (27-286)	6M-2Y	6M-2Y	1Y-3Y	2Y-4Y	3Y-5Y	4Y-6Y	5Y-7Y	6Y-8Y	7Y-9Y	8Y-10Y	8Y-10Y
CDS Distribution Schedule I and II Narcotic (27-286)	6M-3Y	6M-3Y	1Y-4Y	2Y-5Y	3Y-7Y	5Y-10Y	6Y-12Y	8Y-14Y	10Y-16Y	12Y-20Y	15Y-20Y

P = Probation
M = Months
Y = Years

Table 4-2

Comparison of Key Features of the Florida and Maryland Guidelines Grids

Grid Features	Florida	Maryland
Number of grids	6	3
Form of Grids	One-dimensional	Two-dimensional
Grid Axes	Single combined offense and offender score	Property and drug offense grids: offender score by seriousness of convicted offense. Crimes against persons grid: separate offense and offender scores
Scoring Units	Fine	Coarse
Ease of Score Calculation	Difficult (number of factors, amount of arithmetic, calibration of weights)	Moderate (easy except prior record scoring)
Cell Contents	Median and sentence range or OUT	Sentence range or Probation
Width of Ranges	Narrow-moderate	Very wide
Overlap of Ranges	None	Extensive

Factors and scoring. Many common elements are found between the Florida and Maryland factor lists, among them prior adult and juvenile record, legal status at the time of the offense, victim injury and weapon usage. Both systems based their offense scoring on the count or counts at conviction, rather than on counts as charged. In principle, sentencing is thus linked to the seriousness of the offense as proven to the court. In practice, the convicted charge is often a matter of plea negotiation, and may not reflect all charges for which there is adequate proof. As a result, differences in local policies and practices concerning plea bargains can have a strong impact on the scoring of cases that might be considered identical in their facts. Since jurisdictions vary in their charge bargaining practices, it is crucial to design guidelines systems to accommodate local procedures and norms.

The projects defined the factors and gave directions on how to measure them in the guidelines manuals. The manuals gave instructions on the procedures involved in guidelines operation, including the scoring of each factor. In practice, some factors proved persistently ambiguous and open to argument and interpretation; prime examples were victim precipitation and role of the offender in Florida, victim injury in Maryland.[23] To the degree that this kind of factor carries a significant share of the total possible weight, it is likely that bargaining between prosecution and defense over the scoring will override "objective" fact. That is, factors requiring judgment may lend themselves to manipulation. As other evaluators have noted, clear, distinct and well-defined antecedent conditions (facts about the offense and the offender) are necessary if guidelines are to have any chance of increasing sentence uniformity.[24]

There are great contrasts in scoring between the Maryland and Florida systems, primarily in how finely the scores are calculated (calibration). The Florida weights are based on multivariate coefficients, and the total score for prior record requires multiplying out the number of offenses in each statutory class by the appropriate weight. For Category 2 (Figure 4-1), the scores could run from less than 20 to more than 450 points. The Maryland guidelines weight each factor by simple integers, with no multiplication required (Figure 4-2). Such differences in calibration can have important practical consequences. While the Maryland judges were willing to score cases themselves when no PSI was requested (and thus the scoresheet was not prepared by Parole and Probation officers), it would have been both more difficult and more time-consuming for the Florida judges to do so. Under the time pressures associated with negotiated plea cases in urban jurisdictions, differences in ease of scoring can strongly influence the successful incorporation of guidelines in court routine.

Sentence ranges. A number of aspects of the sentence ranges were handled differently between the Florida and the Maryland guidelines. Nonincarcerative sentences were designated "OUT" (at least initially) in the former, "Probation" in the latter. While the original Florida grids separated OUT from jail or prison sentences, the Maryland grids included some cells spanning these options. Both sets of grids

showed upper and lower sentence boundaries, but Florida's showed a median sentence as well. While Florida's ranges tended to be consecutive (i.e. 5-7 years, 7-9 years), Maryland's overlapped considerably. In addition, many of the Maryland ranges were very wide.

The import of the width and overlap of ranges is clear from the debate of the Maryland Advisory Board on these issues. One judge argued strongly that only narrow ranges provided real sentencing guidance. However, others argued that wide ranges would aid in acceptance of the guidelines, especially since the disappointing analytic results had robbed the Board of a major selling point--that by virtue of their empirical basis, the guidelines were no more than past sentencing practice made explicit.

Even so, the width and overlap of the Maryland ranges are substantial enough to raise questions of equity and cast doubt on whether the guidelines could effect any reduction in sentence disparity. For example, for a violent crime, a 10-year sentence would fall within the guidelines for an offender with an offense score of between 4 and 10 and an offender score of 3 or 4; the same sentence would also apply to an offense score between 3 and 8, combined with an offender score of 5 or 6. In fact, a 10-year sentence would be within the guidelines for close to a third of all the possible score combinations. As another guidelines evaluation notes, "There seems little point in creating guidelines of this form, if the cells of the matrix do not mark off clear distinctions between offense and offender types."[25] Since the score differences may represent wide variation in the facts of the crime and the background of the offender, reduction in disparity is not likely to result from adherence to these ranges.

Guidelines sentence ranges can also have important implications for local acceptance of the reform. Under the empirical approach, the ranges are based on an amalgam of the sentencing practices of the participating jurisdictions; they may resemble typical sentences in any <u>one</u> site to a greater or lesser degree. When ranges are empirically derived and also narrow, as were Florida's, they may prove unpalatable to courts where sentencing outcomes differ greatly from the ranges; this was the reaction of one Florida circuit. Maryland's ranges may well have eased acceptance in the four test sites, since their width and overlap would accomodate a range of local patterns; however, such acceptance would be gained by reducing potential impact on sentence disparity.

Relationship to the sentencing event. A final important contrast in the two sets of guidelines grids concerns the ways they accommodated sentencing of multiple counts or offenses at one hearing. A single Florida scoresheet was to be completed for any sentencing event; up to three convicted offenses received separate points,[26] but none of the convicted counts being sentenced were included in prior record. Maryland's scoresheet was to be prepared by count; if more than one count was being sentenced, a separate scoresheet had to be completed for each. The initial policy presumed concurrent sentences for multiple counts, with consecutive ones considered outside the guidelines and requiring a written reason. The arrangement was awkward,

and the consecutive/concurrent issue occasioned continuing debate and later policy changes.

Scoresheet design. Little has been said thus far about the scoresheets on which the scores were to be entered and the sentences recorded. In Florida, a separate scoresheet was developed for each of the six crime categories[27], while Maryland had one for violent crimes and a second for property and drug offenses. Each had a top section recording basic case information (for administrative and monitoring use) and a bottom section for calculating actual guidelines scores. There were spaces for entering the recommended sentence, the actual one, and the judge's reason for an extra-guidelines sentence. It was already pointed out that complex scoring can prove an impediment to guidelines acceptance in pressured courts. The scoresheet itself can play an important role in identifying problems with guidelines acceptance, if it is designed to provide the necessary information. Florida's scoresheet omitted one item of identifying information that proved very important and useful in Maryland. This was the type of disposition--plea or trial and kind of plea. Data on disposition could have greatly aided the Florida staff in monitoring guidelines usage, as Chapter 5 will discuss in more detail.

4.2.3 Design Issues

A number of issues about guidelines design are raised by the above account of the design processes and products in Florida and Maryland. The first concerns the realm of policy-making under the empirical approach to guidelines. The CJRC paradigm, as embodied in the multijurisdictional test design, involved using statistical analysis to identify the factors for scoring, weight these factors, and determine the sentence ranges corresponding to particular scores. While the Advisory Boards were to make the final selection among alternative grids ("models"), the choices were defined by the analyses and reflected the technical expertise of the staffs. Florida adhered more closely to this paradigm than Maryland, and it is not clear whether, in fact, the Florida Board did make the <u>full</u> set of final choices. Because the Maryland empirical analysis was pronounced inconclusive, that Board had a far greater role in shaping the final guidelines. However, because Maryland judges had been sold on the concept of empirical guidelines as an embodiment of current sentencing practice, the judges from the participating jurisdictions in Maryland felt they lacked the mandate to accept the fuller policy-making role. The wide and overlapping sentence ranges were an explicit result of their concerns about such a role.

The empirical approach is ambiguous not only with respect to the division between analysis and policy, it is also ambiguous in its relationship to jurisdictional differences. If guidelines are built from an analysis of sentences rendered in courts with widely differing philosophies and norms, the multivariate analytic techniques will produce a group of models that correspond to no philosophy or norm in particular. The early analyses performed by the Florida staff did show significant differences both in

the factors influencing sentencing in the test sites and in the resulting sentences.[28] However, when the Board selected a common set of factors for the guidelines, the judges apparently did not consider that the resulting score weights and recommended penalties might differ considerably from the practice of the sites. Therefore, no efforts were made in advance to test reactions to the guidelines nor to address the issue in the context of local norms. It would seem more difficult, in such a context, to justify a set of guidelines derived by statistics than a set of normative guidelines that resulted from explicit consideration of alternative philosophies and that were designed through a recognized policy-making process.

A final key design issue involves the implications of the recommended guidelines ranges. We have already noted that the width and overlap of ranges will certainly have a limiting effect on the degree to which compliace with the guidelines can actually reduce sentence disparity. There are two additional concerns. First, given public attention to the issue of sentence severity, it becomes tempting for guidelines designers to raise the tops of the ranges, making the penalities appear more severe. Of course, this does not necessarily produce more severe sentences, but simply increases the potential for within-guidelines disparity. Second, the empirical approach purports to avoid any effect on jail or prison populations. Since the guidelines sentences would, in effect, reflect the average commitment rates of the past and stabilize those rates for future sentences, incarcerative populations should be maintained at a constant level, barring major changes in the number or seriousness of crimes. As a result, designers of such guidelines have rarely tested out whether their final ranges will alter the number or duration of incarcerative sentences. Normative guidelines systems are often designed with particular concern about correctional resources limits. It would seem advisable that designers of empirical guidelines systems for statewide use address this question directly (through forecasting and through use of monitoring data) rather than assuming that the guidelines will have no effect on population levels.

4.3 Revising the Guidelines

A crucial feature (and selling point) of sentencing guidelines in general is their inherent flexibility; they can easily be adjusted to changes in sanctioning norms or to new concerns among the public or the judicary. Unlike legislatively devised sentence structures, guidelines systems can be modified by administrative procedure, making the system less vulnerable to partisan concerns or the lengthy deliberations typical of the legislative branch. When guidelines are mandated by a legislative body, there may be a requirement of acceptance of changes. However, when guidelines are established by the judiciary, only administrative mechanisms are needed to revise the factors, scoring, and recommended ranges.[29]

All guidelines systems require that judges write their reasons for imposing extra-guidelines sentences (sentences outside the recommended ranges). Under the empirical paradigm developed by CJRC and embodied in the NIJ test design, these reasons were to be the primary ongoing source of feedback concerning guidelines design (the choice of factors, relative weights, grid configuration, sentence ranges, etc.). Judicial reasons were to be analyzed by the central projects for this purpose. Indeed, the test design called for both states to undertake revisions approximately one-third of the way through the test year.

4.3.1 Revisions During the Guidelines Test

During the one year test, neither Florida nor Maryland was able to accomplish a systematic analysis of reasons for extra-guidelines sentences (though Maryland conducted this analysis after the test was over). In large part, this was due to compliance problems: gaps and delays in scoresheet completion and submission, as well as missing or vague reasons.[30] Nevertheless, substantive revisions were made. We focus here on the nature of these revisions and the sources of input actually used for the revision process.

In Florida, two sets of revisions were made to the guidelines during the test year. The first occurred during the second month of guidelines operation, and was based largely on verbal comments and questions about the guidelines and the written commentary solicited soon after implementation by the project director. The second set of revisions was made seven and one-half months after implementation. Again, the primary source of information was not the completed scoresheets, but comments, criticisms and suggestions gathered by the project staff in a series of visits to each of the test sites.

As noted earlier, each of the original Florida guidelines grids contained a sentence recommendation of "OUT" for the lowest point ranges. During the initial stages of guidelines implementation, the precise meaning of this category was unclear, perhaps because no definition of the term was included in the draft manual distributed to the sites. While the guidelines project had intended OUT to mean non-incarcerative sentences (particularly probation), some judges interpreted this to include so called "split-sentences," since Florida judges are permitted to impose up to 12 months of local incarceration as a condition of probation. Recognizing this confusion, staff informed the sites that split sentences would not be considered OUT sentences; written explanations would thus be required if a split sentence were imposed on a defendant whose score fell in the OUT category. In addition, the updated version of the guidelines manual distributed after the training sessions contained a definition of the OUT term excluding split sentences.

Judges' and state attorneys' reactions to this definition ranged from consternation to outrage. The project staff received numerous complaints largely arguing that an OUT sentence range would decrease the deterrent value of the

guidelines sentences. It was clear that this issue was fast becoming a major stumbling block for the guidelines implementation. The problem was discussed in May, when the four voting members of the Advisory Board met at the Circuit Court Judges' conference held in Clearwater. Although none of the ex-officio Board members were present for the discussion (a point which apparently caused some bad feeling), the judges decided to modify the guidelines grids by collapsing the lowest and second lowest cells, creating a single range with a recommended sentence of 0-18 months for all the crime categories except Category 3, which had a range of 0-36 months.

This modification had the intended cosmetic effect of eliminating explicitly non-incarcerative sentence recommendations from the grids, and it may have been a reasonable response to a political reality unanticipated before the guidelines implementation. However, the consequences of this choice were significant:

- Even before the modification, a large proportion of the felony caseload was likely to fall within the lowest range in each grid.[31] With the change, an overall average of 73 percent scored in the lowest range.[32] With so many cases in a single cell, much of the "guidance" potential of the guidelines was effectively nullified.

- Ranges of zero to some term of months do little to promote sentence consistency, since they constrain judges' sentences only at the upper bound.[33] As one Advisory Board member later pointed out, this modification opened the door for greater disparity by permitting sentences of probation for offenders who would have warranted incarcerative sentences under the original matrices.

- By developing a broad sentence range and incorporating incarcerative and non-incarcerative sentences within a single cell of the matrix, the guidelines project lost the opportunity to develop a "case law" of sentencing around the crucial IN/OUT decision. This objection was voiced by an ex-officio member of the Advisory Board after implementation of the change.

During the May meeting, Florida Advisory Board members also decided on a procedural revision to the guidelines: they eliminated the requirement to use probation personnel to verify guidelines scores in plea-negotiated cases. Again, this change was largely in response to the extremely negative reactions from many attorneys and judges and the belated awareness of the great difficulty that missing PSIs would cause in completing the scoring in the urban and suburban jurisdictions. While it is possible that discovery of these attitudes and conditions before the

guidelines implementation would not have prevented all implementation difficulties, certainly the negative feelings generated during the first few weeks of the guidelines might have been diminished if direct Board consideration had been given to the issue of non-incarcerative sentences and if the practical implications of the problems encountered with missing PSIs in data collection had been recognized. The consequences of allowing attorneys to complete the scoresheets are examined in detail in Chapter 6.

Staff contact with the sites was minimal during the period of April through August, 1981. According to the Florida progress report for that period, the intention was to allow each site to adapt the guidelines to its own political and procedural environment, without interference or oversight from the central project. However, in August guidelines project staff conducted a series of visits to all the test sites to speak with State Attorneys, defense attorneys, and probation personnel and to obtain their suggestions for change. Interviews with the judiciary were not conducted, since it was felt that the Board member from each site had had sufficient opportunity to solicit opinions of the judges.

In October 1981, the entire Advisory Board met for the only time to discuss major substantive revisions to the guidelines. Two primary sources of information and feedback were presented to the Board: the first analysis of approximately 1200 scoresheets received by the State Courts Administrators Office, and the results of the staff interviews with site personnel in August. Based primarily on the interview results rather than the data analysis, three significant changes were debated by the Board:

- greater emphasis on the current convicted offenses by modification of the scoring procedure;

- the addition of victim injury as a scoring factor for Category 2 offenses;

- reinstitution of an OUT or "probation" range in the guidelines.

After considerable discussion, it was decided not to adopt this last modification, and only the first two changes were accepted by the Board.

Revision of the scoring of the current offense was particularly significant. By removing the original restrictions allowing scoring on only the primary, second, and third offenses, and by assessing points for all counts of the primary offense, the Board effectively increased the maximum scores that could be received for a single sentencing event under the guidelines. Since the sentencing grids remained untouched, this revision had the net effect of increasing the recommended guidelines sentences for some offenders. A similar, though less dramatic effect was expected through the

addition of the victim injury factor in Category 2.

During the test year of implementation in Maryland, the Advisory Board met several times to discuss practical implementation concerns, the general progress of the guidelines, and modifications to them. However, the revisions made to the Maryland guidelines grids were intentionally kept to a minimum. Since the project experienced continued difficulty in obtaining complete and informative reasons for extra-guidelines sentences, substantive revisions on this basis would have been difficult at best.[34] However, the Advisory Board in essence made this point moot; concerned about resistance and confusion in the test sites and the possibility of due process challenges, members repeatedly decided to defer major revisions until the end of the test period.

Nevertheless, some revisions to the guidelines were instituted during the test. As in Florida, the primary reason for these changes was strong negative feedback from the prosecution and judiciary. In August 1981, the Board discussed a number of possible revisions. Since the employment factor had met with heated criticism for the prosecution, defense, and some judges, deletion of this factor was readily voted. The Board's policy on concurrent and consecutive sentences was also addressed. The original decision had been to sentence multiple-count cases concurrently. While consecutive sentences would be allowed, these would have to be treated as extra-guidelines sentences, explained by written reasons. This policy was fiercely attacked during the initial weeks of implementation and was rescinded at the August Board meeting. However, no alternative policy was established, and the question was left to individual judges to decide.

The Board also considered changing of the prior criminal history scoring, but this was viewed as a major revision, and it was agreed that, as such, it should be deferred. The Board did agree to modify the upper sentence boundary on the crimes against persons category for theft offenses in order to bring the penalty in line with newly enacted legislation. Finally, the Board agreed to one procedural modification: adding space to the guidelines worksheet for institutional or parole recommendations of the judge, the name of the person preparing the materials, and the judge's signature. It was argued that this would facilitate the monitoring efforts of the central staff.

During subsequent Maryland Advisory Board meetings, the issues of prior record scoring and consecutive/concurrent sentences were again raised, but no decisions were reached. Finally, in February 1982, a second set of revisions to the guidelines was made. First, in response to continued criticism of the guidelines sentences for rape offenses, it was decided to alter the victim injury scoring in these cases by presuming a permanent injury rather than a temporary injury. The resulting increase of one point for the offense score produced heavier recommended sentences for rapes. The second modification articulated a new policy on consecutive/concurrent sentencing. It was announced that "sentences for multiple

counts growing out of one criminal event will be presumed to be concurrent unless the judge indicates otherwise and gives his reasons . . . Sentences for multiple counts from more than one criminal event will be presumed to be consecutive."[35]

As in Florida, it was clear that the early changes in the Maryland guidelines were a political necessity. Although deletion of the employment factor was a modest change at best, earlier (and broader) review of the guidelines might have allowed staff to avert the difficulty caused by this factor. More important, however, was the impact of the consecutive/concurrent question. The Board was clearly caught off guard by the strong negative reaction to the original policy. While the hasty decision to abandon it without substituting another was politically expedient, it effectively obviated any "guidance" by the sentencing guidelines in multiple offense cases for a period of six months.

4.3.2 Guidelines Revisions Issues

The experience of the test states in revising their guidelines suggest two areas in which revisions may pose a challenge. As we noted, ease of revisions is considered an important advantage of guidelines, in contrast to legislative minima or other kinds of sentencing reforms involving determinate sanctions. However, it is clear that the ability to make revisions systematically depends on judicial compliance with the guidelines--especially the quality of reasons given for sentences outside the recommended ranges--and on adequate monitoring of guidelines use. This is the case whether we are discussing empirical or normative guidelines systems. However, we shall argue in the next chapter that the voluntary nature of most empirical guidelines efforts impedes their effectiveness in these respects.

A second issue about revisions is perhaps ironic. Were they, in fact, too easily made during the Florida and Maryland tests? Absent systematic feedback about factors, scoring and ranges, it was political pressure from test site actors that raised questions about the OUT range (Florida) and treatment of multiple counts and rape scoring (Maryland) and also suggested expedient solutions. Although the political necessity was clear, that does not mean the resulting changes were desirable; in fact, they seem to have reduced the guidelines' potential for affecting sentence disparity. While more concern about implementation might prevent such necessities from developing (as Chapter 5 will suggest), the issue of maintaining guidelines' flexibility to change while in some way limiting the vulnerability to political pressure remains.

4.4 Conclusions on Development and Design Issues

Throughout this chapter, we have used a review of the development, design and revision experiences of the Florida and Maryland projects to raise issues concerning the strengths and weaknesses of sentencing guidelines. In some areas, we have drawn a contrast between the empirical and normative approaches, although a number of issues are relevant to any guidelines effort. This section briefly

summarizes the issues and their implications for future guidelines use.

Four main arguments have been made in favor of empirically based sentencing guidelines:

- they will effectively diminish sentencing disparity, by providing ranges that do not reflect extremes of current practice but do reflect the bulk of that practice;

- they will meet with more ready acceptance than other kinds of sentencing reform, both within and outside the judiciary, because they represent an embodiment of existing, latent sentencing policy rather than a new area of policy-making or a substantive change in current practice;

- they preserve judicial prerogative (although limiting discretion at the extremes), at a time when legislative actions concerning criminal sanctions point strongly toward a reduced judicial role; and

- they need not be developed with explicit reference to impacts on prison or jail populations, because they are based on current practice which is already a factor in the allocation of resources to corrections.

The multijurisdictional sentencing guidelines test in Florida and Maryland suggests a number of problems with these arguments as they pertain to development and design.

4.4.1 Implications for Sentencing Disparity

The potential effectiveness of guidelines for diminishing sentence disparity is reduced in several ways under the empirical approach. Most important, it offers no incentive for recognizing or dealing with inter-judge or inter-jurisdictional differences. Instead, the statistical analyses are a political tool for avoiding conflict, and the likelihood of substantive policy debate is reduced.[36] Note that the separate analysis of sentencing, by judge or jurisdiction, could actually provide a basis for such debate. However, the paradigm for developing the final guidelines requires statistical modeling based on pooled data from all participant sites.

Even with empirical analyses to identify factors and scoring weights, the recommended sentence range widths and the designation of guidelines cells as IN or OUT are still policy decisions.[37] There is no statistical rule on the appropriate width of ranges or on how a predicted average incarceration rate should determine whether

a cell is labelled IN or OUT. This is always a policy exercise. As a result, empirical guidelines do not necessarily reduce disparity; even with full compliance, disparity reduction also depends on width of ranges. As we have seen, the Maryland project produced wide and overlapping ranges, while Florida's effort yielded relatively narrow consecutive ranges. In the current climate of public concern about crime, there is the further temptation to make the guidelines sentences look more severe--by eliminating the OUT cell, or by raising the upper boundaries of the sentence ranges. Changes like these also diminish the effect guidelines can have on disparity.

A final problem with respect to increasing sentencing uniformity is the absence of standards in empirical guidelines to date. There is no limit to how far a sentence may fall outside the recommended range, no rule concerning what reasons justify extra-guidelines sentences. There are also no consequences (such as appellate review) for these kinds of non-compliance. It would not be impossible to design empirical guidelines with standards, although their meaning and impact in a voluntary system would be unclear.[38] In their absence, effectiveness in reducing disparity cannot be anticipated.

By contrast to empirical guidelines, normative systems articulate an explicit policy and create sentence ranges that flow from that policy. One such policy is to maintain correctional populations within capacity through the sentencing function. This goal virtually requires narrow and sharply defined guidelines ranges. Normative systems also offer standards for the level of extra-guidelines sentences and/or the reasons for giving them. Each of these features appears to offer a better chance for effecting actual reductions in disparity.

4.4.2 Implications for Acceptance of Guidelines

The argument for the greater political acceptability of empirical guidelines has its basis in the history of the dialogue between legislature and judiciary over sentencing policy. The denial that such guidelines involve express policy-making often results from concerns about crossing the boundary between judicial and legislative authority. So does the assertion that they will bring about no departure from current practice.[39] Such arguments have, in fact, brought empirical guidelines widespread support and averted a variety of alternative legislative actions (see Chapter 2 and Appendix A).

However, the acceptance these arguments purchase for empirical guidelines is rather shallow. For example, what does it mean for the guidelines to be based on a sort of average of practices across several jurisdictions? Where there are significant local differences in norms and practices, the acceptance may last only until the guidelines are implemented and their consequences for particular cases become evident. Given their weak theoretical basis,[40] there is little to persuade judges that these guidelines should supersede their own views or the norms of their jurisdictions. Certainly, Florida's 4th Circuit judges--uncomfortable with both the sentence ranges

and the guidelines procedures--found nothing persuasive in the process by which the guidelines had been developed. Ultimately, however, the acceptance of any guidelines system depends far more on the politics of implementation than on the basis for guidelines design.

To the degree that the empirical approach does help garner acceptance, it may well backfire for those who fully believe its claims. The empirical analyses do not produce the guidelines whole, even when no technical problems arise. As we noted, neither the width of ranges nor the IN/OUT designation arise directly from the statistical techniques; policy decisions may also be made concerning choice of factors, scoring and sentence recommendations. Judges who expect a minimal policy role may be unprepared to make such decisions. It is clear that the Maryland Advisory Board's voting members, faced with a somewhat expanded policy role, were unprepared to assume it. Since they lacked any explicit policy-making mandate from their colleagues and any philosophical consensus on which to base their choices, their reaction is readily understandable.

By contrast, normative guidelines are developed by a body with an explicit policy-making mandate--in Minnesota and Washington, a sentencing commission. Because they are designed in terms of a particular philosophy, it is possible to make cogent arguments to support the choices of factors, scoring and--most important--the sentence ranges. Reasons for the difference between these ranges and current local practice are clear. While it is undoubtedly more difficult to build the political consensus for a normative guidelines effort than to win acceptance for an empirical system, the payoff in terms of continued support is likely to be significant.[41]

4.4.3 Implications for Judicial Prerogative

Sentencing guidelines do appear to preserve the prerogative of the judiciary, especially in contrast to the various legislative approaches to limiting sentencing discretion. However, the development and design process for empirically based guidelines holds the potential for a shift of authority and responsibility from the policy-making board to the technically trained staff.

It is apparent, from the account of guidelines development in Sections 4.1 and 4.2, that the Maryland Board played a much more active and decisive role than Florida's. The sheer level of involvement of the voting judges, and especially the range of decisions made by the Board, would serve to increase both their understanding and their personal investment in the project. The difference in staff-Board dynamics probably reinforced this pattern. Staff clearly played a more leading role in Florida than in Maryland. For a number of reasons (including the logistics of convening the Advisory Board, the greater technical experience of the staff, and the shortened time frame for grid development), the Florida staff took a greater part in policy decisions and at times simply looked to the Board for ratification. The staff's location and the ongoing communication with the first Board chairman--the Supreme

Court Justice who championed statewide guidelines--may also have served to strengthen staff independence and widen the gap between staff and Board.

Maryland's Board had far more substantive input to guidelines development. It is difficult to say whether the greater acceptance of the guidelines by the Maryland test site judges resulted from this or was a partial pre-condition for it, evidenced in the greater effort made by the Board members. In any event, we believe the contrast is instructive. Where the empirical nature of a guidelines development effort is the paramount consideration in staffing, it seems likely that communication between Board and staff will be more difficult, and staff will be less accepting of a board's authority to overrule technical findings in policy decisions. Although it must be emphasized that the Florida judges did exercise control over the selection and deletion of factors for guidelines scoring, the "success" of the empirical analysis in Florida set the stage for a more limited range of policy decisions.

The development of normative guidelines is also likely to involve sharing authority, but with other interested parties rather than with a hired staff. There may well be representatives of the prosecution, defense, correctional establishment, legislature and citizenry working with judges to develop sentencing policy. The choice between approaches is thus not a matter of whether responsibility for sentencing policy (within limits set by the legislature) will be shared, but with whom and toward what end.

4.4.4 Implications for Correctional Populations

The argument that empirical sentencing guidelines will not affect prison or jail populations is based more on the notion that such guidelines do not change basic current practice than on careful analyses. Indeed, it would be very difficult to test the likely effects of the systems developed in Maryland or Florida, since the ranges were not binding and there were no rules concerning how often or how far sentences could deviate from the guidelines. Still, with the technical expertise required for empirical efforts and with adequate monitoring data on extra-guidelines sanctions, it would seem that such analyses could be conducted.

The claim for neutrality of empirical guidelines vis-a-vis correctional populations does, however, ignore one important problem. Differences between present local practice and the statistical composite embodied in the guidelines could certainly have consequences for local jails, if local judicial compliance with the guidelines was strong. For example, the effort to establish statewide guidelines in Pennsylvania foundered in part because it became clear that the proposed sentencing ranges influenced by the more severe standards of suburban and rural jurisdictions and the general "get tough" climate, would sharply increase the jail populations in Philadelphia and Pittsburgh.[42]

The mandate to limit impacts on prison and jail populations plays an

important role in the development of normative guidelines. In Minnesota, it effectively constrained the options for the sentencing commission without predetermining the nature of the guidelines. It also aided the policy-makers in dealing with the demands of some interest groups, especially demands for more severe sanctions.[43]

The issue of correctional population impacts under empirical sentencing guidelines may well remain in the background, if such guidelines continue to be implemented with judicial compliance voluntary rather than mandatory. The multi-jurisdictional test experience seems to indicate that compliance is lowest where the difference between local practice and the guidelines is greatest. Under the circumstances, changes in incarceration rates--while not impossible--are not likely to occur as a consequence of guidelines implementation.

FOOTNOTES

1. William D. Rich, L. Paul Sutton, Todd R. Clear and Michael J. Saks, Sentencing by Mathematics: An Evaluation of the Early Attempts to Develop and Implement Sentencing Guidelines (Williamsburg, VA: National Center for State Courts, 1982).

2. Chapter 5 focuses on the effects of the voluntary nature of the multijurisdictional field test guidelines.

3. Minnesota's approach is described in Preliminary Report on the Development and Impact of the Minnesota Sentencing Guidelines (St. Paul, MN: Minnesota Sentencing Guidelines Commission, 1982).

4. The Washington guidelines are described in Sentencing Guidelines Commission: Report to the Legislature (Olympia, WA: Sentencing Guidelines Commission, January 1983).

5. The Board ratified this decision long after it was implemented. Staff members reported that they could not get the Board together to make the decision, so they went ahead. In the end, the Board was pleased.

6. Multijurisdictional Sentencing Guidelines Project: Final Report (Tallahassee, FL: Office of the State Courts Administrator, July 1982), p. 5.

7. The full list of offenses in each category is shown with the guidelines grids (see Appendix D).

8. It is interesting to note that a somewhat similar division was used by the earlier Sentencing Study Committee in Florida: crimes against persons, crimes against property, and public order/victimless crimes.

9. "ABA pleas" are authorized under Rule 733 of the Maryland Rules of Procedure. According to this Rule, the prosecutor may present plea agreements which include a specific sentence or other judicial action as part of the bargain. Once the judge approves such an agreement, he or she is bound by its terms and may not impose a sentence more severe than that contemplated in the bargain. However, if the judge rejects the agreement, the defendant is allowed to withdraw the guilty plea and may request that the case be tried before another judge. Non-ABA pleas include those based on charge bargains and those where the judge is not explicitly bound as to sentence.

10. Florida Final Report, p.4.

11. The two rural circuits in Florida (the 10th and the 14th) were combined for purposes of sampling and analysis.

12. Staff presented the Advisory Board with copies of pre- and post-sentence investigations and asked it to determine if part or all of the information in the post-sentence investigation was available to the judge in cases where a pre-sentence investigation had not been completed. Several Board members argued strongly that in many cases, especially plea-negotiated cases, factors and circumstances that were not found on any written document affected the disposition of the case and, thus, the information on the post-sentence investigation had little or nothing to do with the sentence outcome. A majority of the Board members, however, agreed that judges are made aware of the types of information on the post-sentence investigation by discussions with the prosecutor, defense attorney, and others. The Board voted that in the absence of pre-sentence investigations, staff should use post-sentence investigations as the primary source of data.

13. See Rich et al., Sentencing by Mathematics, Chapter 2.

14. Among the factors deleted were: method of adjudication (plea or trial); offense at arraignment rather than conviction; and pending and unverified prior offense behavior, (Florida Final Report, pp. 12-14).

15. Florida Final Report, p. 15.

16. TOBIT analysis was used to derive the point scores, because of the non-normal distribution of the dependent variable (large numbers of zero values, distribution truncated at zero).

17. Weighting for analysis can be a separate step from weighting to reflect differences in sampling fractions (sample design).

18. They were the seriousness of the instant count, the most serious physical injury suffered by a personal contact victim, and the offender's legal status at the time of the crime.

19. The public defender representative (ex officio) to the Board did object to treating all the crimes in the category the same, even though some were second and others third degree felonies. Implicitly, this asked for separate offense scaling. However, none of the voting members seemed concerned about this.

20. The argument was made that weight changes would lead to increasing the prison population. Despite the judges' view that this was not their problem, and despite their avowed disinclination to accept the statistical analysis as "incontestible," the only weight changes they approved concerned prior adult record.

21. Only one of Florida's six scoresheets and grids (Category 2) is shown; formats did not differ for the other categories. The full set can be seen in Appendix D, which also includes the scoresheets after revisions (see Section 4.3). Appendix E contains Maryland's original and revised scoresheets and grids.

22. The Preliminary Report on the Minnesota guidelines presents four hypothetical grids corresponding to different sentencing philosophies. For example, they suggest that utilitarian sentencing goals (e.g. deterrence, rehabilitation) are served by guidelines which weight criminal history more heavily than current offense in terms of incarceration, while a heavier emphasis on the severity of the convicted offense relative to prior record embodies a goal of retribution or "just desserts." (See the Preliminary Report on the Minnesota Sentencing Guidelines, p. 11, for the hypothetical grids.)

23. We are referring to the debate over whether psychological damage can also be permanent (viz. rape).

24. Richard F. Sparks, Bridget A. Stecher, Jay S. Albanese, and Peggy L. Shelly, with a chapter by Donald M. Barry, Stumbling Toward Justice: Some Overlooked Research and Policy Questions about Statewide Sentencing Guidelines (Newark, N.J.: School of Criminal Justice, Rutgers University, 1982), p. 552.

25. Sparks et al., Stumbling Toward Justice, p. 450.

26. Later modification extended the scoring to all current offenses.

27. Florida also developed a supplementary worksheet for the calculation of prior criminal record points and for comments on the work habits and drug or alcohol use of the offender.

28. The impact and compliance analyses presented in this report confirm the existence of substantial differences; see Chapters 7 and 8.

29. Other contrasts between judically and legislatively mandated guidelines are examined in Chapter 5.

30. Analysis of these problems is presented in Chapter 7, where there is also a discussion of the reasons that were articulated by the judges during the test year.

31. In materials distributed to the Advisory Board for its Februay 19-20, 1981 meeting, staff provided information on the distribution of construction sample cases against the proposed guidelines grids. These materials show that from 17 to 78 percent of the cases fell in the OUT cell (varying among the grids). The median was 55 percent (excluding Category 3, for which the score intervals were changed before the grid was approved). If the OUT cell were to be combined with the second lowest cell, these figures would range from 20 to 83 percent, with a median of 71 percent.

32. The actual distributions from test year scoresheets are shown in Chapter 7, Table 8-5.

33. Sparks et al., Stumbling Toward Justice, p. 90.

34. These problems are analyzed in Chapter 7.

35. Maryland Sentencing Guidelines Project memorandum, March 17, 1982.

36. Jolene Galegher and John S. Carroll, "Voluntary Sentencing Guidelines: Prescription for Justice or Patent Medicine?" Law and Human Behavior 7 (1983): 361-400.

37. Sparks et al., Stumbling Toward Justice, p. 88.

38. Chapter 5 expands this argument.

39. Sparks et al., Stumbling Toward Justice, p. 299.

40. Rich et al., in Sentencing by Mathematics, make a variety of philosophical and methodological arguments against interpreting a pooled statistical model as an embodiment of "latent policy." See especially pp. 25-30.

41. Chapter 5 focuses further on support-building beyond the judiciary and legislature.

42. Susan E. Martin, "The Politics of Sentencing Reform: Sentencing Guidelines in Pennsylvania and Minnesota," in Research on Sentencing: The Search for Reform, Volume II, eds. Alfred Blumstein, Jacqueline Cohen, Susan E. Martin, and Michael H. Tonry (Washington, D.C.: National Academy Press, 1983).

43. Martin, "The Politics of Sentencing Reform," pp. 281-282. See also Andrew von Hirsch, "Constructing Guidelines for Sentencing: The Critical Choices for the Minnesota Sentencing Guidelines Commission," Hamline Law Review 5 (June 1982): 176-180.

CHAPTER 5

THE POLITICS AND PROCESS OF IMPLEMENTATION

As the preceding chapter has noted, the restriction of the test to empirical guidelines shaped much of the design activity in the two sites. Clearly, this choice was grounded in the historical experience with the pilot implementation studies conducted by the Albany CJRC. Similarly, the guidelines implementation process was intended to parallel the experience of the early pilot projects, although on a larger scale. It was anticipated that the guidelines would be developed by and for the judiciary, and would be implemented through some form of judicial mandate, whether that be court rule, decree of the Chief Justice of the Supreme Court, or by formal agreement of the participating judges.[1] The varying degrees of formality (and authority) inherent in the implementation mechanisms suggested in the test reinforce the philosophy of the early guidelines research, in which the development of the guidelines was viewed as a consensual exercise among the judiciary. Both states further emphasized this perspective by establishing their systems through the less restrictive implementation mechanisms, essentially creating <u>voluntary</u> guidelines. At the point of implementation, this feature of the test design was perhaps the most influential issue shaping the guidelines project and the operation of the guidelines in the participating jurisdictions.

This chapter considers the process and results of three key implementation issues. All are important to any guidelines endeavor; however, they are particularly crucial in a voluntary system:

- Building support;
- Accommodating local procedures and norms; and
- Monitoring compliance.

The chapter closes by considering the overall consequences of selecting voluntary rather than mandatory guidelines.

To provide a context for these discussions, we begin with a brief overview of the factors involved in the states' initial choice to participate in sentencing guidelines--particularly the guidelines approach envisioned in the multijurisdictional field test.

5.1 The Choice of Guidelines

Sentencing issues were a matter of substantial interest, and even controversy, in both states well before the consideration of guidelines. The legislatures in Florida and Maryland actively considered sentence reforms, especially mandatory minimum sentences and determinate or presumptive sentencing schemes. In both states, legislation was passed providing minimum penalties for certain specific offenses. In addition, bills calling for presumptive sentencing were introduced in both state legislatures. While such measures met with no success in Maryland, the Florida legislature actually passed presumptive sentencing legislation in its 1978 session, although this bill was later vetoed by the Governor.

Spurred in part by the high level of legislative interest, the judiciaries of Florida and Maryland undertook significant efforts to examine the "state of sentencing" in their own courts and to explore possible responses to problems in sentencing. For example, in January 1978 the Florida Supreme Court appointed a Sentencing Study Committee to investigate issues of sentencing disparity and sentencing reform. The 21-person Committee included members of the judiciary from the county, circuit, appellate, and Supreme courts, plus legislators, a member of the private bar, a professor of law, and representatives of the state attorneys and public defenders. Activities of the Committee included a review of the literature on sentencing reform and an examination of Florida sentencing practices. The Committee was responsible for overseeing a federal grant to examine sentencing practices through an empirical analysis of data on approximately 1000 cases from the 20 judicial circuits of Florida. The Sentencing Study Committee was an instrumental force in Florida's decision to experiment with sentencing guidelines. In its interim report to the legislature, the Committee recommended "the development and implementation of structured sentencing guidelines in combination with a sentence review panel that would operate within the sentence parameters prescribed by the legislature."[2] In addition, it was largely due to the efforts of the Florida Sentencing Study Committee that the 1978 determinate sentencing legislation failed to become law: the Governor vetoed the measure pending the results of the Committee's investigations.

The Maryland judiciary also demonstrated a longstanding concern with sentencing issues before the multijurisdictional guidelines field test. In the mid-1970s, the Administrative Office of the Courts developed special sentencing programs in response to judges' requests for instruction on this topic. Presented as part of the state's judicial education seminars, these programs were attended by all judges in the state. In May 1978, the Chief Judge of the Maryland Court of Appeals established an ad hoc Sentencing Study Committee as part of the State's Judicial Conference. It was charged with the responsibility "to review recent developments in sentencing in the United States, to study the major proposals for change, to consider sentencing practices in Maryland and to report its recommendations for change to the Judicial

Conference."[3] Based on the results of these studies, the Committee recommended to the Judicial Conference in March 1979 that the state develop sentencing guidelines, and it urged "the establishment of appellate sentence review, of right, for all imposed sentences not indicated by applicable sentencing guidelines."[4] However, by this time the Maryland Administrative Office of the Courts had initiated the application process for the multijurisdictional sentencing guidelines field test. The Judicial Conference therefore approved a resolution supporting the development of sentencing guidelines, but deferred consideration of their adoption on a statewide level until completion of the field test.

Concern with sentencing issues and the possibility of legislative action was not the only force operating in the states' decisions to proceed with sentencing guidelines. Given that parole agencies, rather than judges, exercised primary control over duration of incarceration in Maryland and Florida, it is not surprising that parole policies--and even the existence of parole--were also significant issues.

The Florida Parole Commission, composed of eight members versed in the fields of criminal justice, corrections, and social science, was given primary responsibility for determining the release date for the majority of incarcerated offenders. Prior to 1978, release decisions were made by the Parole Commission on the basis of parole interviews and information presented at parole hearings.[5] There was substantial public, judicial, and legislative concern about these decisions, from the point of view of both leniency and consistency. As a result, in 1978 the Florida Legislature enacted a statute requiring the Commission to develop "objective parole guidelines" and to implement these guidelines by January 1, 1979. For the most part, the guidelines and the methodology used to develop them resemble those of Gottfredson and Wilkins.[6] The guidelines used a two-dimensional matrix, with one axis presenting "salient factors" that indicate the offender's probable success on parole, while the other axis lists and ranks offenses according to several severity categories.[7]

As stipulated by the Legislature, the parole guidelines were formally implemented on January 1, 1979, only nine months before the initiation of the Sentencing Guidelines Project. During this period, public clamor concerning parole continued, while judges voiced increasing concern about parole release in general and the parole guidelines in particular. While the general concerns still centered around the leniency of the parole guidelines, judges also noted with frustration that the guidelines did not consider the sentence imposed by the judge as one of the factors in deriving the presumptive parole release date.[8] A common complaint voiced by judges, prosecutors, and even some defense attorneys was that sentencing was "meaningless" in the context of parole. This, in fact, was often suggested as the major problem in sentencing in the state. In part, this concern served as a disincentive to adopt sentencing guidelines, since sentence reform was viewed as a meaningless exercise in light of parole. However, others looked to the sentencing guidelines as a possible

solution to the parole question, reasoning that increased consistency in sentencing could eliminate one justification for the parole system.

In many respects, the Maryland parole system suffered similar criticisms and pressures. In 1976, the Maryland General Assembly created a seven member Parole Commission, charging it with the responsibility for authorizing parole releases for incarcerated offenders. Parole could be authorized at any time after incarceration, and the Commissioners were required only to base their decisions on the circumstances of the crime; the physical, mental, and moral traits of the inmate; institutional behavior; probable success on parole; and the welfare of society.[9] However, as in Florida, concern about sentence disparity and inconsistency in parole policies prompted the Maryland authorities to consider parole guidelines. Unlike the Florida parole guidelines, Maryland parole guidelines were developed as an administrative policy rather than by legislative mandate. In 1979, the Parole Commission first adopted its "guidelines for parole consideration," which incorporated information on the instant offense and the prior criminal history of the offender.

As in Florida, the Maryland judiciary expressed considerable concern over the effect of parole on their sentences, often noting that parole was granted after only one-fifth of the sentence was served. Many judges and other criminal justice system officials readily conceded that parole was a significant consideration in judges' sentencing calculus. In addition, judges were concerned about the predictability of parole relase; in fact, the 1979 report by the Committee on Sentencing of the Maryland Judicial Conference had recommended that the Parole Commission establish and publish "guidelines for release of Maryland prisoners on parole."[10] Still, the degree of concern, and indeed rancor, expressed by the Florida judiciary did not find a parallel in Maryland.

In sum, it seems reasonable to suggest that interest in sentencing guidelines in the Maryland and Florida courts stemmed from at least two sources. On the one hand, it would appear that guidelines offered the hope of strengthening judicial sentencing authority and possibly eroding the rationale for a parole system. On the other hand, it seems clear that the judiciary's interest in sentence reform--and guidelines in particular--may have been in part a response to legislative initiatives. In Florida, the prospect of determinate sentencing was a real threat recognized by the judiciary; during our interviews with judges in the Florida test jurisdictions, many expressed the conviction that without a judicial initiative in sentence reform, determinate sentencing would be legislated. Even in Maryland, where the possibility of such measures seemed less immediate, there was continued concern about presumptive sentencing. This tension is evident in the report of the Maryland Committee on Sentencing, which, we believe, expresses the beliefs of judges in both states. The report noted that "the Committee on Sentencing was formed in response to judicial concern about sentencing, but also in recognition of the concern others have about this central judicial responsibility, and to ensure that the Maryland

Judiciary will be prepared to take an active part in determining any changes which are made in sentencing laws and practices. If the Judiciary does not respond responsibly to widespread dissatisfaction with the present systems of sentencing and punishment, others less qualified may initiate changes which are not adequately informed by judicial knowledge and experience."[11]

5.2 The Choice of Judicial versus Legislative Authority

In both Florida and Maryland, the choice of guidelines systems based on judicial rather than legislative mandates might easily be attributed to the same desire to avoid any legislative dictates. If both states had turned to guidelines to counter the prospect of legislated change, they were hardly likely to incorporate a strong legislative role in the implementation of guidelines.

In fact, the choice was never a matter of significant debate. As noted above, the test design reflected the original guidelines concept, asking that the experiment be judicially initiated and controlled. This provision grew from the argument that a guidelines system established by the judiciary will be more effective than reforms mandated by a legislature. According to Wilkins et al.,[12]

> When comparing sentencing guidelines to legislatively mandated sentencing proposals, the most striking positive practical attribute of the guideline system is that it is judicially implemented and judicially controlled. Governmental change is at best a slow process in which overt hostility and resentment or at least passive resistance, can be expected to result from forced change. In a judicially developed and controlled guideline system, however, sentences need not be specifically prescribed by any outside body. This is especially important if one recognizes sentencing to be a legitimate judicial function. When change takes place under the direction of those whose present authority and responsibilities are to be directly affected by its enactment, then future acceptance of it is more likely to be relatively problem-free. The use of sentencing guidelines should lead to less circumvention because it is the existing policies of the court itself that are initially being made explicit.

In practice, the assumptions of the early guidelines efforts about judicial support--and its corollary, judicial compliance--seem to have been overstated. For example, the fact that judges themselves developed guidelines in Phoenix, Denver, and Philadelphia did not appear to create sufficient inducement to ensure their application in all relevant cases.[13] The same can be said about Maryland and Florida, where compliance in some of the test sites was far below expectations.

Is this the fault of judicially-initiated guidelines? Not necessarily. For example, in their original recommendations, both the Maryland and Florida Sentencing Study Commissions called for guidelines in conjunction with sentence appeals. Indeed, had not the test design (with its foundation in voluntary guidelines) been instituted in each state, it is possible that both would have proceeded by other means to implement guidelines with some form of appeal. Clearly, to include such a provision, judicially developed guidelines must have strong implementing authority--perhaps court rule. Just as clearly, an appeals procedure would have done much to ensure the accuracy of the guidelines scoring and compliance with the guidelines requirements.[14]

At the time the multijurisdictional field test was instituted, voluntary, judicially based guidelines were the only available model. However, soon after the start of the test, other states began to experiment with variations on the basic guidelines theme--most notably, turning to legislative authorization for the guidelines and mandatory systems. Over the years, it has become apparent that these approaches offer some benefits not automatically available under the original guidelines concept.

Much of the initial concern about legislative involvement in sentencing reform arose in the context of determinate sentencing. Most commonly, it was argued that legislatively determined sentences cannot hope to account for the legitimate sources of variation in sentencing, and that injustice would occur if the same determinate sentence were applied indiscriminately to dissimilar offenders committing the "same" crime. More apropos to sentencing guidelines, it was noted that legislatively established sentences are too often subject to political pressures, reflecting in their severity a symbolic statement of concern about crime rather than a rational system of sanctioning policy.[15] Through their sentencing guidelines development efforts, a number of states--notably Minnesota, Pennsylvania, and more recently, Washington--have demonstrated that legislative involvement can take a very different form than that suggested by the earlier determinate sentencing experience. In these states, legislatively mandated sentence commissions have been empowered to develop sentencing guidelines which are then implemented by law.

Such an approach offers several potential benefits. For example, the commissions established under legislative authority are given a firm, broadly based mandate to proceed with guidelines development. Backed by the legislature's constitutional authority to set penalties for crimes, this can be a potent force in ensuring that the guidelines are taken for what they truly are--development of new sentencing policy. In addition, with legislative authorization for the guidelines, enforcement of the guidelines can be facilitated. Minnesota offers the most suggestive example in this regard: the guidelines sentence ranges are presumptive rather than advisory, and sentences outside the guidelines may be appealed by the defense or prosecution. To date, such strict enforcement provisions have been developed only under legislative mandate.

In clear contrast, the guidelines developed and tested in Maryland and Florida were voluntary in nature. In Florida, the guidelines were implemented by an Administrative Order of the Supreme Court, while in Maryland, implementation was left to the administrative judges in each participating jurisdiction. In this experimental effort, formal methods were not developed to compel judges to consider the guidelines, sentence within the recommended ranges, or to explain the reasons for extra-guidelines sentences; agreement of the participating jurisdictions to comply with these terms was the primary means of ensuring cooperation. Under such circumstances, activities designed to build support, accommodate local norms, and monitor guidelines usage assume considerable importance for any meaningful test of the guidelines. The sites' efforts in these arenas are examined in the following sections.

5.3 Building Support

At least three groups can be identified as key participants in a sentencing guidelines effort as it will routinely operate. They are the judiciary--whose decisionmaking is meant to be affected by the reform--the prosecution, and the defense (both public and private bar). Because the locus of discretion for sentencing can shift between prosecution and judge, and since there is an intricate and variable working relationship and balance among the three parties, the involvement or at least assent of all three is required to change the way that criminal cases are disposed.[16]

Since sentencing guidelines were judicially initiated in both Maryland and Florida, a number of questions arise concerning the participation of these three groups. How broad was judicial awareness of guidelines in the participating jurisdictions? How prepared were the Advisory Board judges to serve as "change agents" in their circuits? And how was the cooperation of the defense and prosecution secured?

In addition, other groups less directly involved with the guidelines may nevertheless have a substantial interest in their development and implementation. For example, the legislature retains its authority to modify sentencing statutes--a power which may obviate the courts' efforts to reform sentencing on their own initative. Similarly, the division of parole and probation serves a crucial role in providing information and assistance under most guidelines efforts, while the activities of the parole commission can do much to support or detract from the effectiveness of judicially developed guidelines. Thus, the support and understanding of these groups are especially important in the multijurisdictional guidelines effort--particularly since judicially mandated guidelines have little or no inherent authority to require the cooperation of these groups.

5.3.1 The Role of Advisors

In Chapter 4, we noted the role of the Advisory Board in supervising the data collection and analysis phase of guidelines development and providing policy input for the design of empirically-based guidelines. However, the success of the guidelines can also depend on the Board's assumption of other potential roles: providing a forum for education about about the guidelines objectives and benefits, developing some degree of investment in the project by participants, and building support for the guidelines. Both the composition of the Boards and Board members' activities influenced the states' ability to achieve these aims within each of the major groups concerned with the guidelines development: the judiciary, defense and prosecution, and external agencies.

The Judiciary

Recognizing the multijurisdictional sentencing guidelines as a judicially-initiated reform, the test design document placed particular importance on the judicial representation on the Advisory Board. However, in developing their board membership, the two states involved in the test took very different approaches.

In Florida, voting members of the Advisory Board included only the Chief Judges of the four participating circuits.[17] In addition, a Justice of the Florida Supreme Court was appointed to chair the Board and serve as a non-voting member except in the case of a tie. Originally, the Florida project had planned to include several judicial representatives from each site on the Board, in numbers roughly proportional to the caseloads of the circuits. Thus, it was suggested that the urban circuit (4th) would have three voting members; the suburban circuit (15th) would have two members, and the rural sites (10th and 14th) would each have one representative.[18] During the initial meeting of the Advisory Board, however, it was decided instead that each site would be given equal representation, and that only one judge from each circuit would participate.

Maryland established a different approach to site representation. In developing the Advisory Board, the Administrative Judge of each county was asked to nominate appropriate Board members. Representation of the sites was roughly proportional to population; Baltimore had five representatives, Montgomery and Prince George's Counties had two each, and one judge represented Harford County. One of the Baltimore City representatives was designated as Chairman of the Advisory Board.

In terms of providing policy input for guidelines development, either approach could prove to be workable. How, then, did these structures serve the other goals of support building, education, and investment?

Clearly, broad judicial representation offers a number of potential advantages. For example, participation in the Advisory Board in and of itself is likely to generate support and understanding. The extent to which this translates to the judiciary in the participating sites is in part a function of the number of judges on the Board. Maryland's choice--to provide proportional representation of circuit judges--meant that fully 22 percent of the judiciary in its urban jurisdiction were included on the Board. In contrast, only one of the 23 judges (4 percent) from the urban site in Florida was represented. Given the special implementation problems presented by the urban courts, it is not unreasonable to suggest that the contrast in judicial support between Baltimore and Jacksonville may be traced in part to the representation provided on the Board.

Given practical constraints on Board size, however, direct proportionality of representation will be infeasible in any large guidelines development effort. What, then, is the alternative? In large measure, the answer lies in the activities and duties of the Board. The test design--and the sites--clearly focused on the Board's role in guidelines development. Although Wilkins had noted the benefits of judicial participation in building support (or at least minimizing resistance) in the courts,[19] this was viewed primarily as a natural by-product of the design effort. Particularly in Florida, it is not clear that the Board's role as a change agent was ever made explicit. The challenge presented to members appeared to be the development and implementation of a new administrative procedure, rather than the institution of a very significant change in sentencing policy.[20] This view was reinforced by the choice of judicial representatives: as chief judges of the participating circuits, members were vested with the administrative responsibility for their circuits. In addition, the limited number of Advisory Board meetings and the tenor of those meetings (focusing on guidelines development rather than implementation) may have further solidified this perception. The degree to which members could and did adopt the role of change agents on their own initiative varied among the sites.

From the start, the Maryland judiciary viewed the guidelines effort as a more political undertaking. The Board members were extensively involved in the policy decisions for guidelines development, and met frequently and regularly. By virtue of this involvement, it appears that they were well versed in the guidelines development process and could serve as knowledgeable spokespersons for the guidelines within their jurisdictions. They were also more involved in implementation discussions than their Florida counterparts, which may have served to heighten their awareness of their own role in that process. Finally, vested with no special authority among their circuits, the judges of the Advisory Board of necessity had to rely on developing political support within their jurisdictions to ensure the implementation of the guidelines.

The Defense and Prosecution

By suggesting the inclusion of ex-officio Board members, the test design clearly recognized the need to develop political relations with other key groups. Of these, the defense and prosecution are perhaps the most important with respect to the daily logistics of guidelines operation.

In many respects, Florida and Maryland took similar approaches to attorney representation on their Advisory Boards. Each initially chose to have one representative for the prosecution, and one public defender. The Florida Board drew its representatives from individual circuits. The prosecutor served in the Fourth Judicial Circuit, while the public defender served in the Fourteenth. In addition, a private attorney practicing in the Miami area was added to the Board. However, some few months into the project the State Attorney from Jacksonville resigned and was replaced by the State Attorney from the Sixth Circuit, which was not one of the test sites. In Maryland, at least initially, the defense and prosecution were also limited to one ex-officio member each. However, unlike Florida, these members represented central state offices rather than defense or prosecutorial agencies in the test sites themselves.

As we will see in Section 5.3, both states clearly underestimated the importance of enlisting the support and understanding of the prosecution and defense, with the result that guidelines implementation was hampered by overt resistance and, in some cases, subversion through plea negotiations. Since the Advisory Board offers a major means of building support, both states could have used the Board to better advantage with these groups.

As in the case of the judiciary, it could be argued that greater representation of the defense and prosecution could have helped to enlist the support of attorneys in the test sites. However, the test design mandate for judicially-initiated guidelines rendered such an approach unlikely at best. In addition, inclusion of members from each site would have unduly increased the size of the Board. Once again, the solution lies with the nature of members' participation rather than the number of participants.

One obvious concern is the "legitimacy" of these representatives--the extent to which prosecutors and defense attorneys in the sites would agree that these were "their" representatives. Since the selection of these ex-officio members was not brought before either the prosecution or the defense in either state, it is not surprising that many attorneys in the sites felt that they had little input or representation on the Board. This was particularly true of the prosecution in both states. While Maryland obviously attempted to deal with this issue by selecting representatives from central agencies, prosecutors in the sites disagreed strongly with the Maryland guidelines project's choice to invite the State's Attorney Coordinator. Ultimately, after the implementation of guidelines and under pressure from the State's Attorneys, voting members of the Maryland Board agreed to add a representative chosen by the prosecutors themselves.

A second concern is the Board members' perception of their role. Like the judiciary, the prosecution and defense focused on guidelines development issues, especially in Florida. For example, while the Florida representatives participated very actively in the development process and worked hard to provide the Board with the perspectives of their constituencies, it seems doubtful that their role as "change agents" or even as educators was stressed. On the other hand, the Maryland representatives--particularly the prosecutor--participated in the meetings and actively tried to disseminate information to their counterparts in the sites. However, isolation from the sites and (for the prosecutor) the question of legitimacy served to undermine the education function.

Although it was not directly raised as a source of concern in the test sites, the non-voting status of the ex-officio members may also have undercut their effectiveness in building support for the guidelines. Certainly representation on the Board was important, but without the possibility to shape decisions, prosecutors and defenders may have felt--quite legitimately--that the representation was less than totally effective. The possibility of according voting status was probably non-existent under the multijurisdictional field test (and small under any judicially-initiated, voluntary guidelines effort). Nevertheless, given the difficulty (and necessity) of obtaining the cooperation of the defense and prosecution, voting status may prove to be a potent force in future guidelines efforts.

External Support

Other ex-officio members suggested in the test design included judges from non-participating jurisdictions, court administration personnel, correctional representatives from the participating jurisdictions, citizen representatives, and a consulting expert in research methodology.[21] To a greater or lesser degree both states followed these general guidelines. Florida, for example, included a parole commissioner, a representative of the circuit judges conference, a state senator, a state representative, and a criminologist. The Maryland Advisory Board had a similar membership, including a representative of the Department of Public Safety and Correctional Services, the Division of Parole and Probation, the Maryland House of Delegates, the Maryland State Senate, and a Legislative Officer of the Governor's Office. In addition, the Chairman of the Maryland Parole Commissions was later asked to join the Board.

It is through the selection of these ex-officio members that the sites demonstrated their greatest concern with building political support for the guidelines. In light of their positions, the question of "legitimacy" was never an issue with these ex-officio members. The Board secured the involvement of the top administrator or key representative of each group. In addition, by virtue of their positions, the dual roles as policy advisors and spokespersons for the guidelines among their own agencies was perhaps more clear.

Although the sites followed the test design's recommendation to include ex-officio members, some of the agencies represented were clearly not suggested in the design document. For example, both states included representatives of the legislature, no doubt in recognition of the pressures for sentencing reform generated by that group. In Maryland, this proved to be quite successful. The representative of the House played a very active role in the Board deliberations, frequently offering comments on the boundary between judicial and legislative prerogatives or assessing likely legislative reactions to Board decisions. In addition, judicial representatives on the Board maintained an ongoing interest in legislative concerns, and took special care to ensure that the legislature was informed of their activities.

While the Florida Board also included legislative representatives, attendance of these members was rare. However, the legislature was highly aware of the activities of the sentencing guidelines project, perhaps because its 1978 determinate sentencing bill had been vetoed in order to allow the judiciary to examine other sentence reform measures. In fact, the legislature demonstrated such interest in the guidelines concept that there was some concern that it might move prematurely to implement statewide guidelines before the multijurisdictional guidelines were developed. In general, it appears that less formal contacts than Advisory Board meetings werer used to maintain legislative contact, but that this function was carried out quite thoroughly.

The test design document was also silent on the issue of involving the Parole Commission. Still, both states chose to have a representative of this body on their Boards. Given the high level of concern with parole and correctional issues in both states, the choice of such representation is not surprising. However, it was only in Maryland that moves were made to foster a cooperative and long-term relationship between the sentencing guidelines project and the Parole Commission. The Maryland Board established a special committee to work with the Chairman of the Parole Commission to examine the relationship of the parole and sentencing guidelines. In February 1981, the two bodies reached an agreement, with the Parole Commission willing to modify its guidelines to parallel the sentencing guidelines; in addition, judges' reasons for extra-guidelines sentences would be taken under consideration in parole release decisions. The administrative origins of the Maryland parole guidelines no doubt made such modifications feasible. The relationship of the Florida parole guidelines and sentencing guidelines was also a matter of discussion at the Board meetings. However, the tone was far more negative, with many expressing the opinion that parole guidelines would not be necessary once sentencing guidelines were established.

Finally, the test design suggested that a citizen representative be included on the Board, although neither state chose to exercise this alternative. Given that both states expressed concern about public and legislative reaction to the guidelines, this decision may seem surprising. Surely strong political support for the guidelines by an

influential private citizen could have helped to secure a favorable reaction.[22] Again, however, the emphasis on the Boards' role in guidelines development (rather than policysetting) and the focus on sentencing as a judicial prerogative might have made the case for citizen representation less compelling.

This is not to say that neither state made any effort to reach the public. However, their efforts were (perhaps correctly) focused on the media. In Maryland, these contacts were not carried out directly through the medium of the Advisory Board. Florida, on the other hand, used the Board meetings themselves to obtain publicity for the guidelines, opening the meetings to the press on a number of occasions.

Clearly the Advisory Board offers one of the most important forums for building support for the guidelines. However, it is far from the only means available. We turn now to another consideration in support-building: information. While ensuring that all parties are aware of the forthcoming change can do little in itself to build support, the converse is not true--being "caught by surprise" will almost certainly engender resistance, as the experience of the two test sites demonstrates.

5.3.2 Outreach

In Florida, early contacts with the four test circuits were relatively limited. Aside from the initial recruitment efforts before the award of the Multijurisdictional Sentencing Guidelines grant, there was little "public relations" activity carried out in the test jurisdictions. For example, at the time of our initial round of interviews with judges, states attorneys, and public defenders in the winter and spring of 1980, it was clear that most knew little or nothing about the specific plans for guidelines development and implementation, and that many viewed the idea of guidelines with minimal enthusiasm. Although the staff made some efforts to brief site participants on the guidelines effort during the development phase, it was not until the guidelines training (discussed below) that any extensive site contacts were made. Thus, at the point of implementation, the Florida staff acknowledged that it was "surprising, . . . that so many individuals had virtually no conception . . . that they were about to participate in the project . . . It was obvious from the start of the [training] sessions that the majority of the personnel had never heard of the guidelines . . ."[23]

While personal contact with key actors in the test sites was limited, Florida staff made good use of written materials and the press. For example, staff initiated a newsletter to inform participants about the guidelines development. Incorporated in the Judicial Forum, a general newsletter of the State Court Administrator's Office, the first article appeared in April 1980, announcing the grant award to the Florida project and reviewing the Advisory Board membership. Later issues of the newsletter featured the preliminary results of the data collection and an overview of the guidelines developed for Flroida. In addition, staff and Advisory Board members were often interviewed by the press of the test circuits, and a number of articles appeared in such professional journals as the Florida Bar News and the Florida Bar Journal.

The Maryland project took a different approach, demonstrating considerable concern with outreach and education from the start. During the first meeting of the Advisory Board in February 1980, judges discussed ways of developing relations with the legislature, the Parole Commission, the Governor's office, and the general judiciary. Throughout the life of the project, staff and Advisory Board members made numerous presentations to such groups as local bar associations, the circuit court judges conference, the State's Attorneys conference, the legislature, judicial education seminars, and local meetings of the judiciary in each site. As in Florida, the project also benefitted from the attention of the media and placed articles in relevant professional journals such as the Maryland Prosecutor. Finally, at an early stage in the project's history staff developed a newsletter that was widely distributed to the judiciary, state's attorneys, the defense, and probation and parole personnel. Several issues were distributed during the guidelines development phase, providing updates on the project membership and summaries of the guideline project activities.

5.3.3 Training

During the period immediately before the implementation of the guidelines, staff were also responsible for developing and carrying out appropriate training programs to support the guidelines implementation. The test design document mentions only a training session for the judiciary, noting that the "purpose of this presentation is essentially informational, that is, to introduce the guidelines and their supporting documentation to the judges."[24] However, both Florida and Maryland projects ultimately developed far more extensive training programs, designed to reach the judiciary, defense and prosecuting attorneys, probation and parole staff, and other court personnel.

In Florida, preparation for the training did not begin in earnest until after the guidelines models were finalized by the Advisory Board, approximately one month before the implementation date. Staff began preparation of the guidelines documentation, including the sentencing grids, scoresheets, sentencing guidelines worksheets, and a guidelines manual intended for distribution to judges, attorneys, and probation personnel. The manual contained a brief overview of the project history and methodology, definitions of terms, and a description of the scoring procedures for the guidelines. In addition, the manual contained the grids, scoresheets, and worksheets for each crime category and a sample case and scoring for each.

Training sessions in the sites began approximately three weeks before the scheduled implementation of guidelines, with the first sessions being held in one of the rural sites. Separate sessions were generally held for the three major groups involved in the guidelines: judges, attorneys, and probation personnel. At the start of the session, staff distributed copies of the guidelines manual, and the training itself closely paralleled the contents of the manual. Thus, the sessions opened with a summary of the project and the procedures used to develop the guidelines. This was

followed by a brief overview of the scoring process, during which participants often raised particular concerns about definitions or questioned the implications of the guidelines for criminal case processing. Finally, staff reviewed several of the examples included in the manual and answered questions concerning the use of the guidelines. While the presentation for all parties was basically similar, training for the probation staff tended to emphasize the more technical aspects of guidelines preparation and scoring, while sessions for judges and attorneys were more general and often involved questions of policy or practical concerns about the impact of the guidelines on each group's typical procedures.

The training generally assumed a lecture format. Following a brief introduction by the member of the Advisory Board representing the circuit, the guidelines project director assumed responsibility for the presentation. Sessions generally lasted for approximately two hours.

Maryland began preparing for the guidelines training very early in the project, with the first Advisory Board discussion of these issues occurring in September 1980 (approximately nine months before the guidelines implementation). Staff presented preliminary suggestions for both the judicial training and sessions for other personnel, such as the defense, prosecution, and probation staff; the Board actively discussed the options presented. Early in 1981, staff refined the plans for the training, asking Advisory Board judges to serve as trainers, and agreeing on individual sessions conducted by the project staff for other interested personnel. In addition, in February 1981, a subcommittee of the Board was appointed to develop the content and format of the judicial training. At this date it was recognized that although a single training session could be held for all Maryland judges, several training sessions in each site would have to be held to accommodate the other interested parties. Finally, some months before the implementation date, staff and Advisory Board members settled the judicial training format and agreed on the general content and timing of the other sessions.

Preparation of the guidelines documentation began soon after the guidelines grids were finalized, in early May. Training materials included the guidelines manual, sample scoresheets, a completed scoresheet with its associated presentence investigation report, and a package of three sample cases including presentence investigation reports and the appropriate blank scoresheets to be used in scoring exercises for the judges. The manual, similar in many respects to that developed by Florida, included a brief history of the project and general instructions on such topics as eligible offenses, responsibility for completing the worksheet, procedures for recording the sentence and reasons for out-of-guidelines sentences, and instructions for distributing the completed materials. This was followed by scoring instructions for each of the three grids with a sample case for each, a list of guidelines offenses grouped by seriousness category, and copies of the sentencing matrices.

The judicial training, held on May 21, 1981, was the first of the sessions. Unlike the Florida training, staff did not present any of the lecture material; instead several Board members were involved in making the presentations. Specifically, the format was as follows:

- Opening remarks by the Chairman of the Advisory Board, in which the guidelines history, development, objectives, and use were explained.

- A presentation by a visiting judge from the Philadelphia Court of Common Pleas, where guidelines had been developed by the judiciary. This judge explained the operations of guidelines in Philadelphia and lauded the judges of Maryland for their effort in developing guidelines.

- A presentation by a second Advisory Board member in which the specific procedures and rules for guidelines use were explained. The sample guidelines case was reviewed, and specific questions of guidelines policy and practice were answered. This presentation was particularly effective; the speaker drew heavily on the Advisory Board deliberations and his own experience as a judge to review key points and allay the concerns of attending judges.

- A moot court session, in which various Advisory Board members presented a case, argued the aggravating and mitigating circumstances, scored the case using the guidelines, and sentenced the offender.

- Small group sessions led by Advisory Board members, in which participants reviewed and scored the practice cases included with the training materials.

- A concluding general session, in which judges raised particular concerns and questions about guidelines use. Questions covered such topics as development of the guidelines, scoring of the prior criminal history, and--most frequently--use of the guidelines in plea-negotiated cases.

The training session was followed by a general cocktail hour and dinner where judges could informally discuss the guidelines.

Training for other court personnel began after the judges' session and continued during May, June, and July. Generally, the sessions involved a presentation by the project staff, although Board members participated in some of these training sessions. Staff spent two days in Baltimore and one day in each of the other jurisdictions during the initial round of presentations, and follow-up sessions were held for many groups as well.

Several features of the Florida and Maryland training programs are worthy of note. In Florida, it was clear from the reactions of many of the participants that the timing of the training was a significant problem. Although notices concerning the training were sent to each site in advance, no site received its training program more than three weeks before the guidelines implementation. This proved to be particularly troublesome in the urban and suburban jurisdictions, where higher case volumes would demand more significant procedural modifications to cope with the introduction of guidelines and faster implementation of the required changes. While time pressures on the staff to develop and implement the guidelines undoubtedly dictated this schedule, it was clear that the lack of "lead time" served to exacerbate resistance to the guidelines in some sites. In addition, since the training was the primary method of alerting participants about the impending program, the issue of "lead time" was particularly sensitive. In Maryland, training was also provided at a relatively late date, with many of the sessions occurring after the actual implementation. However, only the probation staff and the urban prosecutors mentioned the timing of the training as a problem. Since these two groups were most likely to see a large volume of scoresheets early in the implementation period, this may explain some of their concern. The generally positive reaction of the other groups in Maryland may be due to the earlier outreach efforts of staff and the fact that many respondents found the manual self-explanatory and the guidelines easy to use.

As is evident from the above descriptions, the two states took very different approaches to the training logistics. Consider, for example, the judicial trainings: Florida used short lectures, while the Maryland session for judges lasted a half-day and consisted of a combination of lecture, demonstration, and trainee participation. By keeping the presentation brief and holding separate sessions in each test site, the Florida trainings were very convenient for judges and attendance was good. But even in Maryland, where the training session was held in a central location in the state, attendance was high: 86 percent of the test site judges attended. The longer format of the Maryland session also offered more learning opportunities. In addition, by virtue of the length and location of the Maryland training, those judges able to attend the session generally stayed for the entire proceeding. In Florida, the shorter sessions were often squeezed into busy schedules, and in some sessions participants had to leave before the training was complete.

Related to the format issue is that of the presentors at the session. Again focusing for the moment on judicial training, Florida made use of the Advisory Board

judge representing the jurisdiction and the central staff, with the staff taking the lead early in the session. Maryland, on the other hand, used several different Advisory Board members, while the staff took a supporting role only. This difference probably affected the participants' perception of the "source" of the reform, and may also have influenced their receptivity to the training (and guidelines) as a whole.

Far fewer differences between the states are evident in the training provided to prosecutors, defense counsel, and probation staff. Generally, these sessions were relatively short and were conducted by the staff. However, it should be noted that several respondents in Maryland expressed satisfaction with the follow-up training given, while Florida respondents criticized the lack of such training. In addition, in both states the relative inattention given to involving attorneys in either the development of training plans for the training sessions themselves may explain some of the resistance (indeed, hostility) of the prosecution in the larger sites.

5.3.4 Reactions to the Guidelines

Full implementation of the guidelines was delayed in some of the test sites in both states. However, reaction to the guidelines was swift. In Florida, the staff were soon flooded with questions, calls, and complaints about the guidelines. The guidelines staff also solicited further information on the guidelines process by letter approximately three weeks after the implementation date. In general, reactions of the rural sites were fairly neutral and were generally limited to questions about procedural issues (such as the necessity to wait for completion of a presentence investigation report before sentencing) or concerns about the sentence ranges. However, in the suburban and urban sites, reaction was much more volatile and negative. State Attorneys criticized the sentence ranges, finding numerous examples of cases which, in their opinion, merited sentences much more severe than those prescribed by the guidelines. Particular concern was voiced about the use of guidelines in conjunction with the parole guidelines system and the likely restrictions on attorneys' ability to negotiate guilty pleas. In the suburban site, negative reactions of the prosecution were accompanied by concrete action: announcement of a new "no charge bargain" policy, designed to counteract the perceived leniency of the guidelines. While defense counsel in the urban and suburban sites were not opposed to the guidelines ranges themselves, they shared the prosecutions' fears about the possibility of the guidelines reducing incentives to plea bargain. In addition, several complained that the involvement of the probation department in scoresheet preparation resulted in delays of several weeks. Adaptations made in procedure and policy as a result of the guidelines and these reactions are investigated in Chapter 6.

Reactions of the Florida judiciary were also mixed. Judges in all but the urban jurisdiction expressed no strong reservations about guidelines, although there was some initial adjustment. For example, one respondent in a rural jurisdiction characterized the judges as "perplexed, concerned, but enthusiastic," while another

judge wrote that he "had no quarrel with the guidelines and [found] them helpful." However, judges in the urban site were skeptical from the start, expressing strong reservations about the guidelines during the training session conducted in early April. The actual implementation of the guidelines did nothing to dissuade judges from these opinions, and as the Advisory Board representative from Jacksonville predicted, much of their concern centered around the possible delays guidelines might cause in the plea process. In early May, one judge, responding to a request for opinions on the guidelines, wrote a thoughtful but extremely negative response, criticizing the guidelines on several grounds. Among the arguments cited were that the guidelines were not sufficiently severe, that they were time-consuming to use, that they decreased the deterrent effect of criminal sanctions, and that they might lead to increased post-conviction rights for the defendant. In addition, the judge questioned the philosophical basis for the guidelines, arguing that no two defendants or their crimes are identical, and that defendants themselves should rightfully run the risk of drawing an unusually severe sentence from particular judges. This letter received wide distribution among the judges of Jacksonville and other circuits around the state; it became "a standard for the anti-guideline contingent."[25]

Initial reactions in Maryland showed a similar though far less dramatic pattern. In Harford County, implementation of the guidelines caused little concern over the first few months--in fact, four months after the start date, only four guidelines cases for single convicted counts had been completed and sent to the central staff.[26] Reactions of the judiciary in the suburban counties were similarly subdued, although staff noted that the number of scoresheets prepared during this period fell substantially below their expectations. In Baltimore, judges expressed some initial reservations about the guidelines; however, when guidelines came under attack by the defense and prosecution in the city, the judges demonstrated collective determination to continue with the test and to overcome the attorneys' resistance.

In Maryland, it was the prosecution, rather than the judiciary, which "led the charge" against guidelines. Although in two sites (Harford County and Montgomery County) prosecutors' reactions ranged from "wait and see" to strong public support, prosecutors in the other two jurisdictions voiced strong negative opinions. The Baltimore prosecutors were the first to react. By the third week of guidelines implementation, the State's Attorney actively refused to cooperate with the guidelines effort, basing his decision on two grounds:

- The requirement that attorneys complete the guidelines scoresheets in cases without presentence investigation reports. This, it was argued, was a judicial responsibility, imposed an unfair burden on the State, and presented particular difficulties since the prosecution was unwilling to vouch for the accuracy of the prior criminal hsitory information available to the State.

- Specific elements of the guidelines themselves. For example, the prosecution objected to the scoring of victim injury on rape offenses, the points for weapons usage, the scoring method chosen for prior criminal history (emphasizing the sentence imposed for past offenses), and the inclusion of prior employment history. In addition, the attorneys took issue with the guidelines treatment of white collar crimes, major drug dealers, consecutive sentences, plea negotiations, and crimes with a mandatory minimum sentence.

A clear undercurrent in both these sets of objections was the attorneys' perception that they had been excluded from the decisionmaking process and ill-informed about the plans for the guidelines. Although several outreach efforts had in fact been made vis-a-vis the state's attorneys, it was clear that they had not had the intended effect of increasing awareness and support. As a result, during a special Advisory Board meeting called to discuss the Baltimore State's Attorney's concerns, it was agreed that the prosecution and defense in Baltimore would provide prior criminal information but that they would not vouch for the accuracy of the information or complete the scoresheets themselves. In addition, it was agreed to add a second State's Attorney representative to the ex-officio membership of the Board. Since the defense also complained (though less vociferously) about the lack of input in the development, the Board decided to add a second defense representative to its membership as well.

Resistance in the Prince George's County State's Attorney's Office took a very different form. The prosecution in Prince George's County initially refused to recommend sentences outside the prescribed guidelines range--a policy which ran counter to the central project's intended implementation approach. Nevertheless, the State's Attorney indicated that this policy was adopted as a means of supporting the guidelines. Shortly after the implementation of the guidelines, however, individual attorneys began to notify their superiors of cases in which the guidelines sentences imposed were too lenient. This caused considerable concern in the office. In addition, as in Baltimore, the attorneys complained about their lack of representation in the planning and development stages. Although their "support" for the guidelines was not publicly withdrawn until December 1981, it was clear that practical support for the guidelines faded long before that point.

What of the longer term reactions to the guidelines? We have seen that in Prince George's County, at least, the initial prosecutorial reaction worsened during the test year. Was the same pattern evident in other jurisdictions, or did the initial reaction just reflect the inevitable resistance to change that fades over time?

To answer this question, midway through the guidelines test year in both states we asked judges, prosecutors, and public defenders to report their initial and current reaction to the guidelines. (Copies of the instrument and a discussion of

response rates are presented in Appendix G.) Although asked as two separate questions, by comparing the responses we can assess respondents' perceptions of change in their own reaction to the guidelines.[27] The results of this inquiry in Florida are presented in Figure 5-1. As might be expected, respondents in the 4th Circuit (Jacksonville) show a strong increase in negative attitudes concerning the guidelines; while this is largely due to the prosecutors' reactions, judges and defense attorneys also showed a similar shift. In the other Florida jurisdictions, the shift in opposition was far less dramatic, with one site showing a small increase in the percent opposed and one showing a modest decrease.

Though the attitudes in the Maryland sites also became more negative during the test year, this was largely due to prosecutorial rather than judicial attitudes. As shown in Figure 5-2, in the two larger jurisdictions (Baltimore and Prince George's County), respondents reported that their negative feelings for the guidelines increased over time. However, reactions of the prosecution in the two larger sites accounts for the bulk of this shift--in fact, in both sites judges' opinions of the guidelines became slightly more favorable after several months of experience with the system. In Montgomery County, the percent opposed remained small and stable, while no opposition was noted at either point in Harford County.

The reported sources of satisfaction and dissatisfaction with the guidelines remained relatively constant during the course of the guidelines year in both states. Leniency of the guidelines, philosophical disagreement with categorizing offenders and crimes for the purposes of sentencing, resistance to reduced judicial discretion, increased paperwork, concern about use of sentencing guidelines in conjunction with parole, and fear that the guidelines would result in an additional "defendants' right" or source of appeals were common complaints. The only major addition to these concerns was several respondents' observations that judges often "over-complied" with the guidelines, viewing them as "mandatory sentences."

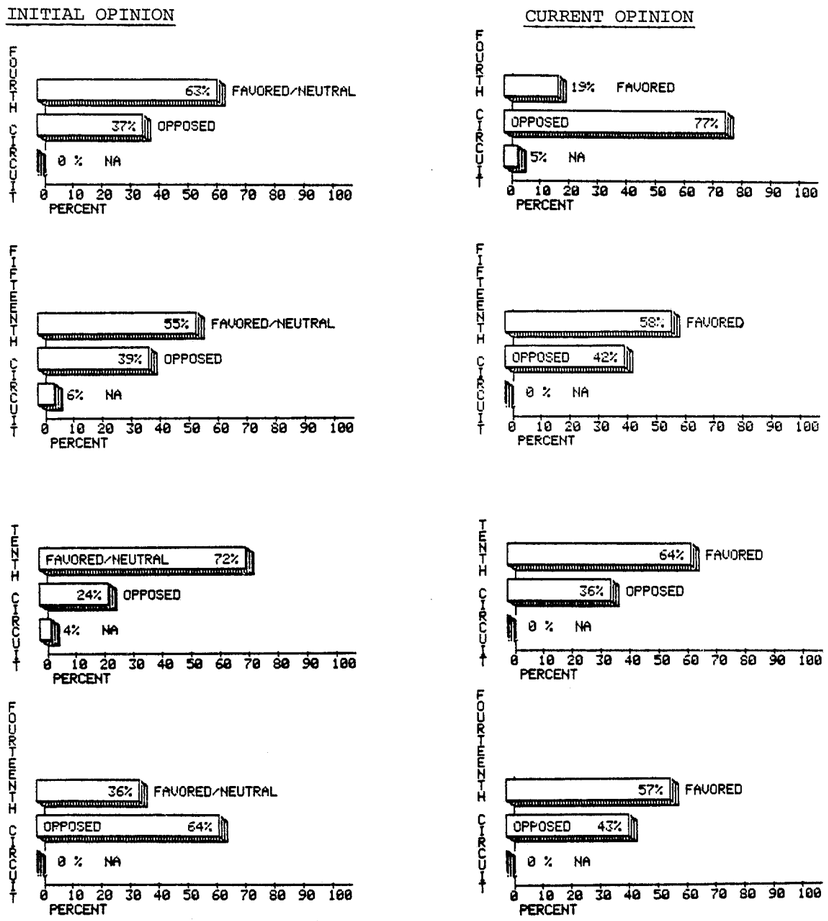
On the positive side, however, many individuals indicated several benefits of guidelines operations, including more uniform sentencing, use of the guidelines as a "training tool" for less experienced judges and other court personnel, use of the guidelines as a starting or reference point in plea negotiations, and guidelines' success in holding off mandatory minimum sentencing legislation in both states.

Many of the reactions discussed in this section not only highlight the importance of efforts to mobilize support for the introduction of guidelines, but also the importance of designing procedures that consider local norms. In the next section, we describe both the planning and implementation activities undertaken by the sites to accommodate the variations in local procedures, practices, and values.

FIGURE 5-1
Questionnaire Responses by Test Site for Florida:

"Did you initially favor or oppose your state's sentencing guidelines when they were first implemented?"

"Do you generally favor or oppose your state's sentencing guidelines now?"



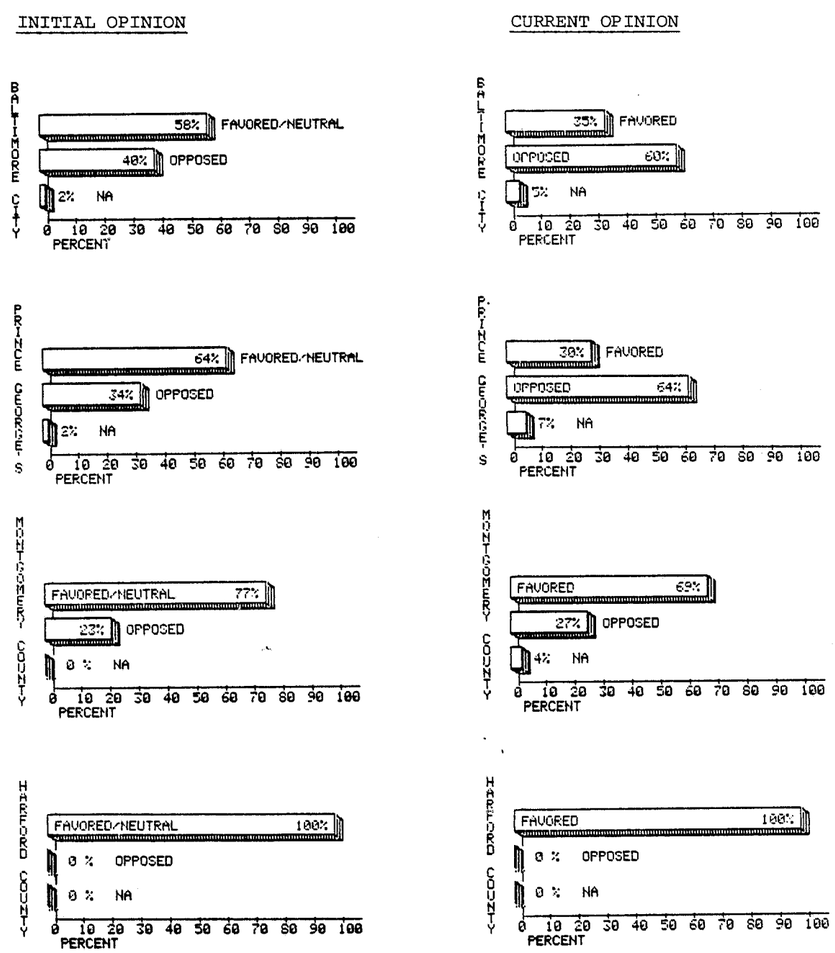
Data Set: Questionnaire responses

Base: All responses. (judges = 31; prosecutors = 45; defense attorneys = 37. No responses received from defense attorneys in the 14th Circuit.)

FIGURE 5-2
Questionnaire Responses by Test Site for Maryland:

"Did you initially favor or oppose your state's sentencing guidelines when they were first implemented?"

"Do you generally favor or oppose you state's sentencing guidelines now?"



Data Set: Questionnaire responses

Base: All responses. (judges = 45; prosecutors = 72; defense attorneys = 24. No responses received from defense attorneys in Harford County).

5.4 Accommodating Local Procedures and Characteristics

Sentencing guidelines systems involve at least two kinds of change: substantive and procedural. Substantively, the guidelines are designed to reduce variations in sentencing among individual judges or local jurisdictions. In most guidelines systems, the compromises and policy decisions necessary to accommodate local sentencing norms occur during the guidelines design phase, examined in Chapter 4. However, sentencing is not a single decision or event: it involves a series of procedures, decisions, and interactions. These, like the sentence outcome itself, may be subject to local variations. Since the introduction of guidelines is also a procedural change, part of the implementation effort must necessarily take these procedural variations into account.

As Galegher and Carroll have pointed out, innovations offering compatibility with past practice and demonstrable improvements over prior methods may be more readily adopted.[28] Ideally, then, a guidelines system would be perfectly tailored to each jurisdiction in which it is to be used, minimizing the changes participants would need to make in their daily operating routines and offering clear improvements as incentives in those areas that must change. Given the voluntary nature of the multijurisdictional sentencing guidelines, designing the implementation effort to accommodate local procedures is even more essential.

What kinds of procedures and characteristics might the guidelines projects consider? Clearly, any which potentially affect judges' sentencing decisions are important, including plea negotiations, presentence investigations, local caseloads, jail conditions, and typical channels of communication. And what kind of guidelines procedures might be affected by these conditions? The primary impact will be on the logistics of guidelines operation--procedures for obtaining the necessary information, scoring the case, and verifying both the information and the score. This section explores Maryland's and Florida's attempts to adapt these guidelines procedures to local conditions through their planning efforts; in addition, it examines the implementation procedures developed through this effort.

5.4.1 Learning About Local Procedures

As a first step in accomodating local characteristics, guidelines designers need a firm idea of the procedures and structures that are likely to influence guidelines operation. In a sense, this is not unlike the data collection effort used to construct the guidelines, in which designers gathered information on sentencing outcomes. The project Advisory Boards, and to some extent the construction sample data collection effort itself, provide two sources of information on local procedures. How well did the sites make use of these resources?

In Florida, staff faced substantial time pressure in the data collection and analysis; although part of this stemmed from the influence of the schedule established

by the field test grant, the ever-present threat of legislatively-mandated sentencing reform added some degree of urgency to the development effort. Given the time pressure and the test design's requirements for a substantial data collection and analysis effort, the staff and Board concentrated the bulk of their energies on the actual guidelines development during the first year of the project. Thus, for example, although staff visited the test sites extensively during the data collection period, their activities were concentrated on data collection tasks rather than learning about local court procedures and norms. The major exception to this observation was the issue of presentence investigation report preparation. As staff supervised the data collection effort, it became clear that presentence investigations were ordered on a smaller percentage of cases than originally anticipated, and this factor was later used in planning for scoresheet preparation.

Efforts to involve other significant parties (such as the Advisory Board) in the planning effort were hampered by time pressures, the limited membership of the Board (discussed above in Section 5.3), and the limited time available in the Advisory Board meetings to discuss logistical concerns. Again, however, there is one significant exception to this last observation. During the fourth Advisory Board meeting, at which members discussed the initial results of the data analysis, the issue of guidelines usage in plea negotiations was raised. The judge for the urban jurisdiction objected strenuously to the application of guidelines in these cases, noting that any impediment to the speedy processing of these cases would engender enormous resistance. As a result, some methods of ensuring more rapid turnaround for scoresheet preparation were explored for the Fourth Circuit, although the extent of this problem was still underestimated throughout the planning effort.

In Maryland, use of the Advisory Board and the data collection effort as planning resources was a bit more systematic. For example, initial plans for scoresheet preparation were discussed during the summer of 1980, when the staff were conducting the data collection. The project was quick to enlist the support and input of the Division of Parole and Probation through the Advisory Board's ex-officio membership, and secured an understanding that probation agents would complete the scoresheets for guidelines cases in which a presentence investigation report was ordered by the judge. This agreement was refined over time, and over the next year was the subject of several Advisory Board discussions and Advisory Board subcommittee meetings between the judiciary and the Director of the Division of Probation and Parole. In addition, it appears that an attempt was made to coordinate planned changes in the PSI format with the information requirements of the guidelines scoresheets, effectively increasing the incentives for probation agents to cooperate with the guidelines effort.

Furthermore, on the basis of the data collection, staff realized that presentence investigation reports would be unavailable on a large percentage of cases in some sites and began soliciting Advisory Board input on the best methods of

obtaining the requisite information in these cases. During the course of several Advisory Board meetings, it was decided that the judge would take responsibility for the scoring in these cases, although other parties such as the attorneys or law clerks were also mentioned as potential sources of assistance in this effort. In fact, staff used the comments and concerns expressed by some Board members (e.g., judicial resistance to completing paperwork and filling out justifications for extra-guidelines sentences) in their planning process. Thus, for example, it was arranged to exempt judges in the participating sites from the statewide requirement to document the reasons for all sentences in excess of three years; completion of the guidelines scoresheet and, if necessary, reasons for the sentence were substituted for this requirement in an effort to secure judges' cooperation.

It is worthwhile to note that both states' efforts to assess local conditions and procedures were relatively limited. Presentence investigation report preparation was, understandably, a primary focus of this effort, in part because it stemmed so directly from the data collection effort, and in part because it was such a central concern for scoresheet preparation. Nevertheless, other issues, such as the possible impact on the pace of justice or the possible reactions or resistance of the defense and prosecution, were not thoroughly addressed during this stage of the planning effort. In addition, neither state undertook special information gathering approaches--community meetings, interviews with key defense, prosecutorial, probation, or correctional staff, questionnaires on potential trouble spots, and so on--that might have helped them to understand the effects of local procedures on the guidelines. Not incidentally, such measures could also have contributed to support building for the guidelines by increasing participants' awareness of and input in the development and implementation process.

5.4.2 Planning for Guidelines Procedures

The second step in accommodating local characteristics is to match the guidelines procedures to local conditions. Essentially, this process raises two questions: how did the staffs use the information obtained from the Advisory Boards and the sites, and what was the extent and nature of their planning efforts?

In Florida, planning for the actual implementation of the guidelines was rather limited during the period before guidelines implementation, and the primary focus of the planning seemed to be the admittedly important issue of who would complete the guidelines scoring. The test design document suggested that probation and parole personnel be given this responsibility, and the Advisory Board members and staff agreed on this procedure, in principle, at least seven months before the implementation of the guidelines. This concept was refined somewhat as the implementation date approached, when it became clear that a relatively low number of presentence investigation reports were ordered in the urban and suburban sites. Since probation agents would have no information on the defendant unless they

prepared a PSI, it was agreed that the defense and prosecution could provide the information necessary for the guidelines computation, and that probation agents would simply use the information provided by the attorneys to compute the guidelines scores. Finally, it was decided that the court clerks in each test site would be responsible for copying the completed scoresheets and worksheets and returning them to the central office.

Although these general procedures were established before the onset of the guidelines, it was clear that local jurisdictions were not assisted in their efforts to translate these general instructions into specific procedures. For example, the need for such routine procedures as the following were not explicitly discussed with the sites:

- how to coordinate and verify scoresheet preparation when no presentence investigation report was ordered;
- the proper distribution of the scoresheets among the judge, defense, and prosecution;
- how completed scoresheets would be transferred to the court clerk and later to the central Court Administrator's office in Tallahassee;
- how to monitor completion of a scoresheet for every eligible guidelines case; and
- how to ensure that all the information needed on the scoresheet was provided.

Clearly, at the point of implementation, most responsibility for adjusting the guidelines to local needs was delegated to the Florida sites themselves. As the staff stated: "The instructions given probation and court personnel regarding the interaction between the two groups were deliberately vague in order to permit each jurisdiction to adapt the guidelines to the existing sentencing procedures. In this manner it was hoped that the outcome would be four models for dealing with the problem which could then be examined and used as a reference for statewide implementation."[29] In effect, the initial implementation period was used to develop the procedures which, under other circumstances, might have been considered part of the planning stage. To some extent, this approach worked; each site did eventually develop its own procedures tailored to its local norms and policies. However, many necessary procedures were never really implemented. The net result of this approach was at best confusion, and at worst, a further source of irritation for those who already questioned the viability of guidelines.

As noted above, the Maryland project began the planning effort early in its history, during the data collection phase. However, as in Florida, the primary focus of these planning efforts was on assigning responsibility for scoresheet preparation. It was decided that the probation department would complete scoresheets when a PSI was ordered, and that the judge would take primary responsibility for completing the materials in other cases. It was also decided that the defense and prosecution would participate in providing the requisite information in these instances--a decision which was not ratified by either agency. Unlike Florida, the Maryland project also focused its planning effort on the proper distribution of the scoresheets. The staff developed a special six-part form for the scoresheets, color-coded as to the distribution. Thus, distribution of the forms to the judge, central staff, probation department, court file, prosecution, and defense was facilitated.

Although the Maryland staff played a role in helping sites through the implementation period, it was clear that development of the "specifics" was largely left to the discretion of individual participants. Thus, for example, procedures for transferring scoresheets to the central staff and local verification of the scoresheets were not established in any of the test sites until the actual implementation period began. Like Florida, this period served as a "trial run," enabling the sites to formalize procedures and to translate the general implementation principles into specific activities which matched the needs of individual sites. In the next section, we examine in detail the net results of Maryland's and Florida's planning efforts, focusing on the logistical arrangements finally developed in each of the test sites.

5.4.3 Institutionalizing Procedures

As might be expected, actual operation of the guidelines developed differently in the two states, and indeed, among jurisdictions in each state. In addition, procedures evolved somewhat during the guidelines period, as the central staff and the individual sites adjusted to the guidelines. In this section we examine both the process of "installing" guidelines operations in the sites and the procedures ultimately adopted to complete the scoresheets.

The initial period of implementation in Florida was marked with significant confusion, particularly in the urban and suburban sites. In the urban jurisdiction, for example, delays in forwarding completed scoresheets were experienced; although the central project had designated the clerks for this task, many were not aware that they had the responsibility to send the materials to the central office. In addition, individual judges developed very different responses to the scoresheet preparation, with some requiring probation officers to be present in chambers to complete scoresheets, while others, at least initially, waited for presentence investigations to be completed. The suburban jurisdiction fared somewhat better, although the probation department soon fell behind in scoresheet preparation due to the heavy case volume. Rural jurisdictions, in contrast, appeared to adapt with little difficulty,

perhaps because PSIs were ordered for most cases and their lower case volumes allowed for a far more gradual introduction of the guidelines.

Any new procedure will inevitably cause some disruption and will require an initial period of adjustment. This is particularly true when the planning effort does not concentrate on methods of easing the transition by accommodating local procedures, as happened in Florida. In addition, the Florida technical assistance effort in the post-implementation period was relatively limited. Although staff visited each site during the training sessions conducted before guidelines implementation, technical assistance during the initial weeks of guidelines was primarily conducted by telephone, as staff answered specific inquiries on the guidelines or scoresheet handling. Approximately four months after initial implementation, staff again visited the sites to assess reactions to the guidelines and to ascertain any specific problems in guidelines implementation and use. The focus of this effort was more reactive than proactive, in that it was designed to identify problems rather than prevent them.

For the first six weeks of guidelines implementation in Florida, the Division of Parole and Probation had the exclusive responsibility for completing the scoresheets; although counsel could provide some of the information needed for scoring, probation agents were still supposed to verify most information items and submit the completed scoresheet to the judge. However, this requirement was modified within six weeks of the guidelines implementation. In cases in which no PSI was ordered, defense and prosecuting attorneys were permitted to prepare and submit the scoresheets without the involvement of the Probation Department. This procedural modification, combined with the minimal requirements established by the central project, allowed a variety of methods for scoresheet preparation to develop in the sites:

> 4th Circuit, Jacksonville. When presentence investigation reports were ordered, the agent completing the PSI typically prepared the guidelines scoresheet. When no PSI was ordered (as was the case in most plea negotiated cases), scoresheets were either:
>
> --filled out by the attorneys;
>
> --completed in the judge's chambers by a probation agent, using only law enforcement criminal history information;
>
> --prepared by the probation department several weeks (or months) after sentence was imposed; or
>
> --not completed at all.

15th Circuit, West Palm Beach. Initially, the Division of Parole and Probation completed all scoresheets. However, once attorneys were permitted to score non-PSI cases, the Division completed scoresheets only in cases in which a PSI was ordered. Otherwise, defense and/or prosecuting attorneys would prepare the materials, often in collaboration with each other.

10th Circuit, Bartow. Since most cases in this Circuit receive PSIs, most cases were scored by the Probation Department. Unlike other jurisdictions, in Polk County the scoresheet preparation procedure was highly centralized; after completing the PSI, agents forwarded it to one agent who prepared all scoresheets in the County. Highland and Hardee Counties adopted procedures more like those of the other test sites, with the agent preparing the investigation given responsibility for the scoresheet preparation. In some plea negotiated cases, however, the defense and prosecution assumed responsibility for scoring the case.

14th Circuit, Marianna/Panama City. Most cases received PSIs, and as a result, most cases were scored by the agent completing the PSI. Procedures varied widely when no PSI was ordered:

--attorneys might complete the scoring;

--probation agents attending the pretrial conference might fill out the scoresheet during that proceeding;

--the scoresheet might be completed after sentence was passed; or

--parties might not complete the scoresheet.

The first several weeks of guidelines implementation in Maryland were also marked with the initial period of adjustment typical of new reforms. In all the test jurisdictions, there was substantial confusion at first. For example, in spite of the training effort, many judges had the misconception that plea negotiated cases were not covered by the guidelines, and some failed to file scoresheets on these cases during the first three months. Still, significant difficulties in implementation developed only in the urban jurisdiction, where the prosecution objected strenuously to the presumption that assistant state's attorneys would provide information needed for guidelines scoring. In the remaining three Maryland jurisdictions, scoresheet preparation proceeded more smoothly; no doubt the lower case volumes and higher

percentage of presentence investigation reports ordered did much to facilitate guidelines implementation. In addition, the Maryland project staff actively visited the test sites during this period to provide follow-up training and technical assistance, which may have helped to ease the transition.

Basically, the Maryland plans for guidelines preparation were very similar to those ultimately established in Florida: the Division of Parole and Probation would be given responsibility for completing the scoring when a PSI was ordered. However, in cases in which no presentence report would be available, judges themselves were to assume responsibility for the scoresheets. In practice, this resulted in the use of a variety of approaches, as the sites adapted to their own requirements and norms:

> Baltimore City. Presentence investigation reports were not ordered in approximately half the felony cases in the city, according to many respondents. In these instances, procedures for completing the scoresheet tended to vary by judge; among those who might be given responsibility for scoresheet preparation were the judge's law clerk, the law clerk in consultation with the defense and prosecution, the defense only, and the defense and prosecution in collaboration.

Prince George's County. Again, probation agents prepared the scoresheets in the majority of cases, since PSIs were ordered relatively frequently. However, when no PSI was available, judges themselves, the attorneys, or the judge's law clerks would complete the scoring. In such cases, the information necessary for scoring would sometimes be obtained by putting the defendant on the stand to testify about his or her prior criminal record.

Montgomery County. While the high volume of PSIs ordered in this site meant that probation agents, by and large, completed the scoresheets, in non-PSI cases the judge would complete the scoring with information provided by counsel. In some cases, law clerks would be given responsibility for scoring.

Harford County. Judges assumed a substantial degree of responsibility for completing the scoresheets. While probation agents would generally complete the scoring when a presentence investigation report was ordered, judges themselves would often complete the scoring in other cases, sometimes in conjunction with counsel. A unique procedure of "pre-trial investigations" was also developed in this county; these reports were prepared

by the probation department in advance of a plea and could be used to estimate the guidelines score and sentence before negotiations.

It is clear that in both states, procedures were ultimately developed which matched the needs and pace of the local test sites. By allowing individual sites, and in some cases, individual judges to develop their own procedures, the projects may have succeeded in maximizing the compatibility of the guidelines with prior procedures--a crucial factor in acceptance of the new process. However, the short-term effect of this approach, especially in the larger jurisdictions, was a period of adjustment and confusion which may have solidified resistance to the guidelines, especially on the part of the prosecution. In addition, this approach allowed a diversity of procedures to develop, some of which were ultimately detrimental to the projects' success. Given the procedural variations in the sites and the voluntary nature of their participation, the need for monitoring becomes readily apparent. The states' activities in this crucial service are explored in the following section.

5.5 Monitoring and Compliance

In addition to support building and accommodating local norms, guidelines implementation involves one final component: monitoring guidelines use. Ultimately, the success of the guidelines project is dependent not only on its acceptance by participants, but on correct use of the new approach. Administrative procedures designed to assess the ways in which the guidelines are being used offer the only means of identifying and correcting mistakes or misinterpretations before they permanently derail the guidelines effort. The test design document clearly recognized this possibility, and it called for the development of a monitoring system to review the "accuracy and completeness" of the scoresheets and to permit "continued education and information" should problems be identified.[30]

However, monitoring also serves an important function ensuring compliance with the guidelines. In voluntary systems, like those tested in Florida and Maryland, monitoring is the main mechanism--though a weak one--for reminding all players of their parts in the effort. Given the loose or non-existent hierarchical relationship of the judiciary, defense, and prosecution,[31] and given the diversity of procedures established in the Maryland and Florida sites, it seems likely that monitoring would play a crucial role in the multijurisdictional sentencing guidelines test.

Of course, monitoring can occur at many different levels, with different consequences for guidelines success. At a modest level, monitoring can involve local efforts to verify the scoresheet preparation before sentencing. As we shall see in the following sections, the test sites often established informal procedures to check scoring. A more substantial monitoring effort (and one more in line with the recommendations of the test design) would be to establish a central administrative function

charged with review of (1) scoresheet completion and filing, (2) the case scoring and the degree to which actual sentences fall within the guidelines ranges, and (3) judicial reasons for extra-guidelines sentences. Finally, monitoring can also be broadly construed to include a sentence review function to determine the validity of departures from the recommended sentences. This was clearly not a feature of the field test, although the early planning efforts in both states had suggested such a capability when recommending guidelines.

This section examines the sites' efforts to monitor guidelines implementation and operation. After a brief review of the verification activities carried out at the local level, we describe the central monitoring efforts. Finally, we investigate guidelines use in the sites and summarize the analysis of guidelines compliance presented in detail in Chapter 8.

5.5.1 Verification Procedures

The integrity of the guidelines process depends to a great extent on the accuracy and completeness of the scoresheets. While the central staff can monitor their completion, this does little to eliminate potentially important errors in scoring (and consequently, sentencing). Depending on who completes the scoresheets, there are at least three opportunities for ensuring scoresheet accuracy: within the probation office (if a PSI was ordered), by the judge, and through review by the defense and prosecution. What kinds of procedures did the local jurisdictions establish to ensure the integrity of the guidelines scores?

In the Florida test sites, probation supervisors generally reviewed the scoresheets before they were submitted to the judge, although in many instances it appeared that this review was limited to a "math check." Procedures in the 10th Circuit differed somewhat, however. In Polk County, only one individual prepared scoresheets; because of his extensive experience with the guidelines, no internal review of the scoring was reported to be necessary. In Highland and Hardee Counties, it was reported that agents cross-checked each others' work on the scoresheets.

In most Florida sites, defense and prosecution were furnished with a copy of the scoresheet in advance of the sentencing. In the two rural jurisdictions, it appeared that materials were carefully reviewed before sentencing, and any objections to the scoring would be raised during the sentencing hearing or directly with the probation officer. In the suburban jurisdiction, most scoresheets were actually prepared by the defense and prosecution, and the relatively small number prepared by the probation department were usually reviewed by counsel. Finally, in Jacksonville it was reported that attorneys reviewed scoresheets during the first few months of implementation but later lost the incentive to do so when it became clear that guidelines were not a major factor in judges' sentencing.

Whether the scoresheets were prepared by the Division of Parole and Probation or by the attorneys, it appeared that judges generally reviewed them at or near the point of sentencing, when the defense and prosecution would raise objections or point out errors in scoring. Since attorneys would be unlikely to raise objections when they had collaborated in the scoring effort, judges relying on this approach would have little assurance that the scoresheets were, in fact, accurate. While some judges conducted independent checks, it was clear that in the two highest volume courts, many judges could not afford the time to verify scoresheet information.

Procedures for scoresheet verification in Maryland were similar to those in Florida. Representatives of the Division of Parole and Probation generally reported that supervisory personnel would review scoresheets, although the degree of attention given to this function appeared to vary among sites. Copies of the completed sheets were sent to defense and prosecuting attorneys before sentencing, and both parties were free to point out errors or misinterpretations before sentence was imposed. As in Florida, attorneys' review of the scoresheet was an important check on the accuracy of the scoring. However, some judges also indicated that they checked the scoring themselves or assigned their law clerk this function.

When no presentence investigation report was prepared, judges, attorneys, and law clerks could all be involved in preparing the guidelines scoresheets. In these instances, judges generally indicated that they or their clerk would check the scoring. In any event, such cases were relatively infrequent in all but the urban sites.

Local verification of scoring is certainly an important feature of any guidelines effort. But how can the sites examine the broader issues of guidelines use--questions concerning the extent of their application, the appropriateness of sentencing ranges, and compliance with guidelines requirements? The most logical response to these needs is to develop a central monitoring function.

5.5.2 Guidelines Monitoring

At least three major elements of guidelines use are potentially subject to administrative tracking:

- completion and filing of scoresheets;
- scoring and sentence distribution; and
- presence and content of reasons for extra-guidelines sentences.

In this discussion the Maryland and Florida approaches to and experiences with monitoring are examined, including the administrative procedures developed by the sites and the efforts to gain the cooperation of local court actors.

Tracking the completion and filing of scoresheets cannot be accomplished by the central staff alone, since the first step in this process is to relay completed scoresheets from the sites to the central project. How did the sites enlist and later support the cooperation of administrative personnel in the sites in this important activity?

As noted earlier, Florida delegated the responsibility for collecting and batching the scoresheets to the court clerks in each site. Unfortunately, the exact procedures involved in this effort were not well defined before the actual implementation, and as a result this process lagged during the initial months of implementation. It seems likely that the procedural problems were compounded by the scoresheet design: the Florida scoresheet was a single page with information to be recorded on both sides. Court clerks were supposed to make a copy of the scoresheet for dispatch to Tallahassee before the original went into the defendant's file.[32] In high volume courts, this could constitute a substantial burden on the clerks.

In Maryland, the process for returning scoresheets to the central staff varied among sites. In two jurisdictions, the clerk ultimately forwarded the scoresheets, while in the other two the judges themselves took responsibility for sending in the materials. This variety of procedures was no doubt made viable by the strong monitoring procedures established by the central staff and the design of the scoresheet. As noted above, Maryland used a six-part, color-coded form with the routing specified on its face: copies were to be distributed to the judge, central project staff, probation staff, court file, the prosecution, and the defense.

The second step in monitoring completion and filing involves checking scoresheets against dockets to assess shortfalls and improve filing rates. The Maryland project developed a manual system for these checks. During the year-long test, the Maryland guidelines project staff traveled to the four sites on roughly a monthly basis to list out guidelines-eligible cases from the court records. These lists were used to create a card file for checking in scoresheets. As scoresheets were received, the docket cards were manually sorted; after a period of time, cards without scoresheets were pulled for follow-up, which involved mailing requests for missing scoresheets to the sentencing judges.

Perhaps because of geographical constraints and its larger case volume, the Florida project did not send staff to the sites to list out eligible cases. However, neither did it arrange for such a listing to be compiled by the circuit court clerks or other court personnel. As a result, the project had no means to determine precisely which cases should have received scoresheets. Of course, such a listing has little utility unless the staff check incoming scoresheets against the master list. Manual checking in Florida would have been extremely difficult, given the high volume of cases. However, the case identifiers (such as docket number) needed to conduct such a match were available on the scoresheets. Unfortunately, the project lost a valuable opportunity to computerize this matching process: although docket numbers were

listed on the scoresheets, these numbers were not entered on the computerized records maintained on each guidelines case. Thus, even had a master list been available, it is unclear that Florida could have used the only feasible approach to monitoring completing and filing rates--computerized matching based on case identifiers.

A third series of monitoring activities can address scoring accuracy and the distribution of sentences within and outside the guidelines. This can include checks for:

- use of the proper form (where different scoring or grids cover different offenses);
- completion of all (or key) items;
- arithmetic accuracy and logic (e.g., correct point sums, scoring of juvenile record only for offenders under 26); and
- checks to determine the number of cases falling within and outside the recommended ranges.

As required by the test design, both states developed a computerized data base containing information on the scoresheets received from the sites. In the course of preparing this data base, both the Florida and Maryland guidelines project staffs performed a number of these checks. In Maryland, for example, staff did a preliminary check of scoresheets as they were received. Cases were checked to ensure that the correct guidelines form was used; if the wrong form was used the case was excluded from the project data base. In addition, the staff performed checks on the arithmetic accuracy and logic of the scoring, and reviewed scoresheets to ensure that the correct grid cell was used to determine the guidelines sentence. Again, if it was found that the wrong grid cell had been used, the case was excluded from the project data base. If certain necessary items of information (such as the actual sentence imposed) were found to be missing, the staff would notify the individual judges and attempt to obtain the missing data. Finally, using their computerized records, project staff assessed how many of the cases fell within, above, and below the recommended ranges.

In Florida, the staff also checked scoresheets as they arrived to determine that the correct guidelines form had been used for scoring, and excluded cases from their computerized records if the wrong forms were used. In addition, arithmetic and logic checks were performed, as was a check for missing data items. Although no systematic procedures were developed to obtain missing data items, approximately mid-way through the test year staff gave Advisory Board members a list of cases missing sentencing information and asked them to ensure that clerks in their

jurisdictions forwarded the information to the central staff. However, many scoresheets still lacked information on the actual sentence imposed. Finally, though the Florida project checked the distribution of sentences against the guidelines, this check was done using only the net sentence (i.e., incarcerative time imposed, minus any suspended time or probation). Possible errors in computing the net sentence (and thus possible errors in determining whether the sentence was within or outside the guidelines) could not be monitored using the computerized data base. In addition, because the scoresheets contained no information on mode of disposition (plea versus trial), staff could not detect systematic errors due to plea negotiations. As we shall see, such errors were in fact a major problem in the Florida guidelines use.

A final step in monitoring guidelines is to examine the reasons for extra-guidelines sentences, both in terms of judges' compliance with the requirement to provide reasons and in terms of the content of these reasons. In developing their computerized records of scoresheets, both states entered detailed information on the reasons provided by judges for extra-guidelines sentences. For example, the Florida data set included some 87 different response categories on judicial reasons, while the Maryland data base contained 218 separate response codes.

It was evident that the projects experienced a fair amount of difficulty in obtaining reasons for out-of-guidelines sentences and in ensuring that these reasons were sufficiently informative to aid in the analysis of sentences required by the test design. Although staff reminded judges of the importance of this function throughout the test year, neither project returned scoresheets to the sites when reasons for extra-guidelines sentences were missing or vague. Both states did conduct analyses of judicial reasons. However, as seen in Chapter 4, these analyses were not used to inform revisions to the guidelines during the test year, though the Maryland project did use its analysis to revise the guidelines after the test year was over.

Clearly, the monitoring efforts of the sites suggest a number of difficulties: there were problems in verifying scoresheets before sentences were imposed; missing information items, or indeed, missing scoresheets were not always detected or corrected; and only limited monitoring was carried out with respect to judicial reasons for extra-guidelines sentences. How did the sites actually fare in terms of the accuracy of scoresheet information and judicial compliance with the requirement to consult the guidelines? The next section answers this question by presenting qualitative information on guidelines use obtained from site participants and by summarizing the evaluation findings on compliance presented in greater detail in Part II of this report.

5.5.3 Compliance

The final key to guidelines implementation in the field test lies with the judge. Did judges appear to comply with the general mandate to "consult the guidelines when determining sentences?" Did judges across the sites and within sites

vary their approach? How often were sentences imposed within the guidelines, and did judges provide reasons for extra-guidelines sentences? In this section we briefly review judges' reports of their own guidelines usage[33] and the opinions of prosecutors and defense attorneys on this concern. Of course, self-reported patterns of use may not provide a full picture; to this end, empirical evidence of guidelines use is also reviewed in this section. For a more complete discussion of our empirical findings on guidelines compliance, readers are referred to Chapter 7.

Florida

As in other aspects of the guidelines implementation, it appears that judges' use of the guidelines varied between states, sites, and individual judges. Florida judges generally seemed to be highly aware of the guidelines; however, use of the guidelines in one site was reported to be limited at best. Individual responses of the judiciary concerning use of the guidelines present a picture of considerable diversity of use. For example, many judges indicated that they would first develop a sentence and then consult the guidelines; others used the guidelines to assess the recommendations of the defense and prosecution concerning sentence. Finally, judges in three of the four jurisdictions reported that they would occasionally sentence before reviewing the scoresheet, particularly if the defendant were willing to plead guilty and they could dispose of the case right away. In these instances, judges would sometimes complete the scoresheets after the fact, justifying extra-guidelines sentences when necessary.

Empirical findings on guidelines use generally confirmed our qualitative observations. As noted above, the Florida project developed a computerized data base on all guidelines scoresheets received. We were able to analyze these data to determine the percentage of cases falling within and outside the recommended ranges. However, determining the percentage of eligible cases that actually received scoresheets was more difficult. The Florida guidelines project conducted no independent check on filing rates. Still, we were able to obtain such an estimate by using data we collected on all burglary cases sentenced in the four sites during the guidelines test year. (See Chapter 8 for a more detailed discussion of the burglary data collection and analysis carried out as part of our evaluation.) By comparing our burglary data with the Florida project's data base on cases receiving scoresheets, evidence of the percentage of eligible cases receiving scoresheets could be obtained. Of course, it is possible that some scoresheets were completed but not forwarded to the central office. To the extent that this occurred, our comparison would underestimate the actual use of the guidelines in the sites.

Generally, for all types of burglary, scoresheets were filed for 57 percent of the eligible cases. Filing rates varied by site. In addition, we found that approximately 78 percent of the sentences for all cases (not just burglary cases) fell within the guidelines. Of the remaining cases, 10 percent were sentenced above the

guidelines, while about 12 percent were given sentences below the recommended ranges.

Combining these quantitative data with the results of our on-site observations, we note the following patterns in each of the test sites:

- 4th Circuit, Jacksonville. Guidelines were used primarily by judges with less experience on the criminal bench. Experienced judges tended to refer to the guidelines after imposing sentence, to use the guidelines only in the relatively few cases in which a presentence investigation report was ordered, and/or to ignore the guidelines for the majority of their cases. These observations were confirmed by our analysis of filing rates, which showed that only about 40 percent of the eligible burglary cases received scoresheets. In terms of compliance with the guidelines ranges, it was found that 78 percent of the cases filed with the central staff were within the guidelines. However, it should be recalled that many eligible cases did not receive scoresheets.

- 15th Circuit, West Palm Beach. Judges were highly aware of the guidelines and used them extensively in determining their sentences. Judges received scoresheets on virtually every case and reportedly complied with the guidelines ranges to a high degree. However, as we shall see in the next chapter, the effect of such compliance may have been moot due to score bargaining on the part of the defense and prosecution. The high filing rate reported by respondents was corroborated by our data: 100 percent of the eligible cases received scoresheets. Finally, the scoresheets filed with the central staff indicated that approximately 78 percent of the sentences were within the recommended ranges, while almost 95 percent were within or below these ranges.

- 10th Circuit, Bartow. Judges and other personnel in Polk County reported a great deal of reliance on and compliance with the guidelines. The comparison of burglary data and scoresheet data, however, indicated that only 60 percent of the eligible burglary cases received scoresheets. Anecdotal accounts suggested that much of this shortfall may have occurred in the two rural counties of this circuit: Highland and Hardee. There, it was reported that judges used the guidelines far less systematically than in Polk County. Some respondents indicated

that judges tended to overcomply, since three of the judges participating in the test were new to the felony bench and turned to the guidelines for assistance on sentencing issues. However, it appears that overcompliance may not have been as much of a problem as reported by respondents: approximately 76 percent of the scoresheet cases were sentenced within the guidelines (about the same as the 4th and 15th Circuits). However, if one considers sentences within and below the guidelines, this figure rises to over 90 percent.

- 14th Circuit, Marianna/Panama City. Judges expressed a great deal of commitment to using the guidelines and staying within the guidelines ranges. However, our analysis of filing rates did not corroborate the high filing rates cited by respondents: only 41 percent of the burglary cases received scoresheets. Given the rural, decentralized nature of this circuit, this rate may reflect problems in transmitting completed scoresheets to the central staff. However, some judges also indicated that scoresheets might not be prepared when the defendant was willing to plead guilty and be sentenced immediately; since this was reported to occur with some frequency, this may also contribute to the shortfall in scoresheets. Finally, of those cases that were filed with the central staff, 84 percent were sentenced within the guidelines. Cases sentenced below or within the guidelines accounted for 92 percent of all the scoresheets filed.

In seeking to explain the shortfall in case filing in Florida, we were led to consider the possible effects of plea negotiations. The scoresheet data show that (with the exception of West Palm Beach) almost all scoresheets filed were prepared in conjunction with a presentence investigation. Since PSIs are rarely ordered on pleas, it is reasonable to suggest that much of the shortfall in filing might have occurred in plea negotiated cases.

Using our burglary data, we were also able to estimate the accuracy of the scoresheet data by simulating guidelines scores for the burglary cases and comparing them to the actual scores noted in the project data bases. Again, a note of caution is in order: all the data elements used for scoring were not available in our burglary data. In addition, differences in scoring may be due to data entry errors in the project data base, rather than actual scoring errors. Even with these caveats, we found that the information on the project scoresheets was generally corroborated by our independent scoring, except in West Palm Beach. The pattern of scoring errors in that circuit corroborated reports of score bargaining between defense and prosecution (see Chapter 6).

Finally, we used both empirical and anecdotal evidence to explore the degree to which Florida judges complied with the requirement to give reasons for extra-guidelines sentences. Using the scoresheet data collected by the Florida project staff, we found that for all test sites combined, just under 80 percent of the extra-guidelines cases were accompanied by reasons. Significant differences were observed among the test sites in the rate of filing reasons; in addition, in three of the four circuits, cases falling above the guidelines were more likely to receive reasons than were cases sentenced below the guidelines. Evidence obtained from interviews with site personnel generally supports these findings. For example, judges in the 4th Circuit often indicated that they might not consider the guidelines before sentencing. However, if the sentence imposed fell outside the guidelines range, the probation department would return the scoresheet materials and the judge would complete the reasons for the extra-guidelines sentence--particularly if it fell above the recommended range. In the 15th Circuit, judges reported that they had no difficulty with the requirement to provide reasons for extra-guidelines sentences, although many also stated that such reasons were often completed by the defense or prosecution (rather than the judge) in plea negotiated cases. Judges in both the 10th and 14th Circuits reported that they complied with the requirement to justify sentences imposed outside the guidelines.

Essentially, the Florida project placed no restriction on the kinds of reasons judges could cite for extra-guidelines sentences. In examining the substance of these reasons, we found that 38 percent actually challenged the guidelines. That is, they repeated factors already accounted for in the scoring, or cited "plea negotiations," which were clearly intended to be covered under the existing ranges. The remainder of the reasons cited aggravating or mitigating circumstances of the case.

Maryland

As in Florida, Maryland judges indicated substantial variation in their use of the guidelines. By and large, judges reported ordering PSIs in a large proportion of cases and as a result, scoresheets were most often prepared by the probation department. Judges rarely indicated that they would sentence immediately upon plea and forego scoresheet preparation. Instead, however, several judges emphasized that they reached an independent sentencing decision before consulting the guidelines.

Again, self-reports of guidelines use in the site can be supplemented by our analysis of two data sources: our own data collected on every burglary conviction occurring during the test year in the Maryland test sites, and the project's data on all scoresheets filed with them during the test year. While a more complete discussion of our compliance analyses is presented in Chapter 7, this discussion previews those findings in conjunction with the findings from the numerous interviews conducted with judges, prosecutors, and defense attorneys during the test year.

As in the Florida analysis, our review of Maryland guidelines compliance includes an analysis of the project data to determine the percentage of cases falling within and outside the recommended ranges. We found that the overall rate of compliance with the guidelines ranges in Maryland was 72 percent. An independent check of the Maryland data against the ranges showed a similar pattern: 68 percent of the cases were sentenced within the guidelines. A second aspect of the compliance analysis was to compare our burglary data with the project's data to assess filing rates for buglary cases under the guidelines. We found that roughly 70 percent of the eligible cases received guidelines scoresheets.[34] However, since the Maryland project excluded cases from their data base due to fatal errors, some caution must be taken in interpreting this analysis. The Maryland project itself estimated that almost 85 percent of the eligible cases received guidelines scoresheets.

Again combining our empirical findings with the results of qualitative analysis, we observed these patterns of guidelines use in each of the Maryland test sites:

- Baltimore City. Most judges reported that they arrived at their sentencing decision before referring to the guidelines, but noted that the scoresheets were generally available. The figures on filing rates obtained from our comparison of project data and burglary data lend mild support to judges' self reports of filing rates: approximately 72 percent of the eligible cases received scoresheets. In addition, judges reported that they rarely sentenced outside the guidelines--a point corroborated to some extent by prosecutors' bitter complaints that judges rarely exceeded the recommended ranges. Project data also confirmed this point, showing that 71 percent of the cases were sentenced within the guidelines, while fully 93 percent were sentenced within or below the guidelines--a factor which may account for the prosecutors' concerns about guidelines sentences.

- Prince George's County. There were a number of reports that judges would sentence first and consult the guidelines afterward. In these instances, a scoresheet was sometimes completed after the fact, which would give the appearance (but not the fact) of compliance. In addition, judicial concern with the guidelines in general was reported to diminish over the test year. This picture of moderate use was supported by our analyses, which showed a filing rate of only 55 percent. In contrast to these findings, it was generally reported that judges complied with the recommended ranges. Many noted that judges felt obligated to follow the guidelines, although one judge observed

that the ranges were sufficiently broad that cases fell within the guidelines even if no reference was made to them. The overall compliance rate evidenced by the project data corroborates these findings: 73 percent of the cases fell within the guidelines, while 93 percent were either within or below the recommended ranges.

- Montgomery County. Judges appeared to view the guidelines as one of many sentencing considerations. Some reportedly used them only as an aid for the IN/OUT decision, while others noted that they decided on the sentence and then checked the guidelines. However, overall consideration of the guidelines was reported to be high. Our comparison of the burglary data and the project data strongly support these observations, with a filing rate of 88 percent--the highest of all Maryland test sites. Compliance with the actual guidelines ranges was reported to fade over the course of the test year, though judges indicated that they typically sentenced within the guidelines. Compliance rates were generally similar to those of the other jurisdictions--73 percent of the cases were within recommended ranges. Ninety-four percent of the cases were given sentences which were either within or below the guidelines.

- Harford County. Judges in Harford County were reported to use the guidelines and to participate extensively in the scoresheet completion process. With a filing rate of 79.5 percent shown through our data analysis, this observation is supported. Although respondents reported rather high compliance with the recommended ranges, actual rates were very similar to those in other jurisdictions (approximately 74 percent). It is interesting to note that Harford county judges exceeded the guidelines more often than judges in other sites. Nevertheless, the percentage of cases within or below the guidelines was still 90 percent.

Because our buglary data set was not suitable to a scoring simulation of the type carried out for the Florida data, no external validation of the accuracy of the scoresheets could be conducted for Maryland. However, an internal check using the scoresheet data was conducted. In general, it appears that the scoresheets entered in the project data base were accurate--only 5.8 percent showed any discrepancy between the point elements and the summary offense and offender scores. Still, it should be recalled that the project did not include cases in their data base if so called "fatal errors" committed in scoring affected the choice of guidelines range.

CONTINUED

2 OF 4

Judicial compliance with the requirement to provide reasons for out-of-guidelines sentences was also investigated in our interviews and quantitative data analysis. In general, Maryland judges did not indicate any particular reluctance to provide reasons for these cases. However, a large number stated that their high rates of compliance with the recommended ranges meant that they were rarely required to justify extra-guidelines sentences. Finally, some judges complained that the process of justifying non-guidelines sentences was unduly time consuming. Overall, then, judges and other site personnel presented a picture of general compliance with this guidelines requirement. However, on several occasions staff of the Maryland project requested judges in the sites to pay greater attention to this requirement and provide more detailed responses concerning extra-guidelines sentences. This, of course, would seem to indicate that judicial compliance was not as high as that reported by the sites. Indeed, in examining the quantitative data on the reasons provided by the judiciary, it would seem that the staff's concern was justified. In the aggregate, just over 50 percent of the extra guidelines cases included the reasons for these sentences.

As part of the compliance analyses, we also examined the content of the reasons provided by the judiciary for out-of-guidelines sentences. In general, the findings showed that a high percentage of the reasons cited factors already considered by the guidelines. For example, though the prior record was considered in the scoring, seriousness of the prior record was often cited by judges. Similarly, judges cited the fact that the sentence was imposed as part of a plea bargain, even though guidelines sentences were explicitly structured to cover both pleas and trials. These reasons, in essence, challenge the guidelines by repeating the ranges or scores allocated during the developmental process. Fully 45 percent of the reasons cite such challenges, while the remaining 55 percent of the reasons note aggravating or mitigating conditions.

Taken as a whole, the results of the Maryland and Florida compliance analyses are somewhat disappointing. In part, problems in establishing adequate monitoring systems may account for some of this difficulty, especially in Florida. However, it seems unlikely that more careful monitoring in and of itself would have enhanced compliance. Instead, it appears that the voluntary nature of the guidelines removed many of the incentives for compliance. In addition, the other mechanisms available to enhance compliance--support-building and accommodating local norms in the implementation process--did not (and perhaps could not) compensate for a voluntary guidelines system. In the concluding section of this chapter, we examine the implications of these findings for future guidelines efforts.

5.6 Conclusions

As evaluators, we have had the unusual good fortune to be able to interview both the guidelines developers and the ultimate consumers of the product during both the development and implementation stages of the test. Shortly before the guidelines were to go into effect, we asked one judge for a frank opinion about the need for this

kind of sentence reform. The judge turned to us, and with a slow smile answered, "Well, you know what I always say . . . Don't fix it if it ain't broke."

Reactions of this sort often typify attempts at reform, and the guidelines projects in both states braved such attitudes and worse in the course of the test period. The question for our evaluation was not "is implementation difficult?" The answer, almost invariably, is yes. Instead, our evaluation sought to explore two related issues: how difficult was the implementation, and why?

The short answer is "very difficult." We have seen that both projects experienced resistance, and indeed, hostility, to the guidelines. In addition, use of the guidelines was certainly more limited than either they or the sponsors of the field test might have hoped. Throughout this chapter, we have also seen some evidence of the reasons for these limitations--reasons grounded in the political process of implementing guidelines, the empirical rationale for guidelines development, and the voluntary nature of their implementation. In this section we reiterate these findings and cast them in the broader context of sentencing reform via guidelines.

Guidelines as a Change in Sentencing Policy

In both the original empirical sentencing guidelines research and the later developmental efforts for the multijurisdictional sentencing guidelines, guidelines were clearly viewed as a means of uncovering the "implicit" sentencing policy apparent in judges' previous sentencing decisions. The validity of this approach from a methodological standpoint has been thoroughly critiqued by others[35] and we will not review these concerns here. However, the effect of this claim on the guidelines implementation process is a matter of critical interest in this evaluation.

Judges in both test sites were concerned about legislative moves toward sentence reform and wanted to preserve the traditional judicial responsibility for this function. A system which promised only to reflect past practices, which offered voluntary implementation, and which could be implemented by the judiciary on their own initiative was likely to be viewed as the best of all possible alternatives, in that the basic "problem" could be addressed without instituting massive change or restriction in the system. However, the consequences of accepting guidelines on this basis appear to be numerous, and in many respects, detrimental to successful guidelines implementation.

The early guidelines efforts and the test design as originally envisioned clearly overemphasized the extent of judicial authority and the benefits judicial involvement would have in the development effort. Judges in the test sites did not automatically support the guidelines, nor, in many cases, did they comply with the guidelines requirements. Clearly, the need for substantial and ongoing support building even among the judiciary was not given sufficient emphasis. Furthermore, the project's emphasis on judicially-initiated reform tended to minimize the

participation of the defense and prosecution, who can play a key role in the sentencing process through plea negotiations.

The Advisory Board offers one means of overcoming some of these difficulties. However, to be truly effective, the Board must be truly representative. Members must be perceived as legitimate spokespersons among the groups they represent. Ratification of Board appointments by the groups they represent would be one means of achieving this, as would appointment by a third party with clear authority to make such selections--the governor or the legislature. In addition, Board members must have the power to shape the direction of the reform. In many respects, this means broadening the basis and authority for the reform beyond the judiciary itself. Furthermore, the role of the Advisory Board must be clearly defined as developing a new policy. To the extent that Board members understand this role and participate on this basis, they will be more effective advocates of the reform and can serve as change agents within their own constituencies. Broad representation on the Board is certainly important in this regard; however, given natural constraints on Board size, definition of the Board's policymaking role is probably the more important consideration.

Use of the Advisory Board members as change agents is one part of a larger outreach effort that must be undertaken if participants are to understand and support the new sentencing policy. However, outreach was another implementation feature underemphasized in the guidelines test. The major focus of the initial year of guidelines development was on the data collection and design of the guidelines. Although the Maryland project demonstrated continued concern with outreach, the staff in each state (and the test design document itself) generally emphasized the analytical component of the guidelines over the political and motivational components. In addition, in both states, the "selling" efforts that did occur seemed aimed primarily at the judiciary. As the experience in both states demonstrated, given the realities of case processing and plea negotiations, other groups such as the prosecution and defense must also be informed and involved in the guidelines process.

Both states used conventional methods such as the press and newsletters for their outreach efforts. It would seem that other more direct methods could also be used, including open Advisory Board meetings held in several locations throughout the state; soliciting input via working papers; public hearings on the guidelines; and more personal contact with key community leaders. This would have the additional benefit of widening the outreach effort beyond the criminal justice community to include the public at large.

Training for guidelines use offers another opportunity to point out the benefits of guidelines for the various participants. This would suggest that at least a portion of the training should focus on motivational issues. Issues of disparity reduction and legislative pressure to restrict judges' sentencing discretion were two such "motivators" used more or less successfully during the training of the judiciary in

both states. Another frequently mentioned selling point was the empirical basis of the guidelines and the fact that the matrices represented little more than past practice made explicit. However, training sessions for other groups tended to focus only on the procedures for guidelines use, rather than on motivational issues.

The need to sell guidelines to participants would also suggest that those conducting the training need high credibility and authority with each group. Maryland's use of the judicial members of its Advisory Board offers a positive example in this regard. In a judicially initiated system, project staff--however knowledgeable they may be in the development and use of the guidelines--are unlikely to possess sufficient credibility with judges, prosecutors, or defense counsel to serve as the primary trainers. Thus, it is probably advisable to solicit the involvement of representatives from all participating groups as trainers. In addition, training sessions cannot be conducted on a one-time basis if the implementation is to proceed smoothly and if unintended variations in procedure are to be avoided. Training sessions should be scheduled so that there is no need to rely only on hour-long lectures; a series of sessions provided for each group, with several follow-up technical assistance sessions, would seem to be a fairly minimal requirement.

By recognizing that guidelines represent a significant shift in policy and procedure, effective planning for the logistics of guidelines operations can be facilitated. Plans for guidelines logistics must take into account the current characteristics of each implementing jurisdiction. As demonstrated in both Maryland and Florida, case volumes, plea negotiation practices, jail populations, and use of presentence investigations are all likely to affect the way that sites must implement a sentencing guidelines procedure. For example, the combination of high case volume and low rates of presentence investigations will bode ill for any procedure relying exclusively on probation officers to complete the scoresheets; similarly, the combination of high case volume and extreme jail populations (found, for example, in Jacksonville) will demand procedures that do not increase pretrial detention times for defendants willing to plead guilty. While both states eventually addressed the issue of how to obtain completed scoresheets when no presentence investigation report was prepared, a more thorough awareness of the extent of this problem would have averted many difficulties. To address this and other issues of "match" between existing site procedures and the guidelines, staff might have made increased use of the Advisors' "technical knowledge" of their jurisdictions' procedures and norms, and supplemented the Advisors' experience by contacts with such groups as the prosecution, defense, and probation agency in each site. Expanding (or redirecting) the data collection effort to gather process data as well as sentencing data from each site could also provide this information with relatively modest effort.

Implementation plans may also be more effective if they offer some advantage over existing procedures as an incentive for change. Maryland's agreement to substitute guidelines scoresheets for the already-required explanations of sentences

over three years offers one such example; in addition, observations by some Florida judges that the guidelines scoresheets offer more information in plea-negotiated cases than was previously available might be another. However, it is certain that to secure successful implementation, these procedures must benefit other groups in addition to the judiciary. Again, a more thorough knowledge of the procedures and problems of each key group in the various sites can facilitate this effort.

Finally, a smooth transition to a guidelines system will require that some method be developed to identify logistical problems as they occur and before they damage the effort. By leaving the implementation process largely to the individual sites during the initial months of guidelines, Florida lost this opportunity and experienced a number of negative consequences. Still, technical assistance alone will not detect all existing and potential difficulties in guidelines operation; for that, the sites need an ongoing monitoring effort.

Monitoring Guidelines Operations

Accuracy of scoring, consistent use of the guidelines on all eligible offenses, and adherence to the requirement that extra-guidelines sentences be explained in writing constitute the basic elements of guidelines operation. Under the multijurisdictional field test, the central staff was to play the primary role in monitoring these functions, although it was noted that the accuracy of the scoresheet was the responsibility of the preparer.

The central monitoring function is indispensible for assessing filing rates, for observing systematic problems in guidelines scoring, and for collecting and analyzing the reasons for extra-guidelines sentences. A minimal requirement for accomplishing this is the development of a system that can identify guidelines-eligible offenses and track scoresheet completion for these cases. Hand-listing of cases on a local basis offers one means of identifying eligible cases, although systems with computerized court records might also incorporate such a listing into their records functions. This list can be compiled by the central staff; however, if it is possible to enlist the cooperation of local court personnel, delegating this responsibility to them is probably much more practical. Matching the master list of cases with scoresheets received can then identify cases in which the scoresheets may not have been completed. This checking process can be facilitated by certain elements of scoresheet design, such as use of multipart forms to ensure that scoresheets are forwarded to the central staff and incorporation of appropriate case identifiers on the scoresheets and in the data bases constructed from them. Such a monitoring process can help to ensure that guidelines materials are completed on all eligible cases, since monitoring in itself creates subtle pressures to comply and since reminders can be sent to local court personnel whenever scoresheets are not filed. Finally, using the data from the scoresheets, staff can then conduct checks to determine the extent of compliance with the recommended ranges, the accuracy of the guidelines scoring, and judges'

compliance with the requirement to explain extra guidelines sentences.

Under the multijurisdictional field test, all such monitoring was essentially carried out "after the fact"--that is, after the sentence was actually imposed on the defendant. Given the voluntary nature of the sentencing guidelines established in both states, errors or omissions were considered to be procedural problems more than matters of substantive concern relating to sentencing equity or due process. However, one can seriously question the value of a guidelines system in which scoring errors, rather than the unique circumstances of the case, produce departures from the recommended sentences. Similarly, a guidelines system in which some, but not all, defendants are sentenced with reference to the guidelines also appears to present serious questions of equity.

One solution to this difficulty might be to have the central staff check all scoresheets before sentence can be imposed. Such a procedure was followed during the implementation period in Minnesota,[36] for example. However, it is unlikely that the courts would be able to tolerate the delays in sentencing inherent in such a procedure. An alternative procedure (and one used to some degree in the sentencing guidelines test) would be to allow attorneys to check the scoresheets, and to rely on the adversarial system of justice to detect and correct errors before sentencing. Still, this does little to prevent "intentional" mistakes made to facilitate plea negotiations. One way to take advantage of attorneys' checks while deterring intentional scoring errors might be to combine attorney review with a system of "spot checks" in which some fixed number of cases are randomly chosen and checked before sentence is final. Other methods of dealing with this issue are explored in the final paragraphs of this chapter.

A second problem in the multijurisdictional guidelines test was judges' compliance with the requirement to provide written reasons for extra-guidelines sentences. In neither state was the monitoring effort designed to compel judges to comply with this requirement, and as we have seen, a substantial number of extra-guidelines cases were submitted without the associated justification. The central staffs were aware of missing reasons, and judges were requested a number of times to improve their compliance with the requirement and the quality and detail of the reasons provided. However, given the voluntary nature of the test, neither state was able to develop any means for assessing the validity of the reasons provided or even ensuring compliance with the requirement. This, too, undercuts the value of a guidelines system; without some check on the presence and validity of these justifications, judges are essentially free to continue the very practices that guidelines were designed to prevent.

The Question of Authority

Essentially, even the best monitoring system is toothless when implemented under a voluntary guidelines system. Monitoring can provide a means to detect systematic problems, and no doubt many well-intentioned criminal justice personnel will correct the errors they may have committed once they are notified of their mistakes. Still, absent the authority to enforce correct use of the guidelines, monitoring efforts cannot overcome resistance to the extra effort and more limited discretion that inevitably accompany this reform.

Once again, we return to the question of authority for the guidelines and the viability of a voluntary guidelines system. Based on the experiences of Maryland and Florida (and on the experiences of other guidelines efforts described in previous evaluations[37]) it seems that purely voluntary systems are unlikely to achieve the kinds of compliance needed to make guidelines a meaningful approach to sentence reform. Strong requirements for compliance with the basic conditions of guidelines are necessary, whether that be through court rule or legislative mandate. In addition, we would suggest that requirements for procedural compliance alone are likely to be insufficient. Without some capacity to review the validity of reasons for extra-guidelines sentences, and to limit the extent of the allowable departures from the recommended ranges, guidelines ultimately offer relatively little as a means of structuring judicial discretion.

FOOTNOTES

1. National Institute of Law Enforcement and Criminal Justice, Law Enforcement Assistance Administration, U.S. Department of Justice, Multijurisdictional Sentencing Guidelines Program Test Design (Washington, D.C.: Government Printing Office, 1978), p. 39.

2. Interim Report of the Sentencing Study Committee to the Florida Supreme Court (1978), cited in Kenneth J. Plante, "Florida's Multijurisdictional Sentencing Guidelines Project," a paper presented to the Chief Justices' Conference in Boca Raton, Florida, August 4, 1981. (Typewritten.)

3. "The Judicial Initiative in Sentencing Reform in Maryland: A Report of the Committee on Sentencing to the Maryland Judicial Conference," March 13, 1979, p. 2. (Typewritten.)

4. Ibid., p. 17.

5. Vincent O'Leary and Kathleen J. Hanrahan, Parole System in the United States, Third Edition (Albany, N.Y.: National Parole Institutes and Parole Policy Seminars, 1976), pp. 128-132.

6. Don M. Gottfredson, Leslie T. Wilkins, and Peter B. Hoffman, Guidelines for Parole and Sentencing (Lexington, MA: Lexington Books, 1978).

7. Florida Parole and Probation Commission, 39th Annual Report, July 1, 1978 to July 1, 1979, pp. 5-6.

8. See for example, Morgan Lyons, G. Alec Steele, and Robert E. DiGiacomo, "Objective Parole Guidelines for the Florida Parole and Probation Commission: Technical Report," (Tallahassee, FL: Florida Research Center, Inc., 1978). (Typewritten.)

9. M.D. ANN. CODE art. 41, § 112.

10. "The Judicial Initiative in Sentencing Reform in Maryland," Report of the Committee on Sentencing to the Maryland Judicial Conference, March 13, 1979, p. 18. (Typewritten.)

11. Ibid., p. 2.

12. Leslie T. Wilkins, Jack M. Kress, Don M. Gottfredson, Joseph C. Calpin, and Arthur M. Gelman, Sentencing Guidelines: Structuring Judicial Discretion (Washington, D.C.: U.S. Department of Justice, National Institute of Law Enforcement and Criminal Justice, 1978), p. 30.

13. William D. Rich, L. Paul Sutton, Todd R. Clear and Michael J. Saks, Sentencing by Mathematics: An Evaluation of the Early Attempts to Develop and Implement Sentencing Guidelines (Williamsburg, VA: National Center for State Courts, 1982).

14. Of course, the ability of mobilize sufficient support to institute such a procedure is a crucial question. It is worthwhile to note that in in Maryland, Appeals Courts judges objected strongly to the possibility of instituting sentence appeals. In Florida, the legislation authorizing the development of statewide guidelines initially included a provision for sentencing appeal. However, before final passage of the bill this provision was deleted, largely due to the significant opposition it faced among the judiciary and prosecution. Sentence appeals were ultimately included, however, once the statewide guidelines were actually implemented.

15. See for example, Franklin E. Zimring and Richard S. Frase, The Criminal Justice System: Materials on the Administration and Reform of the Criminal Law (Boston: Little, Brown and Company, 1980), p. 938; and Andrew von Hirsch, "Constructing Guidelines for Sentencing: The Critical Choices for the Minnesota Sentencing Guidelines Commission," Hamline Law Review 5 (June 1982): 167-168.

16. See, for example, Malcolm M. Feeley, Court Reform on Trial: Why Simple Solutions Fail (New York: Basic Books, Inc., 1983); Jolene Galegher and John S. Carroll, "Voluntary Sentencing Guidelines: Prescription for Justice or Potent Medicine?" Law and Human Behavior 7 (1983): 361-400; Rich, et al., Sentencing by Mathematics; Douglas McDonald, "On Blaming Judges--Criminal Sentencing Decisions in New York Courts: Are Guidelines Needed to Restrain Judges?" Summary Report (New York: Citizens' Inquiry on Parole and Criminal Justice, Inc., 1983); Albert W. Alschuler, "Sentencing Reform and Prosecutorial Power: A Critique of Recent Proposals for 'Fixed' and 'Presumptive' Sentencing," University of Pennsylvania Law Review 126 (1978): 550-577.

17. In the case of the 15th Circuit, the Chief Judge's designate.

18. Florida Multijurisdictional Sentencing Guidelines Advisory Board Meeting Minutes, December 14, 1979, Orlando, Florida.

19. Wilkins et al., Structuring Judicial Discretion, p. 30.

20. Of course, this is in keeping with the view that empirical guidelines do not represent a change in policy, but simply "past policy made explicit."

21. NILECJ, Multijurisdictional Sentencing Guidelines Program Test Design, p. 13.

22. For example, the Minnesota Sentencing Guidelines Commission was chaired by a citizen representative, as was the legislatively mandated Sentencing Guidelines Commission in the state of Washington. See Minnesota Sentencing Guidelines Commission, Preliminary Report on the Development and Impact of the Minnesota Sentencing Guidelines (St. Paul, MN: Minnesota Sentencing Guidelines Commission, 1982); State of Washington, Sentencing Guidelines Commission, Report to the Legislature (Olympia, WA: Sentencing Guidelines Commission, 1983).

23. Multijurisdictional Sentencing Guidelines Project, Final Report (Tallahassee, FL: Office of the State Courts Administrator, Supreme Court of Florida, 1982), p. 24.

24. NILECJ, Multijurisdictional Sentencing Guidelines Program Test Design, p. 38.

25. Florida Supreme Court, Sentencing Guidelines Project, Discretionary Grant Progress Report No. 7, April 1, 1981 through June 30, 1981.

26. Materials presented to the Sentencing Guidelines Project Advisory Board, October 9, 1981.

27. The first of these questions asked respondents "Did you generally favor or oppose your state's sentencing guidelines when they were first implemented?" Three response categories were then listed: favored, neutral, and opposed. In the second question, respondents were asked "Do you generally favor or oppose your state's sentencing guidelines now?", and two response categories were available: favor and oppose. The response categories in the two questions are thus not exactly comparable. In both Figures 5-1 and 5-2, we have therefore combined the favored and neutral responses to the first question and compared them with the percent opposed, since our focus is primarily on the potential for increased opposition to the guidelines over time. The figures do not reveal whether there have been shifts between the favored and neutral responses.

28. G. Zaltman, R. Duncan, and J. Holbeck, Innovations and Organizations (New York: John Wiley & Sons, 1973), cited in Jolene Galegher and John S. Carroll, "Voluntary Sentencing Guidelines: Prescription for Justice or Patent Medicine?" Law and Human Behavior, 7 (1983): 361-400.

29. Multijurisdictional Sentencing Guidelines Project, Final Report, p. 27.

30. NILECJ, Multijurisdictional Sentencing Guidelines Program Test Design, pp. 39-40.

31. For a discussion of the effects of court structure on guidelines compliance, see Galegher and Carroll, "Voluntary Sentencing Guidelines: Prescription for Justice or Patent Medicine?," pp. 389-390.

32. Florida Supreme Court, Sentencing Guidelines Project, Discretionary Grant Progress Report No. 7, April 1, 1981 through June 30, 1981.

33. During the guidelines test year, evaluation staff conducted in-depth interviews with felony judges in both states. In Florida, 24 judges were interviewed; the judges accounted for 76.7 percent of all sentencings for which scoresheets were collected by the guidelines staff. The percentages by circuit are as follows: 67.7 percent in the 4th Circuit, 86.5 percent in the 10th, 94.9 percent in the 14th, and 74.2 percent in the 15th. In the Maryland test sites, a total of 34 judges were interviewed. These 34 judges account for 86.6 percent of the sentencings represented by scoresheets in the guidelines project's data base for the test year. By jurisdiction, the percentages are: 91.4 in Baltimore City, 91.0 in Harford County, 80.3 in Montgomery County, and 78.7 in Prince George's County.

34. Note that our analysis of scoresheet filing rates does not include the first three months of guidelines operation. Startup effects are thus eliminated. If the initial three months of activity were considered, the filing rate would be noticeably smaller.

35. Rich et al., Sentencing by Mathematics; Richard F. Sparks, Bridget A. Stecher, Jay S. Albanese, and Peggy L. Shelly, with a chapter by Donald M. Barry, Stumbling Toward Justice: Some Overlooked Research and Policy Questions about Statewide Sentencing Guidelines (Newark, N.J.: School of Criminal Justice, Rutgers University, 1982).

36. Minnesota Sentencing Guidelines Commission, Preliminary Report on the Development and Impact of the Minnesota Sentencing Guidelines, p. 15.

37. Rich et al., Sentencing by Mathematics; Sparks et al., Stumbling Toward Justice.

CHAPTER 6

SYSTEM ADAPTATION TO GUIDELINES

Throughout the guidelines implementation period, one thing became increasingly clear: only in the broadest sense were we observing a change taking place in two states. A much more accurate description, and one which we came to use with greater frequency as the test unfolded, was that we were observing a change in eight distinct jurisdictions, each with its own procedures, norms, and "culture." Not surprisingly, we have seen that reactions to the guidelines and judges' compliance with guidelines requirements differed substantially among the eight test sites. This chapter examines the ways in which the sites ultimately adapted to the guidelines system. It focuses not just on sentencing, but on the broader set of activities entailed in the felony disposition process: initial prosecutorial activities in screening and charging, plea negotiations, sentencing decisionmaking, provision of presentence information to the judges, and post-sentencing concerns.

In examining local system adaptation to the guidelines, we actually focus on two distinct considerations. The first is based on the near truism that any new procedure may involve changes--anticipated and unanticipated--in existing procedures. We seek to describe what those changes were, so that others may be aware of the kinds of modifications they may face under a guidelines system. The second consideration is that many of these modifications actually involve far more than simple changes to accommodate a new procedure. Instead, they may represent reactions designed to maintain existing systems or prerogatives in the face of change--in essence, reactions which change the form of the system in order to maintain its previous function.

These adaptations are, we believe, a crucial focus of our evaluation. Several studies over the years have illustrated that sentence reform efforts such as minimum mandatory sentences, habitual criminal offender statutes, and sentence review often face an uncertain future when implemented. At least in part, this is due to unanticipated adaptations on the part of judges, prosecutors, and/or defense attorneys.[1] Efforts to implement sentencing guidelines have also been subject to these same pressures and concerns, as shown in recent evaluation studies.[2] The limited success of many of the earlier guidelines efforts is of particular concern for this evaluation, given that the multijurisdictional sentencing guidelines field test is based in large part on those demonstration efforts. As discussed in more detail in Part II of this report, Florida and Maryland also experienced difficulties in fully achieving the avowed goals of sentencing guidelines during the initial test year of implementation. Thus, part of the intent of this chapter is to explain these findings by identifying factors in the implementation which may have led to limited success.

Of special interest, of course, are the effects of jurisdictional differences on the operation of guidelines. To support this investigation, a variety of methods was used to obtain site-level data on procedures and modifications. Approximately one year before the guidelines were formally implemented, evaluation project staff visited each of the sites in both states. Semi-structured interviews were conducted with judges, prosecutors, defense attorneys, staff of the division of parole and probation, and other court personnel. Based on these interviews, we developed descriptions of the courts' operations prior to the introduction of guidelines. These were circulated to site participants shortly before the guidelines implementation, both as a check on their accuracy and to obtain information on any changes in procedure that had occurred since the initial site visits. Once guidelines were implemented in each state, participants' reactions and adjustments were tracked by means of personal interviews at the beginning, mid-point, and end of the guidelines test year. (See Appendix K for information on the on-site interviews.) In addition, several months into the guidelines test, we distributed written questionnaires to all judges, prosecutors, and defense attorneys in the sites, requesting information on guidelines operations and their perceived impacts on court processes.[3] (Copies of the questionnaire and a discussion of response rates are included in Appendix G.)

Based on these materials, this chapter explores several features of the guidelines implementation. In the first section, we briefly examine the influence of guidelines on the prosecutor's functions of screening and charging. This is followed by a discussion of plea negotiations and the influence of the guidelines on these procedures. Given that plea bargains are by far the most common means of disposing of cases in both Florida and Maryland, and given the crucial relationship of this process to sentencing, it is likely that the negotiation process will be strongly influenced by the introduction of guidelines. In addition, we will explore guidelines impact on other processes related to sentencing, including presentence investigation report preparation. We also examine both the perceived effects of the guidelines on sentencing and summarize our own findings on sentencing impacts. These are examined in greater detail in Chapter 8, Part II of this report. Finally, we briefly discuss the implications of our findings on system adaptation for sentencing guidelines in general.

6.1 Prosecutor Screening and Charging

Through the screening and charging function, prosecutors are given a major role in determining the ultimate sanction imposed on the offender. Leaving aside for a moment the considerable influence on sentencing exerted through plea negotiations, by the charging decision alone the prosecution controls the "maximum exposure" of the defendant to certain criminal sanctions, and in many cases, the minimum exposure as well.

In conversations with prosecutors, judges, and defense attorneys before the guidelines introduction, most agreed that screening and charging procedures would probably remain stable under a guidelines system. However, some alternative hypotheses were also noted:

- Prosecutors may increase filing of non-guidelines offenses to circumvent the sentence restrictions inherent in a guidelines system. Under these conditions, we might expect to see increased use of more serious misdemeanor charges, for example.

- Charging decisions may be made with some view toward the matrices or scoring procedures: that is, prosecutors may adjust charges to maximize the number of points received by the defendant. This could entail charging certain kinds of offenses and not others, or making increased use of multiple offense indictments.

- Prosecutors may decline to prosecute certain minor offenses, if it appears that the maximum penalty to be awarded would not be commensurate with the effort involved in prosecution.

Despite these suggestions, it appears that in fact, the guidelines had very little effect on prosecutors' screening and charging behavior in the test sites in Florida and Maryland. Even recognizing that this assertion is subject to certain limitations--we rely primarily on respondents' reports of change and descriptions of their procedures before and after the start of the guidelines, rather than direct observation or analysis of prosecutor records--questionnaire responses and interview results present a remarkably consistent picture. The questionnaire responses of Florida and Maryland personnel are displayed in Figure 6-1. Evidently, most respondents indeed found that there were few if any changes in this prosecutorial function. The only major exception to this observation rests in Florida's 15th Circuit (West Palm Beach), although other individuals noted some changes on an isolated basis.

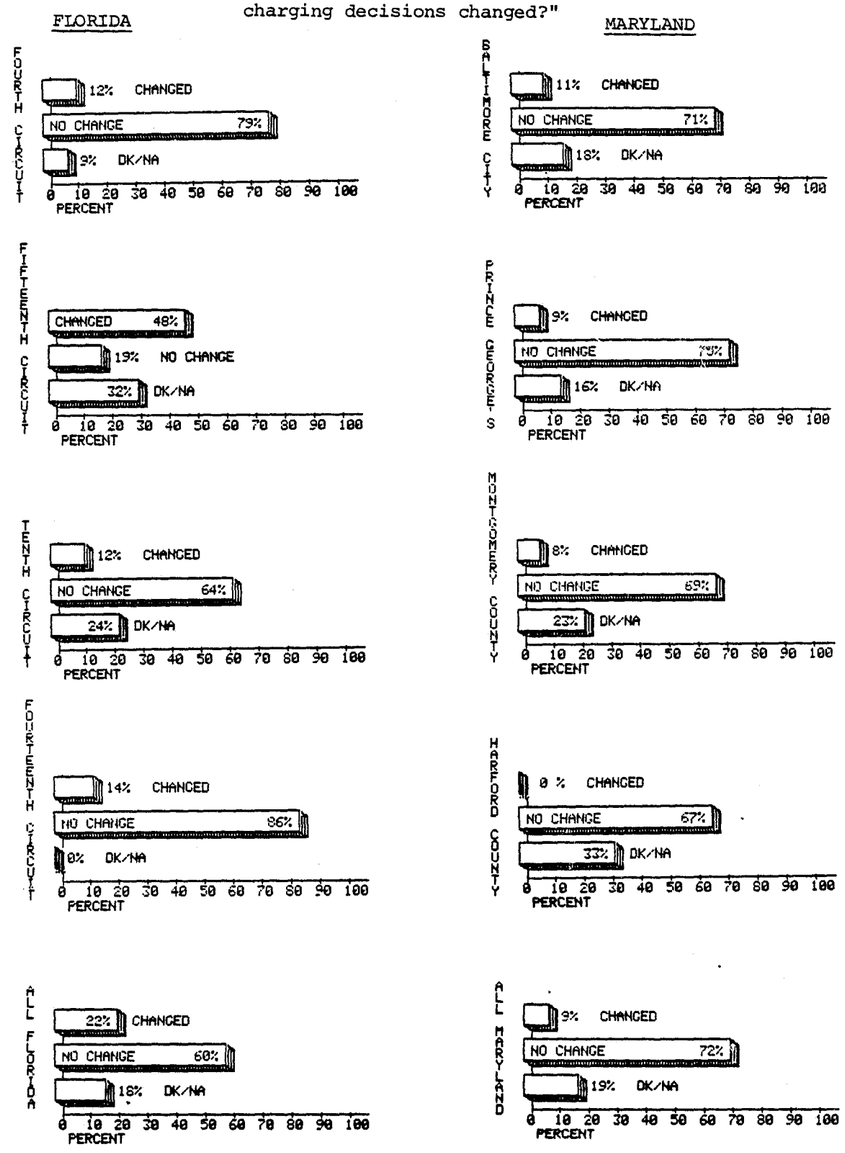
What was the nature of the change described by respondents in West Palm Beach, and what can account for the major shift occurring there? In addition, what is the meaning of the isolated reports of change in other jurisdictions?

In Florida's 15th Circuit, the answer lies primarily in the prosecutors' reactions to the introduction of sentencing guidelines. In Chapter 5 we noted that the reaction to the guidelines on the part of prosecutors was mixed in both states, and was generally more negative in the more populous jurisdictions. In West Palm Beach, negative reaction was accompanied by a concrete policy shift: with the introduction of sentencing guidelines, the State Attorney in the 15th Circuit announced a new

FIGURE 6-1

Questionnaire Responses by Test Site for Florida and Maryland:

"Since the implementation of guidelines, have the prosecutors' initial charging decisions changed?"



Data Set: Questionnaire responses

Base: All responses. (For Florida: judges = 31; prosecutors = 45; defense attorneys = 37. For Maryland: judges = 45; prosecutors = 72; defense attorneys = 24. No responses received from defense attorneys in the 14th Circuit (FL) and Harford County (MD).

policy eliminating charge bargains. According to many respondents, in practice this meant that once the Filing Office formally filed the case, trial assistants would not be permitted to reduce charges or drop counts without consultation with the Chief of Felonies in the State Attorney's office. Defendants pleading guilty were thus required to plead to every count of the indictment.

As might be expected, this policy shift had serious consequences for the intake and filing functions of the State. Since reductions in charge or number of counts was unlikely during a plea bargain, the State Attorney's filing unit made more conservative decisions, filing fewer counts and lower charges. Cases were also reviewed more carefully after the initial filing. (The effect of this policy change on plea negotiations is examined in Section 6.2 below.)

Use of limited charge bargaining policies in Florida was hardly a new procedure in the Florida test sites: even before the introduction of the guidelines the State Attorney in the 15th Circuit had limited charge bargains on certain types of offenses and had restricted plea negotiations in general for career criminal offenders. In addition, before the introduction of guidelines the State Attorneys in two other Florida test sites--the 10th Circuit and the 14th Circuit--had also constrained charge bargaining to some extent, by insisting that guilty pleas include a plea to the first count of the indictment. However, only in West Palm Beach did the prosecutor institute a no charge bargain policy which resulted in significant changes in filing policies. It appears that the less stringent limitations on charge bargaining in other jurisdictions may account for the relative stability in their filing practices.

Although by far the most notable changes in filing practices were confined to West Palm Beach, a small number of the respondents in other jurisdictions did note some adjustment in these functions. The general direction of these changes is consonant with the hypothesized changes presented above. For example, one prosecutor in Florida's 14th Circuit noted that he might increase the number of counts charged in order to get a higher guidelines score.

There were also some reports of changes in screening and charging policies in Prince George's County during the last few months of the guidelines test. According to some respondents, more non-violent first offenders were being prosecuted in the lower courts under misdemeanor charges, to relieve pressure on the Circuit Courts. Some attributed this policy change to pressures brought about by the guidelines, stating that more cases were going to trial. However, at least one respondent conjectured that increased use of the lower courts resulted when the State changed its procedures and began to file separate indictments on each count and on each defendant in a criminal incident. A shift to use of the lower courts was also noted in Florida's 10th Circuit, although the purported reasons for the change differed from those in Prince George's County. One judge and several prosecutors noted that during the final months of the guidelines, less serious felony cases were often filed as misdemeanors. Prosecutors apparently found that under guidelines, the typical

sentence for a minor felony offense was probation. However, as a misdemeanor offense the case would not be heard by judges participating in the guidelines test and a short incarcerative sentence would be imposed.

In spite of these latter reports of procedural changes, it was clear from most respondents that the guidelines did not exert a strong influence on charging, and that where such influence was noted, it was more a result of changes in plea negotiation policies or case volume than of the guidelines per se. In part, the stability of filing procedures may lie in the guidelines design in Maryland and Florida: by basing the scoring on charges at conviction rather than charges as filed, by ensuring that like crimes are generally scored in the same grid (and thereby subject to the same maximum penalties), and by basing the scoring system on the sentencing event rather than individual charges or crimes, incentives for "adapting" the filing process were reduced.

However, it must be recalled that our conclusions are drawn on the basis of only one year's experience with the guidelines. It is possible that adaptations in charging procedures could occur, given more experience with the guidelines. In Minnesota, for example, the preliminary year evaluation did not reveal major shifts in charging policies under the sentencing guidelines.[4] However, after three years of experience, there is some evidence that prosecutors may be modifying their policies in an attempt to increase the scores, and hence, penalties, imposed under the guidelines.[5]

6.2 The Plea Negotiation Process

It is now common wisdom that in many jurisdictions--perhaps most--guilty pleas constitute the most frequent means of criminal case disposition. Estimates of the percentage of cases "pleading out" vary widely among jurisdictions, but it is not uncommon to read figures of 60, 80, or even 90 percent. These figures are certainly in line with the experience of the test sites in Maryland and Florida; respondents often estimated that between 75 and 95 percent of the cases were settled through guilty pleas.

More often than not, guilty pleas are the result of negotiation between the defense and prosecution--negotiations aimed at limiting the "exposure" of the defendant to criminal penalties, either through some discussion of sentence or some adjustment of the charges filed against the defendant. Clearly, then, through the medium of plea negotiations, the lines of responsibility for sentencing become somewhat indistinct. For this reason, as well as the sheer volume of such dispositions, plea negotiations become a crucial concern in the felony disposition process, and by extension, in the operation of sentencing guidelines.

Perhaps the first feature of note in discussing the plea negotiation process in Maryland and Florida is that we cannot, in fact, refer to a single process at all.

Procedures differ between states; more accurately, they differ among jurisdictions and even among courts within a jurisdiction. However, for the most part we have tried to capture the typical process(es) in each of the eight sites both before and after the guidelines implementation, comparing the two time periods to assess the possible impact of guidelines on the negotiation process. Thus, much of the discussion that follows will focus on changes in individual sites. In the concluding paragraphs of this section we will attempt to draw more general conclusions about the impact of guidelines on plea negotiations and to explain some of the changes in individual jurisdictions.

What kind of changes might we expect in the plea negotiation process as a result of guidelines? Again, several possibilities were suggested by respondents in the test sites themselves, including the following:

- Guidelines might alter the nature of the negotiations between defense and prosecution. For example, the discussions might center around the guidelines factors or the guidelines sentence.

- Use of charge bargains would increase under guidelines. This is certainly consonant with current thinking on sentence reforms, which suggests that any restriction in judicial sentencing discretion will transfer that discretionary power to the prosecution.[6]

- Guidelines might decrease the willingness of the defense to engage in plea bargains. For example, it was noted that if the guidelines do not take guilty pleas into account, there would be little incentive for the defense to plead, since the sentence would be virtually identical after trial. As a result, some predicted that trials would increase.

- In contrast, knowledge of the probable sentence might actually encourage pleas, since the defense would have greater certainty about the case outcome. A reduction in the number of trials might result.

- Under guidelines judges might become more involved in plea negotiations or at least more willing to provide an indication of the sentence to be imposed.

How did the guidelines affect plea bargaining in the test sites? To answer this question, we must first explore the typical plea negotiation practices in the sites before the introduction of guidelines.

6.2.1 Negotiation Procedures Prior to Guidelines Implementation

It was apparent from the pre-guidelines interviews that, in most sites, negotiations generally involved some discussion of sentence. For example, all jurisdictions in both states made use of some formal plea mechanism in which the judge hears the terms of the plea negotiation, including sentence, and then agrees to be bound to these conditions. Under this type of plea, the defendant is allowed to withdraw the plea of guilty and the case may be tried before another judge if the original judge finds that he or she cannot accept the terms of the agreement. However, most respondents in all sites indicated that use of this type of plea arrangement was relatively rare.

In addition to formal agreements concerning the sentence, a variety of methods--some of them quite imaginative--were also used. Most jurisdictions made use of the prosecutor's recommendation as one means of conveying the sentence agreement. In some jurisdictions (Florida's 10th Circuit: the southern counties in Florida's 14th Circuit: Baltimore City) this was a primary means of conveying the information to the judge. Other jurisdictions relied heavily on judicial involvement, with the judge explicitly or implicitly indicating the maximum sentence that would be imposed. This approach was used extensively in Prince George's County, and was often used by some judges in Jacksonville, West Palm Beach, and the 14th Circuit in Florida. In Montgomery County a different approach was used: defense attorneys and the prosecution would agree on the judge who would take the plea. Since judges' "going rates" were generally common knowledge, this helped set the sentence boundaries in the plea negotiation.

This is not to say that charge or count negotiations were not used in the test sites in Florida and Maryland. For example, before the introduction of guidelines, several sites reported that it was relatively common practice for the prosecutor to drop second or third counts of the indictment as part of the negotiation.[7] Often, this was done in conjunction with the sentence recommendation or a sentence cap specified by the judge. Relatively few respondents claimed that accepting pleas to a lesser count of the indictment was typical, although it occurred occasionally in Baltimore, Montgomery County, Harford County, and to some extent the 10th and 14th Circuits in Florida.

6.2.2 Plea Negotiations Under the Guidelines

Given these plea negotiation practices, we can now assess the degree to which changes in plea negotiation behavior actually occurred. In the preceding chapter, judicial use of the guidelines in plea negotiated cases was examined. In Florida, we found a very mixed picture. In Jacksonville, where the filing rate of guidelines scoresheets fell far below expectations, there was very suggestive evidence that many, if not most, of the cases in which scoresheets were missing were plea bargains. Just the opposite was true in West Palm Beach, where judges used the

scoresheets consistently in both pleas and trials. Based on interview and quantitative data obtained for the two rural test sites, there was also evidence that scoresheets might not be prepared on some plea bargains in which no PSI was ordered and the judge sentenced immediately. While there was some indication in Maryland that plea bargained cases might receive proportionately fewer scoresheets than cases going to trial, the results were far less conclusive.

Still, these results dealt primarily with judicial use of the guidelines in plea bargains. Given the central role of the defense and prosecution in plea bargains, we sought to obtain a wider perspective on the degree to which guidelines had become a part of the negotiation process. One of the most basic indicators, of course, is the discussion of the guidelines during the actual negotiations. Figure 6-2 shows the responses of judges, prosecutors, and defense counsel in both states when asked if the guidelines were discussed during plea negotiations. As the Figure shows, in most jurisdictions, guidelines were indeed incorporated in the plea discussions. In all but one site, over 80 percent of the respondents indicated that the guidelines were "usually" or "sometimes" discussed as part of the plea negotiation. Virtually identical responses were obtained during in-person interviews conducted at the close of the test year.

As might be expected, the major exception was Jacksonville. There, a third of the respondents indicated that guidelines were never discussed in plea negotiations, while only 19 percent indicated that the guidelines were "usually" discussed. As noted above, use of the guidelines in plea bargained cases was relatively rare in that circuit. In fact, on the basis of interviews conducted at the close of the test year, it was determined that many participants in the 4th Circuit understood that negotiated cases had been explicitly exempted from the guidelines a few months after the implementation date.[8] Thus, in some Jacksonville court divisions, guidelines were almost never discussed or consulted as part of the plea bargain or sentencing in general; even in one division where guidelines were used in plea bargains, the judge indicated that the attorneys lost interest in discussing the guidelines because his sentences generally did not vary from the recommended range.

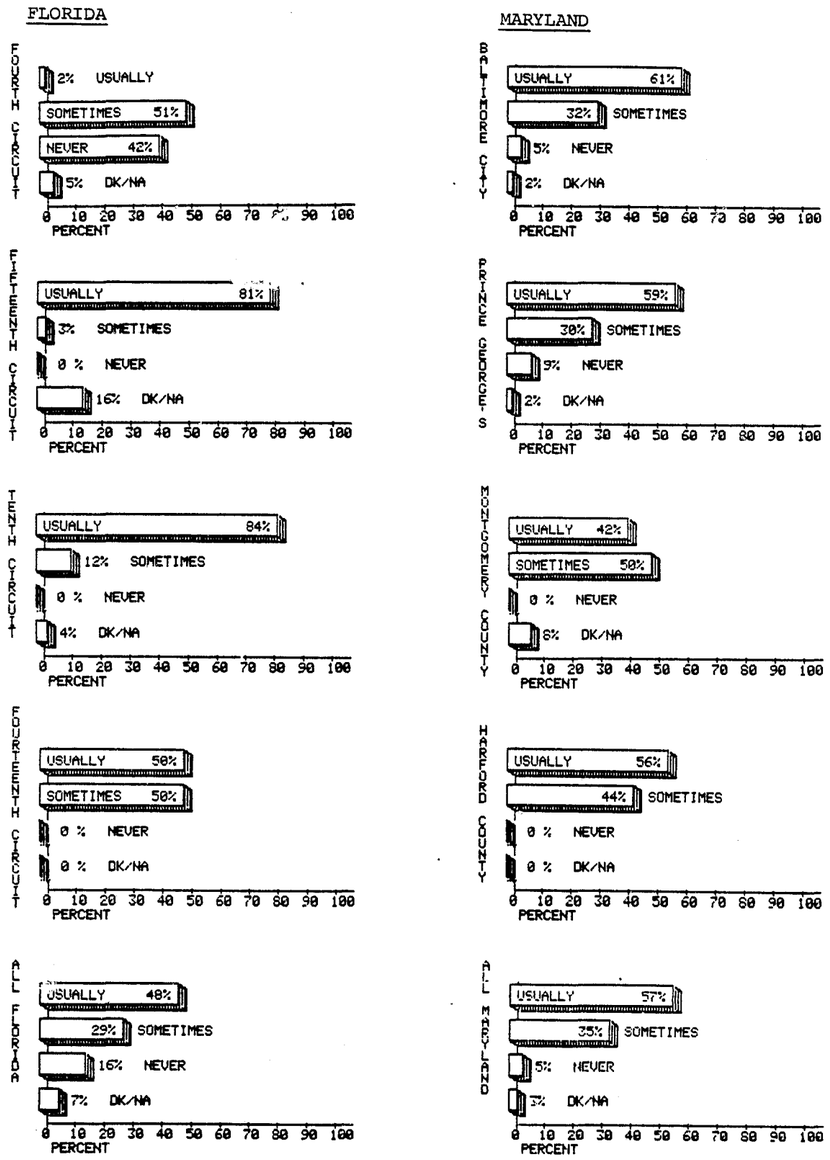
Knowing that the guidelines were discussed as part of the negotiation says much about participants' awareness of the guidelines and indicates that a new factor had been added to the negotiation strategy for many attorneys. But what were the practical consequences of these discussions? What, if anything, changed in the plea negotiation process?

A general answer to this question was obtained by asking judges, prosecutors, and defense attorneys in each jurisdiction if the plea negotiation process had changed since the implementation of guidelines. Figure 6-3 shows the responses by site for Florida and Maryland. While opinions tended to be divided in each jurisdiction, some general patterns do emerge. For example, in Florida the majority of respondents in the 4th Circuit (Jacksonville) and the 14th Circuit (Marianna/Panama City) indicated

FIGURE 6-2

Questionnaire Responses by Test Site for Florida and Maryland:

"As far as you know, are the guidelines discussed as part of plea negotiations?"



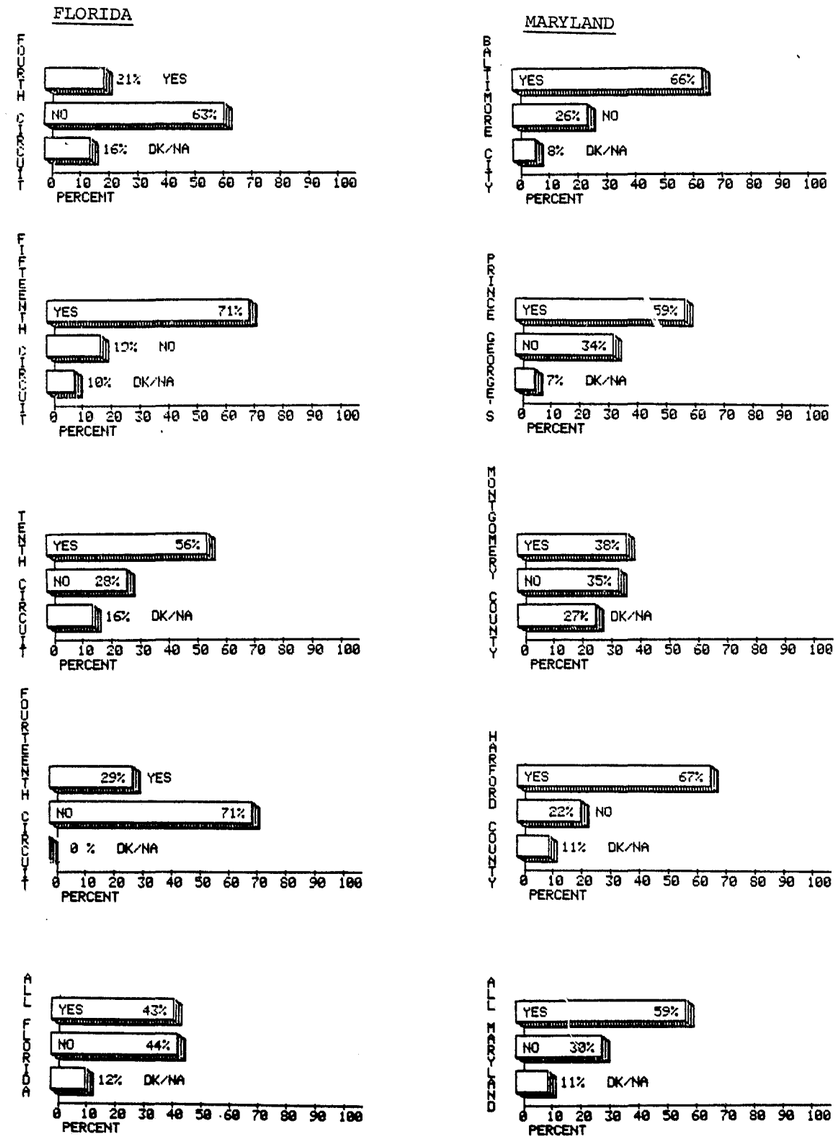
Data Set: Questionnaire responses

Base: All responses. (For Florida: judges = 31; prosecutors = 45; defense attorneys = 37. For Maryland: judges = 45; prosecutors = 72; defense attorneys = 24. No responses received from defense attorneys in the 14th Circuit (FL) and Harford County (MD).

FIGURE 6-3

Questionnaire Responses by Test Site for Florida and Maryland:

"Since the implementation of guidelines, has the plea negotiation process changed?"



Data Set: Questionnaire responses

Base: All responses. (For Florida: judges = 31; prosecutors = 45; defense attorneys = 37. For Maryland: judges = 45; prosecutors = 72; defense attorneys = 24. No responses received from defense attorneys in the 14th Circuit (FL) and Harford County (MD).

that no changes had occurred, while in the 10th and 15th Circuits there was a fair degree of consensus that change had taken place. Similarly, three of the four Maryland jurisdictions (Baltimore, Prince George's County, and Harford County) affirmed some changes while respondents in Montgomery County were almost evenly split among those who noted a change, those who found no change, and those who had no knowledge of the situation. For the most part, these patterns are consonant with those reported in Figure 6-2 on discussions of the guidelines during plea negotiations.[9]

To examine the nature of these changes more closely, we turn to specific written comments provided by various respondents and to the results of our interviews conducted before and after the guidelines implementation. We will attempt to describe changes at a variety of levels: those necessary only to operate the guidelines, those that appear to be an adaptation intended to manipulate the guidelines, and those which, in our opinion, had little or no relationship to the guidelines at all.

As noted above, sentence caps (whether indicated by the judge, recommended by the prosecutor, or agreed to by the defense and prosecution) play a key role in the negotiation process in most of the Maryland and Florida test sites. One of the most frequently reported effects of the guidelines was in shaping the cap or recommendation. This effect was specifically noted in three of the four Florida jurisdictions (the 10th, 14th, and 15th Circuits), where it was often reported that plea negotiations presented to the judge were typically "within the guidelines." The precise method of structuring these sentence agreements naturally varied from site to site (and from judge to judge). In the 10th Circuit, for example, the prosecutor's recommendation usually conformed with the guidelines range. In the 14th Circuit, a variety of approaches was used: often, the sentence agreement presented by the defense and prosecution would be within the guidelines, and several individuals noted that defendants may "plead to the guidelines sentence." However, judges also used the guidelines as a cap in many instances, indicating to counsel that they would "probably follow the guidelines" when presented with a plea. Even in the 15th Circuit (discussed in more detail below) it was reported that plea offers of the state were "dictated by the guidelines." As noted above, the guidelines had little impact on plea negotiations in Jacksonville; no such shift towards the use of guidelines sentences was observed in that site.

The experience in Maryland was generally similar. In Baltimore, for example, it was noted that the state intensely disliked the guidelines, and would not generally agree to the guidelines sentence. As a practical matter, however, most Baltimore cases were sentenced in or below the guidelines--a fact well known to the defense. Thus, it was reported that sentences suggested during plea agreements were generally within the guidelines and that once again guidelines served as a de facto cap. One judge even said he insisted that plea agreement recommendations be within the guidelines. Montgomery County respondents noted a variety of practices, including

pleas to the guidelines ranges and sentence recommendations in line with the guidelines.[10]

Although the prosecution in Prince George's County initially recommended the guidelines sentence in every case (and insisted on a plea to the first count of the indictment), this procedure was changed approximately halfway through the test year. A policy against offering sentence recommendations was instituted, and as a result it was rare for a sentence agreement incorporating the guidelines to be presented by the defense and prosecution. However, a compensating strategy developed; according to many respondents, judges, rather than the state, would often indicate the guidelines sentence as a cap, with the result that negotiations often appeared to exclude the prosecution. In Harford County, it appeared that explicit sentence agreements between the state and defense had been far less common than in other jurisdictions. This pattern continued into the guidelines period; however, it appeared that the guidelines took the place of the informal indications of sentence which were often provided by judges in that jurisdiction, and served as a de facto cap where none had existed before.

At first blush, it may seem somewhat unlikely that prosecutors in so many of the jurisdictions would agree to the guidelines sentences in developing plea negotiations, particularly since prosecutors almost universally reported that they found the guidelines ranges too lenient.[11] Why, then, would they agree to the guidelines or recommend the guidelines sentence as part of a plea negotiation? While we can give no definitive answer, some suggetions may be offered on the basis of our on-site observations. For the most part, we believe that this approach represents little more than a practical adjustment to the guidelines--one adopted simply to "keep the system going." Recall that sentence negotiations are a major factor in almost every test site. If we assume that the sentences typically recommended in these instances were usually below the prosecutorial "ideal" in order to reach a compromise with the defense, it is possible that the use of guidelines sentences may have been roughly comparable to past practice, and thus grudgingly accepted. In fact, comparability with past plea negotiation behavior is not unlikely in empirically based guidelines where plea negotiated cases make up the bulk of the construction sample data used to develop the guidelines.

Alternatively, it may be that the prosecution is, in fact, relatively constrained in providing above-guidelines recommendations, given the ever-present need to move the docket. Knowing judges' reluctance to sentence outside the guidelines--particularly to go above the guidelines--and recognizing that the defense may refuse to plead to anything other than a sentence within or below the guidelines, the prosecution may simply have bowed to practical constraints and recommended the sentence most likely to be accepted. Again, this may not represent a considerable change from the previous situation, where the state's recommendation often balanced the judge's sentencing preferences against the demands of the defense. Finally, there

was evidence that even when prosecutors refused to accept the guidelines ranges, judges' use of the guidelines as an upper bound would result in continued plea negotiations.

While the use of the guidelines as a de facto sentence cap is perhaps the most common change, it is by no means the only one cited by respondents. And in fact, changes in the prosecutorial recommendation or cap were often accompanied by other modifications in the process. For example, two of the Florida sites (the 10th and 14th Circuits) reported occasional use of charge negotiations in which attorneys would agree on a sentence and then adjust the charges to ensure that the guidelines score would match the agreement. However, this was not reported to occur as a general pattern.

Matching charging negotiations to the guidelines scoring was not reported in any of the Maryland sites, although it is certainly possible that some "targeted" charge negotiations occurred there. However, given the relatively low contribution of the instant offense to the total score in the Maryland guidelines and the high degree of overlap among the guidelines ranges, it is likely that this approach offered less leverage in Maryland than in Florida, where the sentence range was more dependent on the charge at conviction. In addition, as noted in Chapter 3, the Maryland criminal code does not offer a strict felony degree structure and prescribes broad penalty ranges (or no ranges at all) for many offenses, which may decrease the importance of charge negotiations as a means of constraining sentence.[12]

The above discussions would seem to indicate that few dramatic changes in the plea negotiation process occurred in the test sites, and for the most part this is true. However, in two of the eight sites--one in Florida, and one in Maryland--the introduction of guidelines was accompanied by major policy shifts on the part of the prosecution. Since these changes had important consequences for plea negotiations and the use of guidelines in these sites, we describe them in some detail below.

As noted in Section 6.1 above, on the day of guidelines implementation, the State Attorney in West Palm Beach announced a new "no charge bargain" policy. While charge bargains had in fact been limited to some extent even before the introduction of guidelines, and were not in any case the primary means of plea negotiation in that circuit, the new policy had one major impact: defendants would be required to plead guilty to all counts as filed. Trial assistants would not be allowed to drop second or third counts of the indictment, nor could they accept pleas to lesser included offenses. While filing practices were adjusted to make certain that the charges were truly reflective of the offense, this new policy enraged defense attorneys, who refused to enter guilty pleas under these conditions. Thus, during the initial months of guidelines implementation, plea negotiations virtually ground to a halt, angering all parties, especially the judiciary. Trial backlogs increased dramatically: one judge's trial caseload was reported to have doubled during 1981 (which included the first eight months of guidelines operation).

With this degree of pressure, some adaptation to the guidelines was inevitable: two approaches were actually adopted. As one defense attorney put it, "the situation definitely caused more trials until people realized it would not be practical unless we went under the guidelines or played with the score." Thus, although the prosecution argued that the guidelines ranges were too lenient, it became common practice over time for defense and prosecution to agree to a sentence below the guidelines. Judges overcame their initial resistance to sentencing outside the guidelines, and began accepting such pleas.

Such sentence negotiations may be within the bounds of "acceptable" guidelines practice, in that judges may disagree with the guidelines sentence and impose another, so long as the reasons for this departure are documented. Arguably, these reasons could include the judge's belief that the sentence was unduly severe given the conditions of the case or the norms of his jurisdiction, which seems to be the situation in the 15th Circuit. However, in West Palm Beach, attorneys, rather than judges, often completed the reasons for out-of-guidelines sentences. In addition, as noted in the preceding chapter, the reasons often cited "plea bargains" rather than the sources of disagreement with the sentence.

The other form of plea negotiation--score bargaining--was clearly a means of subverting the guidelines process. As noted in Chapter 5, soon after the guidelines implementation the Florida project modified its procedures to permit the state and defense to complete the guidelines scoring themselves for cases in which no PSI was ordered. The change was made to create some way of using the guidelines in all cases. Practically speaking, this meant that most plea negotiated cases in West Palm Beach would be scored by the attorneys. (Note that PSIs were ordered more frequently for plea bargains in the rural jurisdictions in the test: probation agents rather than attorneys were thus able to derive the guidelines score for many plea bargained cases.) Since the defense was not enthusiastic about accepting pleas within the guidelines ranges, and judges were equally unenthusiastic about sentencing above the guidelines, score bargaining provided a logical solution: the prosecution would not have to modify its no charge bargain policy, judges could sentence wtihin the guidelines, and the defense would obtain a lighter sentence than if the case were scored correctly. Thus, while completing the scoresheets attorneys would negotiate the scoring of some of the more "flexible" factors in order to arrive at a point score that would match the sentence agreement of the defense and prosecution. Among the factors most commonly subject to negotiation were role of the defendant, extent of victim injury, and victim precipitation. In addition, scoring of the prior record was also open to manipulation, either through assuming that offenses from other states were not as serious as their Florida counterparts or by ignoring dispositions, which meant that the prior conviction would not be counted in scoring. Finally, in some instances, the attorneys would deliberately score cases in the wrong crime category if this would result in a sentence agreeable to both the defense and prosecution.

As noted by the attorneys themselves, it is clear that this adaptation was made possible by the lack of verification procedures for the scoresheets, either within the circuit or by the central staff. However lamentable the score bargaining procedure may seem, one can only speculate what would have happened in the high-volume West Palm Beach courts had more stringent monitoring procedures precluded this option. At best, it might have forced an extensive charge bargain system or induced high numbers of extra guidelines sentences as judges struggled to control case volume by granting sentences below the guidelines to induce pleas and imposing higher sentences on those who go to trial--in effect creating a two-tiered guidelines system with different norms for trials and plea negotiations. At worst, the West Palm Beach courts could have been faced with intolerable trial volumes or with the kind of noncooperation experienced in the Jacksonville courts.

Of all the participating counties in Maryland, only Prince George's experienced substantial changes in the plea negotiation process, again largely due to new policies of the prosecution. In mid-1979, approximately one year before the guidelines implementation, the administrative judge of the county established a system of pre-trial conferences for felony offenses. During the year immediately prior to guidelines implementation, these proceedings were the primary means of arranging sentence negotiations. Typically, each felony case was scheduled for a conference before one of three judges, and at that time the judge would generally indicate a sentence cap for the offense. While a different judge could (and usually did) take the actual plea, the second judge would typically comply with the cap suggested at the pretrial conference. Initially, at least, judges continued their pre-guidelines practice of conducting pretrial status hearings. Thus, it appears that during the first half of the guidelines test, the major change in the plea negotiation process originated with the State's Attorney. In all cases, the state adopted a policy of recommending the guidelines sentence and a plea to the first count of the indictment. In fact, assistants were not permitted to recommend a sentence other than the guidelines range. According to the State's Attorney, this policy was adopted to "support" the guidelines. This did not meet with a very favorable response from defense counsel, and judges were also reported to disapprove of this policy, charging that the prosecution was using the guidelines as a negotiation tool. The central staff also expressed considerable concern, noting that the prosecutor's policy was more likely to hamper the guidelines effort than to help it.

Over time, prosecutors themselves became increasingly dissatisfied with the guidelines and their own policy of recommending the guidelines ranges. Although no single reason seems to explain this development, respondents cited increases in trials (since defendants were unwilling to plead unless a "better deal" than the guidelines sentence was offered), and dissatisfaction with the guidelines ranges prescribed for several offenses. As a result, in January 1982 the State's Attorney revised his policy: Assistants would no longer be bound to recommend the guidelines ranges, and in fact would not be permitted to provide a sentence recommendation until all the facts on

the case were available--usually at the sentencing hearing.[13] Also, during this period the pretrial status conference procedure fell into disuse, with both the defense and prosecution believing that it served little purpose given the respective positions of the two offices.

Through the discontinuation of the pretrial status conference and the elimination of the state's participation in sentence recommendations before entry of a guilty plea, sentence bargaining in Prince George's County seems to have changed in form and perhaps decreased in frequency. Among the specific changes noted by respondents in Prince George's County were the following:

- Several reported that the terms of the plea were no longer discussed at a formal hearing, but informally in the judges' chambers, at the motions hearing, or on the morning of the trial. This may also mean that many pleas were entered later in the process, since the pretrial conference hearing was not generally available.

- A greater number of judges appeared to be involved in providing some indication of sentence. No longer were all cases passed by only three pretrial conference judges.

- Several individuals noted judges' increased willingness to discuss terms directly with the defense, and to provide an agreement to a sentence--often the guidelines range. Prosecutors noted that this often meant that agreements were reached without their input or consent.

- Finally, several judges indicated that the percentage of cases settled through plea negotiations had generally decreased in the county.

Thus, as in West Palm Beach, the experience of Prince George's County demonstrates some important lessons concerning the impact of guidelines. In both sites, the prosecution adopted new policies which had a major disruptive effect on the operation of plea negotiations. It is almost a question of "chicken and egg" to ask if the disruption was more the result of the guidelines or of the prosecution policies. Indeed, it is difficult to determine the extent to which the new policies were actually "caused" by the introduction of guidelines. In either case, however, it is clear that in the face of judicial reluctance to sentence above the guidelines (and defenders' general agreement with the guidelines ranges) new prosecutorial policies may be implemented at the risk of introducing other unanticipated changes in plea negotiations or the guidelines process. Thus, for example, prosecutors in West Palm

Beach eventually began to recommend sentences below the guidelines ranges (and to negotiate scores, yielding lower sentences), while the prosecution in Prince George's County was faced with a reduction in the number of plea negotiations and increased judge/defense cooperation in determining plea arrangements.

As noted earlier in this Section, many respondents anticipated that the guidelines might change the percentage of cases going to trial. To some extent, we have already touched upon this question in the discussion of West Palm Beach and Prince George's County, where increases in trials were frequently reported. However, through our written survey of judges, prosecutors, and defense attorneys in the two states, we obtained information on the effect of the guidelines on the number of trials in all test jurisdictions. Figure 6-4 below displays the results. Clearly, these responses confirm the picture developed above. In only two jurisdictions--West Palm Beach and Prince George's County--were increases in trials reported with any consistency. In other jurisdictions where the guidelines were integrated less painfully into the accepted procedures for plea negotiations, no increase in trials was generally noted.[14] It is interesting to note that data collected by the Maryland guidelines project also tend to confirm these general observations.[15]

6.2.3 Implications of the Guidelines for Plea Negotiations

What are the major implications of the guidelines for plea bargains? Just as important, what are the impacts of plea bargains on the guidelines? We will attempt to answer both these questions by examining five key dimensions of the sites' experience: degree of judicial commitment to the guidelines; the prescribed sentence ranges; local caseload; the scoring mechanism and factors; and policies of the prosecution.

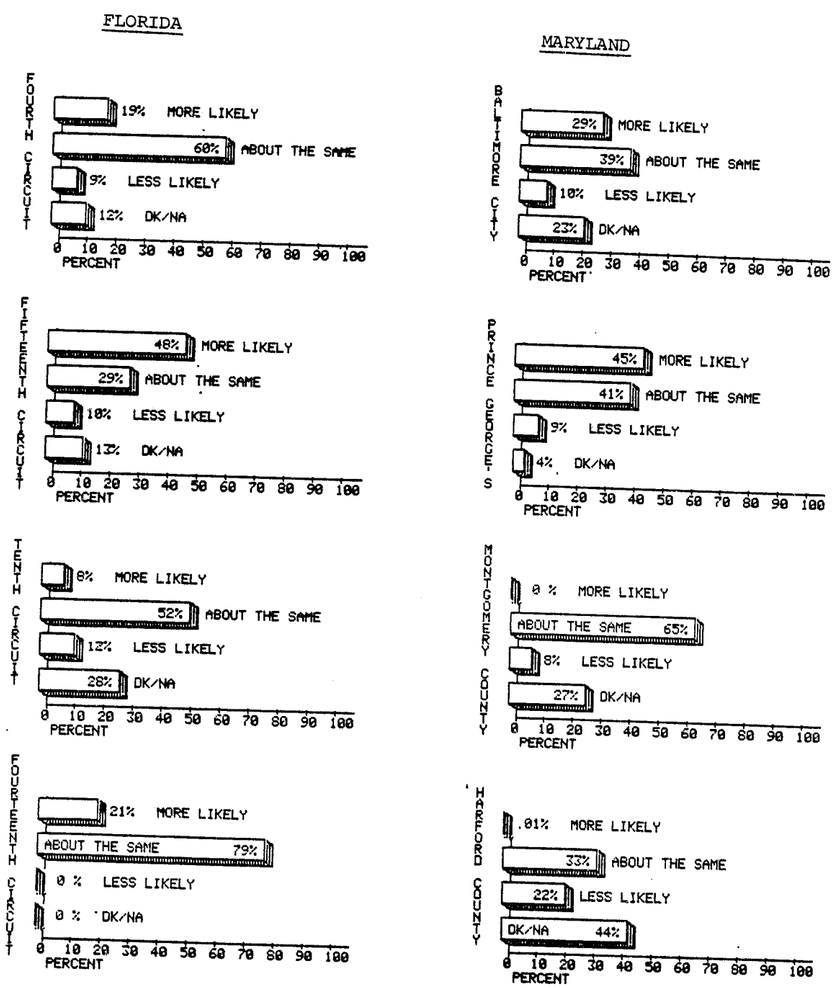
Degree of judicial commitment. After the initial period of adjustment, judges in most of the test jurisdictions expressed considerable commitment to the guidelines test, although many still expressed reservations about the guidelines themselves. In addition, judges were usually very aware of the guidelines ranges, and especially during the initial half of the test, were extremely cautious about exceeding the recommended sentence. Whether this was due to agreement with the ranges or reluctance to justify the sentences in writing is at this point irrelevant; in practical terms, judges' reluctance to exceed the guidelines was one of the major influences on plea negotiations.

As noted above, without judicial commitment we cannot expect guidelines to have much of an impact on plea negotiations, much less on sentencing in general. The experience of Jacksonville amply demonstrates the overriding effect of the judges on guidelines adaptation in the sites. In the absence of judicial attention, the guidelines were considered irrelevant to the case disposition process. Given judges' adherence to the guidelines ranges, however, a variety of impacts is possible.

FIGURE 6-4

Questionnaire Responses by Test Site for Florida and Maryland:

"Since the implementation of guidelines, are cases more or less likely to go to trial?"



Data Set: Questionnaire responses

Base: All responses. (For Florida: judges = 31; prosecutors = 45; defense attorneys = 37. For Maryland: judges = 45; prosecutors = 72; defense attorneys = 24. No responses received from defense attorneys in the 14th Circuit (FL) and Harford County (MD).

One of the most common effects we have noted is that the guidelines begin to serve as reference point in the negotiations largely by functioning as a "sentence cap" for pleas. Prosecutors often complained that the guidelines served as a "new maximum sentence," effectively constraining the sentences for plea bargains. The data on guidelines sentences support the prosecutors' assertion: in the four Florida test sites, just over 90 percent of the sentences were within or below the recommended ranges. Maryland had a similar experience, with almost 93 percent of the sentences falling in or below the guidelines ranges. Whether the sentencing matrices were used willingly or unwillingly, their influence was felt throughout the plea negotiation process.

Sentence Ranges. Even with judicial commitment, however, it appears that the ranges specified in the guidelines play a crucial role in determining the degree of adaptation to the guidelines. For example, in most jurisdictions it seems that the defense was not dissatisfied with the ranges, although they consistently noted particular factors or situations in which they felt the guidelines sentences were inappropriate. Given that these sentences were at least "workable," there were relatively few refusals to plead guilty under the guidelines. Many judges and prosecutors in Florida complained about the empirical basis of the guidelines, noting specifically that inclusion of plea negotiated cases in the analysis would skew the guidelines towards the low end of the scale. However, it appears that this may have served to compensate for the project's decision not to include any specific consideration of plea versus trial in the guidelines themselves. In Maryland, the guidelines were only loosely based on past sentencing patterns (see Chapter 4). However, the width of the ranges and the high degree of overlap between cells may have served the same function of ensuring that the guidelines did not prescribe sentences too divergent from past practice in plea negotiated cases.

Had plea negotiated cases been excluded from the analysis in guidelines development, it is likely that the sentence ranges would have been much more severe than current practice. The potential effects of such a choice are illustrated to some degree by the experience of West Palm Beach. Defenders in that county appeared to find the Florida guidelines ranges excessive for plea negotiated cases, and initially increased the number of cases brought to trial. As one defender noted, the logical approach at this point might have been to bargain charges, and it would not have been too surprising had this occurred in West Palm Beach. However, the prosecutor's introduction of the no charge bargain policy effectively precluded this option. Instead, we see two other adaptations: sentencing below the guidelines in a large number of cases, and score bargaining. As noted earlier, it is not possible to predict what might have happened had score bargaining been prevented in West Palm Beach, but it is likely that the other responses--sentences outside the guidelines, trials, and charge bargaining--might have been used much more frequently. The importance of making explicit arrangements in the guidelines for plea negotiated cases or of having guidelines sentences which are acceptable for the majority of dispositions--that is, plea bargains--cannot be underestimated.[16]

Caseload. Complicating the issues of commitment and appropriateness of the guidelines ranges is court caseload. In itself, it does not appear that the guidelines will necessarily "fail" in high volume courts. For example, Baltimore City judges filed roughly the same percentage of scoresheets as participants in other Maryland jurisdictions, and complied with the recommended ranges as frequently as other judges (See Figure 7-6 and Figure 7-7). However, caseload pressures were often cited by Jacksonville judges as a reason for their noncompliance with the guidelines, especially in plea bargained cases. In addition, since high volume courts are the least likely to be able to absorb significant changes in the trial volume, they may, like West Palm Beach and Jacksonville, be the most likely to avoid such increases by noncompliance or adaptation.

The most time consuming feature of scoresheet preparation is gathering information on the prior criminal history of the defendant. In sites where PSIs are ordered as a matter of routine, it will pose no overwhelming burden to translate this information to a guidelines score. However, in high volume courts where PSIs are rarely ordered, any new information requirement will slow the process.

A partial answer to the dilemma of high volume courts may then lie with the scoresheet design. Use of complex formulas for prior record computation increases preparation time with relatively minimal benefit in terms of sentencing. In addition, inclusion of social history factors which can only be assessed through background checks will also increase scoring time. The fact that Maryland guidelines could be scored by the judge or his clerk probably facilitated their use in that state--even in the higher volume courts. In the long run, however, improved access to and accuracy of disposition records and a shift of resources and duties among probation staff will probably be required.[17]

Finally, creating incentives to use the guidelines through adequate monitoring and enforcement may be yet another factor in ensuring guidelines use in high volume courts. Still, without other measures to speed processing, the problems of delay due to guidelines may ultimately outweigh their benefits.

Guidelines Design. As noted earlier, it appears that certain features of the guidelines scoring process can affect the use of guidelines in plea negotiations. For example, changes in filing procedures are discouraged when the guidelines scoring is based on counts at conviction. On the other hand, inclusion of ambiguous or unverifiable factors in the guidelines offers an opportunity to manipulate the scoring and thereby the sentences. In a similar vein, permitting attorneys to score the guidelines without adequate validation procedures is an invitation to abuse. In the words of one prosecutor, "if the attorneys are doing the scoring, there will be score bargaining."

Prosecutor Policies. Finally, it is worth noting that the policies of the prosecutor can exert an enormous influence on the shape of plea negotiations in general and on their use under guidelines in particular. In two sites--West Palm Beach and Prince Prince George's County--prosecutorial reactions to the guidelines and adjustments to plea negotiation policies upset a balance already jolted by the introduction of guidelines. As observed in Chapter 5, there are few enough incentives for guidelines participation on the part of prosecutors; education and involvement of the prosecution in the development and implementation of the guidelines still appears to offer one source of protection against unanticipated and potentially disruptive policy changes.

6.3 Sentencing Considerations

Although plea negotiations are one of the most significant considerations in exploring the effect of the guidelines on the felony disposition process, other factors are also important. For example, the information available to the judge generally exerts a considerable influence over the sentencing decision. In most of the Florida and Maryland test sites, presentence investigation reports are ordered for the large majority of felony cases, while statements of the defense and prosecution and special reports on the defendant or the victim may also be provided. Under a guidelines system, judges are presumably given yet another information source: the scoresheet containing information on the factors considered in the guidelines and the resulting scores. Thus, the first aim of this section is to examine the effect of the guidelines on other kinds of information usually provided to the judge. Among the hypotheses suggested by respondents were the following:

- Guidelines will cause a shift in the preparation of presentence investigation reports: more emphasis will be given to those factors included in the guidelines, and the probation and parole recommendations contained in the PSI may reflect the guidelines sentence.

- Fewer presentence investigation reports will be ordered, since the guidelines will provide much of the same information in a briefer format.

- Presentence reports may be ordered more frequently, since these reports provide information necessary for the guidelines scoresheets and since scoresheets are prepared in tandem with them.

- Reliance on other sources of information will decrease, as judges refer primarily to the information considered by the guidelines.

The second goal of this section is to assess shifts in the influence of nonguidelines factors on sentencing. We recognize that sentencing dispositions can also be affected by a number of considerations external to the case at hand. For example, knowledge of prison conditions or local jail populations can cause judges to modify their choice between incarcerative and nonincarcerative options. Awareness of parole release can exert some effect on the sentence, as judges factor the likely time served into their decisions. Finally, sensitivity to public attitudes and concern about public image may also affect judges' decisions. Since under a guidelines system one might expect that consideration of these factors would decrease, we have attempted to identify which ones were important before the guidelines and whether judges continued to consider them after the guidelines implementation.

6.3.1 Changes in Presentence Investigation Reports

The degree to which the information provided in the guidelines may supplant or influence other types of information depends, of course, on the degree to which judges consider the guidelines in their sentencing decisions. We have seen that for burglary cases, the percentage of scoresheets filed with the Florida project staff was fairly modest in the two rural sites, high in the suburban site, and low in the urban site. Thus, with the exception of Jacksonville, there should at least be the possibility that guidelines might influence presentence investigation report preparation or other information sources considered by the judge. In Maryland, the use of the guidelines is less clear: although many judges did not report consulting the guidelines before determining the sentence they would impose, it would appear that the guidelines information was available to them and was consulted at some point in the majority of cases. Thus, we might still expect to see some influence on the Maryland presentence reports.

As noted above, the first level of change we might expect to see in presentence investigation reports is a decrease or increase in their use. Both before and after guidelines implementation, we asked judges, defenders, prosecutors, and parole and probation staff in each jurisdiction to estimate the percentage of cases on which PSIs are typically ordered. Estimates varied widely, and many appeared to be overly optimistic. However, before the introduction of guidelines, Jacksonville and West Palm Beach reported low percentages of PSIs (30 to 50 percent) while the two rural Florida jurisdictions generally reported much higher percentages, ranging from 70 to 90 percent. In addition, respondents in Jacksonville and West Palm Beach stressed that PSIs are almost never ordered on plea bargained cases. Before implementation of guidelines in Maryland, a similar pattern was noted. Baltimore respondents reported that as few as 50 percent of the cases received PSIs and noted that they are rarely ordered on cases where a disposition is arranged in the plea negotiation, while the other three jurisdictions all estimated that PSIs were prepared for 90 percent of the cases.

At the close of the guidelines period we again asked respondents to estimate the percentage of PSIs ordered and to assess whether this had changed during the test year. Jacksonville respondents noted a slight decrease, but indicated that this was largely because one judge had modified his policies: prior to the guidelines he ordered PSIs on most cases, but during the test year he ordered PSIs only after trials. However, a representative of Probation and Parole noted that some judges had begun to use the guidelines scoresheets as "replacement PSIs." Given reports of guidelines usage in this city, we would not expect much change in the Jacksonville system; it seems that the observed decrease may reflect changes (such as the judge's shift in policy) generally unrelated to the guidelines. In West Palm Beach, judges noted no decrease in the percentage of PSIs ordered, but the Parole and Probation agents noted a dramatic decrease. According to one, the percentage dropped by about half. Although we cannot be sure, it appears that this change may in fact be related to the guidelines. Given attorneys' willingness[18] to complete the guidelines when there was no PSI, and given that scoresheets could be prepared by them immediately, without the 30-day wait typically needed for PSI completion by Parole and Probation, it is very possible that judges turned to the guidelines in lieu of PSIs to speed case dispositions.

In the rural Florida jurisdictions, estimates of the percentage of PSIs ordered were virtually identical to the pre-guidelines estimates, and most individuals said they observed no change in the number of reports ordered. It seems likely that the lower case volumes in these courts may have enabled judges to continue their past practice of ordering presentence investigations on a high percentage of cases. Thus, there was little incentive to use the guidelines as an alternative information source, and indeed some disincentive, since judges or attorneys would have to complete many more scoresheets if fewer PSIs were ordered.

Maryland's experience appears to be far more mixed. In no site was there a clear consensus that the volume of presentence investigation reports had changed since the introduction of guidelines, and percentage estimates provided before and after the guidelines test were roughly comparable. However, in the two larger jurisdictions--Baltimore and Prince George's County--some probation agents did note a decrease in the percentage of cases receiving PSI reports. In light of respondents' comments that judges often determined the sentence and then consulted the guidelines, it seems unlikely that guidelines scoresheets served as a substitute for other information sources; decreases in PSIs in these two sites may therefore be due primarily to general caseload pressures.

Although it appears that in most sites judges did not appreciably alter their use of the presentence investigation report, it was clear that Probation and Parole agents made a number of adjustments to the guidelines. With the exception of Jacksonville, probation agents often indicated that they had given increased emphasis to investigating information items necessary for scoring the guidelines--specifically, dispositions of prior criminal offenses. A number commented that as a result, judges

were getting better information on the criminal record of the defendant. In addition, most agents indicated that the guidelines had exerted a subtle influence on the sentencing recommendations included in the PSI. Whether consciously or unconsciously, it appears that the PSI recommendations were often shaped by the recommended guidelines range.

Finally, it is worth noting that one significant procedural modification in the presentence investigation process occurred in Harford County as a direct result of the guidelines. Recognizing that defendants willing to plead guilty would often want some indication of the likely sentence prescribed by the guidelines, judges began to order "pre-plea investigations" from the Division of Parole and Probation. Keyed specifically to the elements needed to score the guidelines, these brief reports were frequently ordered when parties believed that the defendant might offer a plea. The only other jurisdiction to use a similar procedure was Florida's 14th Circuit, but it was noted that use of the pretrial investigation was not encouraged there. One can surmise that other jurisdictions might also have instituted this procedure had not attorneys been able to estimate (and in West Palm Beach, negotiate) the guidelines score for a plea bargain, and had their probation departments been able to accommodate the increased workload this would entail.

By all accounts, it appears that there was very little change in the use of information sources other than the presentence investigation report. Both before and after the guidelines, judges noted that the predominant sources included information from the prosecution and defense, and criminal histories provided by law enforcement agencies. However, it is interesting to note that both before and after the guidelines, Maryland judges differed from their Florida counterparts in citing additional sources such as psychological evaluations and special reports on sentencing alternatives arranged by the defense.

6.3.2 Other Factors Considered in Sentencing

As noted earlier, factors other than those relating directly to the case at hand can also affect judges' sentencing decisions. Examples include knowledge of parole release policies, awareness of conditions at state and local correctional facilities, and public opinion. The legitimacy (or at least advisability) of considering these factors in sentencing is a matter of some debate which we will not review here. Suffice it to say that if such factors are given different weights by different judges, they may constitute a source of sentencing variation. With this in mind, before and after the guidelines implementation we questioned judges and other respondents about the degree to which judges consider these external factors.

We observed earlier in Chapter 5 that in both states, the parole system engendered much hostility and ill will from judges, who felt that their sentences were "meaningless" in the face of the "resentencing" done by the parole board. As a response to the perceived leniency of the parole commission, many judges in both

states indicated that they "factor in" the probable parole release date when imposing sentence, in essence deciding how long they want the offender to serve and then inflating the sentence to allow for parole. However, this practice was far from uniform among sites or among judges in a single site; some respondents rightly pointed out that as a result, parole consideration was in itself a source of disparity.

Introduction of the sentencing guidelines appears to have done little or nothing to allay judges' concern about parole or to unify their response to the system. In virtually every site, some judges reportedly inflated their sentences to account for parole, while others claimed that they ignored parole when sentencing. Based on these reports, we can only conclude that parole consideration remained as a source of disparity during the guidelines period.

In many respects, consideration of prison and/or jail conditions is as problematic as considering parole: unless all judges maintain the same policy, some defendants might be spared incarcerative terms solely due to crowded conditions while similarly situated defendants may be given incarcerative sentences. In both Maryland and Florida there was ample reason for judges to be concerned about crowding at state facilities both before and during guidelines. For example, by March 1981, two state penitentiaries in Maryland and the entire state prison system in Florida were under court order to remedy crowding and substandard confinement practices.[19]

Judges' responses to these conditions was quite varied. During both time periods, some judges suggested that they increased their use of nonincarcerative sentences or the use of local jail facilities in view of the crowded prison system. It was particularly common for judges and others to report that prison conditions were an important consideration when sentencing young first offenders, both because of their special vulnerability and because judges were aware that commitment of a new offender could mean earlier parole for a more serious offender. However, just as frequently judges asserted that they refused to consider prison conditions. In the words of many, this was the legislature's concern, not the courts'. Clearly, since the guidelines offered no guidance on this issue, another important source of variation remained unconstrained.[20]

It is interesting to note that respondents did, in fact, report some general trends over the two-year period in which our interviews were conducted. In the period before guidelines, crowded conditions at the state facilities often led judges to make increased use of the local jails. Some 18 months later, six of the eight sites specifically reported that they were now faced with severe crowding at local facilities, and that judges were very concerned with maintaining their jail populations within acceptable bounds. Institutional populations at the state level also continued at record levels.

Finally, during the interviews conducted at the close of the guidelines test, we asked if there had been any other significant factors affecting sentencing. This question was intended only to assist us in interpreting the quantitative data, since

similar questions had not been asked in the pre-guidelines period. However, since responses within each state were quite similar and reflect some interesting trends, we report them here.

Overall, respondents in virtually every site noted increased pressure for harsher sentences, particularly for the crimes of burglary and robbery. Not surprisingly, much of this attention was focused by the media, and several individuals reported the significant (even alarming) effect of media coverage on judges' sentences for these and other crimes. Jacksonville offers one of the most telling examples: one newspaper in that city publishes the dispositions of all Circuit Court cases every afternoon, and periodically ranks judges by severity of sentence for several key crimes. As a result, one judge reportedly increased his sentences for burglary, moving from least severe to most severe in the newspaper's ranking. Other sites, including West Palm Beach, Baltimore, Montgomery County, and Harford County also cited media pressure as a major concern for judges. The effect of such publicity on judges' sentencing behavior cannot be underestimated, especially since judges and prosecutors in both states are elected, not appointed. In fact, 1982 was an election year for Maryland prosecutors and for some judges as well. Public and media pressure for harsher sentences thus was a potent sentencing consideration for many. It is interesting to note that the empirical sentencing guidelines model provides for changes in judges' sentences due to such concerns, by allowing judges to impose extra-guidelines sentences on this basis and ultimately, by factoring these sentences into the guidelines revision process.

6.3.3 Implications of the Guidelines for Sentencing Considerations

Generally, it would seem that the guidelines' primary impact on sentencing considerations is procedural. In some of the higher volume courts, guidelines may serve as an alternative source of information, replacing the use of the more detailed but less timely presentence investigation report. If PSIs had generally been ordered in the past, a shift to reliance on the scoresheet could result in a decline in the quality of information available to the judge, since the accuracy of the guidelines scoresheets information may be open to question. However, in some courts where judges typically sentenced without PSIs, the guidelines may actually make more information available to the judge.

Some adjustments in PSI preparation were noted, including greater emphasis on verification of the information items needed for scoring and incorporation of the guidelines sentences in the recommendations of the investigating agent. Improved verification of criminal history information may have added to probation agencies' workload pressures; however, some respondents reported that the benefit of improved criminal history information on many cases outweighed this drawback.

Finally, it appears that consideration of factors external to the instant case continued to vary by judge. As noted earlier, responses to parole, prison populations, and public opinion are all potential sources of variation in sentencing when considered by some judges and not others.

6.4 Uniformity and Severity

Until this point, we have looked primarily at those activities leading up to the sentencing decision. In this section we shift our focus, examining reports of the guidelines' influence on two distinct concerns: uniformity of the sentencing decision and sentencing severity. As in other sections of this Chapter, one primary information source is interviews conducted both before and after the guidelines, and questionnaires administered to judges, prosecutors, and defense attorneys during the guidelines test.

In previous sections, we could place a fair degree of reliance on these results, at least in the sense that respondents were describing their own experience with the guidelines or reporting on changes which they might reasonably be expected to observe. Here, however, many of the observations must be interpreted with caution, particularly as we explore changes further removed from the daily business of the courts. In many respects, it is useful to view the qualitative results presented in this section as an investigation of changes in perceptions rather than changes in process. Given the substantial influence of such perceptions on the operation and acceptance of the guidelines, this may be a useful inquiry in itself. However, interview data are not our only source of information on severity and uniformity of sentencing under the guidelines. Chapter 8 examines these questions in detail and presents our empirical findings of guidelines impact on these concerns. This Chapter summarizes our findings on guidelines impact, based on our analyses of comprehensive data collected on all burglary cases sentenced before and during the guidelines test year, in all test sites and in a matched set of jurisdictions which did not participate in the guidelines experiment.

6.4.1 Sentencing Uniformity

The avowed goal of sentencing guidelines is to reduce unwarranted variation in sentences, while still allowing judges to respond to the specific exigencies of individual cases. Ideally, of course, we would observe increases in sentencing uniformity after the introduction of guidelines, accompanied by respondents' perceptions of an increase in sentence uniformity. However, it is entirely possible that the two would operate independently: that is, for example, that respondents hostile to the guidelines might fail to observe increased uniformity where it occurred, or that supporters might cite an increase when in fact sentences remained the same. In practice, how did this operate in Maryland and Florida?

In the written questionnaire administered during the test year, we asked judges, prosecutors, and defense attorneys to assess the degree to which "unwarranted disparity" had changed after the introduction of guidelines. Their responses, shown in Table 6-1 below, seem to demonstrate that in both states, the majority of respondents believed that the amount of disparity was about the same as before the guidelines. Interviews conducted at the close of the guidelines test year produced the same pattern of results.

Two observations on the questionnaire responses are worth making, however. First, there are a large number of nonresponses. In part, this is due to those who professed to have no knowledge or opinion. However, a substantial number of respondents took umbrage at the term "unwarranted disparity" used in our questionnaire, asserting that it was not valid to assume that unwarranted disparity existed. Given this opinion, it seems that use of a less controversial term might have prompted these individuals to respond that sentences before and after the guidelines demonstrated about the same amount of variation, avoiding the question of its sources or legitimacy.

Second, it appears that in some jurisdictions, a substantial minority or even a slight majority of the respondents did note more uniform sentencing. This is particularly noticeable in Florida's 10th and 15th Circuits and in two Maryland jurisdictions: Montgomery County and Harford County. However, only in the 10th Circuit was there a general consensus. A substantial number of judges, prosecutors, and defense counsel all believed such a reduction had occurred. In the 15th Circuit, only judges consistently asserted that disparity had decreased. Defense and prosecution were far less optimistic in their responses, and one questioned whether a real decrease was possible, given the score bargaining that was prevalent in that jurisdiction. In Montgomery County, prosecutors and some defenders observed an increase in uniformity, while judges were far less certain. A similar trend is found in Harford County: judges indicated no opinion or knowledge on this issue, while prosecutors reported a reduction in disparity. To some extent, it would appear that perceptions of more uniform sentences are most prevalent in groups supporting the guidelines; we cannot tell if their observations of increased uniformity led them to support the guidelines, or whether their favorable opinions of the guidelines shaped their observations of guidelines impact.

How do the observations of judges, prosecutors, and defense attorneys in the test sites compare with our own empirical findings? In Florida, respondents' observations corresonded closely with our own. Based on the analysis of burglary cases, we initially found a slight increase in uniformity. However, once we accounted for case differences (such as seriousness of the defendant's prior record) there was no increase in uniformity. Thus, the empirical data, like the interview data, show no consistent move towards increased uniformity in Florida.

Table 6-1

Perceptions Concerning Sentence Disparity Under Guidelines: Florida and Maryland Test Sites[a]

BY TEST SITE

Florida	4th Circuit (Jacksonville)	15th Circuit (West Palm Beach)	10th Circuit (Bartow)	14th Circuit (Marianna/Panama City)	All Test Sites
Reduced	7%	35%	44%	7%	23%
About the same	67	42	32	50	50
Increased	2	3	0	7	3
Don't know	23	19	24	36	24
Total	100	100	100	100	100
Number of Respondents	(43)	(31)	(25)	(14)	(113)

Maryland	Baltimore City	Prince George's County	Montgomery County	Harford County	All Test Sites
Reduced	23%	20%	38%	33%	26%
About the same	44	50	35	11	42
Increased	10	5	0	11	6
Don't know	24	25	27	44	26
Total	100	100	100	100	100
Number of Respondents	(62)	(44)	(26)	(9)	(141)

BY RESPONDENT GROUP

Florida	Judges	Prosecutors	Defense	All Respondents
Reduced	39%	9%	27%	23%
About the same	19	62	62	50
Increased	0	7	0	3
Don't know	42	22	11	24
Total	100	100	100	100
Number of Respondents	(31)	(45)	(37)	(113)

Maryland	Judges	Prosecutors	Defense	All Respondents
Reduced	24%	26%	25%	26%
About the same	22	51	50	42
Increased	0	10	8	6
Don't know	53	13	17	26
Total	100	100	100	100
Number of Respondents	(45)	(72)	(24)	(141)

Data Set: Guidelines period questionnaires

Base: All respondents (N=125 Florida, 141 Maryland)

a. Responses to the question, "Since the implementation of guidelines, has unwarranted disparity been reduced or increased?"

In Maryland, analyses of the burglary data showed a modest increase in uniformity under the guidelines. However, this effect was due to the change in only one site: Baltimore City. While sentences became less variable in Baltimore, no such change was observed in the other three sites. Thus, the perceptions of the Montgomery County and Harford County respondents were not supported by our quantitative analyses.

6.4.2. Severity

The effects of guidelines on sentence severity (and their possible role in stabilizing recent trends towards increased severity) were not a major issue in the design of the Maryland and Florida guidelines, although severity was raised at some points during the development process. For example, staff of the guidelines project in Florida expressed some concern about modifying the factor weightings derived from their analysis, for fear of unduly increasing sentence severity and thereby increasing prison populations. Both states probably relied to some extent on the empirical nature of the guidelines to ensure that the guidelines would do no more than continue past levels of severity. How did the guidelines actually influence sentence severity in the two test sites? Again, we can examine both empirical evidence and respondents' perceptions on this question.

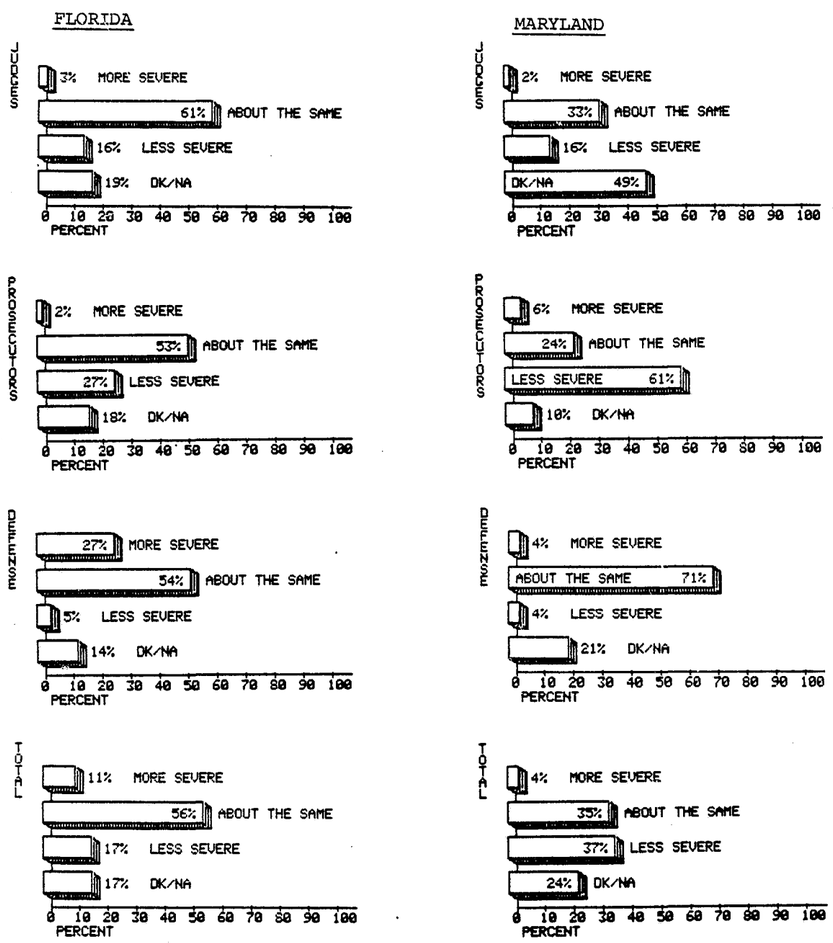
Based on our analysis in Florida, net sentences for burglary were more severe during the guidelines period, although some of this may be due to changes in case characteristics over time. When we developed models which controlled these characteristics, we were able to predict the sentences identical cases would be given both before and during guidelines sentencing. In general, severity increased in the 4th Circuit; both increases and decreases in severity were found in the other test sites. In Maryland, the pattern from our empirical findings was more striking: while burglary sentences increased somewhat during the year in test sites, the rate of increase was greater in the comparison sites. When we controlled for case characteristics and used these equations to predict sentence severity for identical cases, the test sites exhibited varying patterns--some showed more severe sentences and others less severe sentences than before guidelines introduction.[21]

When we examine these findings in light of the interview results, the picture becomes somewhat more complex. On the one hand, interviews with judges, prosecutors, and defense attorneys in both states would suggest that sentences may have become more severe: many noted that sentences for all crimes (and particularly burglary and robbery) were increasing as public concern and pressure escalated. However, when we asked these same individuals about the influence of guidelines on sentence severity, the picture becomes far less clear, as shown in Figure 6-5. Most respondents in both states generally believed that sentences stayed about the same. However, a sizeable minority suggested that sentences had become less severe, and in interviews many attributed this directly to the guidelines. How do we reconcile this with our empirical findings and the other interview responses?

FIGURE 6-5

Questionnsire Responses by Judges, Prosecutors, and Defenders in Florida and Maryland:

"Overall, have sentences become more severe or more lenient since the implementation of guidelines?"



Data Set: Questionnaire responses

Base: All responses. (For Florida: judges = 31; prosecutors = 45; defense attorneys = 37. For Maryland: judges = 45; prosecutors = 72; defense attorneys = 24. No responses received from defense attorneys in the 14th Circuit (FL) and Harford County (MD).

Most likely, it would appear that the difference stems primarily from the widely divergent perceptions of the different groups responding to our questionnaire. As Figure 6-5 shows, prosecutors were most likely to believe that guidelines sentences were less severe, while defenders sometimes cited an increase in severity. These opinions are certainly in line with the traditional adversarial positions of these two groups, and it is likely that these findings reflect these roles more than any overall changes in severity resulting from the guidelines.[22]

6.5 Conclusions

From a one-year test period, it is difficult to determine how guidelines implementation will ultimately fare. The initial year was marked by periods of rapid change and adjustment; development and refinement of a complex set of procedures; and the tensions and time pressures typical of any new effort. From this, one can obtain a detailed picture of the problems involved in implementing a new program. More difficult, but perhaps more rewarding, is the attempt to determine implications of the initial test year for guidelines operations in general.

Given the role of the prosecution in shaping sentences through the initial filing decision, it is plausible that filing policies would be profoundly affected by the introduction of sentencing guidelines. However, during the first year of implementation we saw little evidence of changes in filing policies directly attributable to the guidelines. In part, this may have been due to the limited period of the test: given more time and experience with the guidelines, prosecutors may have adapted filing procedures to obtain a better bargaining position or to increase potential penalties. Still, it appears that scoresheet design played some role in discouraging adaptation in charging practices. By focusing on charges at conviction, treating like offenses in the same sentencing matrix, and basing the scoring on the sentencing event, the likelihood of successful charge manipulations was somewhat diminished.

Concerns about delays due to scoring was a major problem in the test sites, especially in high volume courts where a large number of plea bargained cases are sentenced without reference to presentence investigations. Permitting attorneys to score the guidelines in plea bargained cases offers one means of speeding the scoring process. However, without verification of the resulting scores, the guidelines can be subverted through negotiations on scores or by legitimate errors made by the attorneys. At a minimum, it is important to reduce the ambiguity of the scoring factors, both to reduce errors and to minimize the opportunity for "interpretation" during negotiations. However, the longer term answer may be to minimize or eliminate attorneys' role in preparing the final scoresheet.

Instead, we would suggest a fuller role for the probation and parole agents. By virtue of their position, they may be more objective than either the prosecution or defense. In addition, they have the advantage of established access to information on criminal histories and current offenses. However, use of the probation department cannot be accomplished without some shifts in their priorities and resources. Since guidelines scoresheets generally reflect the same kinds of information contained in a presentence investigation, shifting the emphasis from PSI preparation to scoresheet preparation would be one alternative that might result in only a small decrease in information available on many cases.

Due to the importance and pervasiveness of plea negotiations as a disposition mechanism, it is also likely that the success of guidelines will be determined by the extent to which (and the ways in which) pleas are accommodated under the guidelines. As Rich et al. noted in discussing plea negotiations in the pilot implementation sites, "the judge's dependence on the discretionary actions of the prosecutor and defendant, if not adequately taken into account, may frustrate the reformers' attempt to structure judicial sentencing discretion through guidelines."[23]

Of course, one reaction of judges and attorneys might be to ignore guidelines for plea negotiated cases. Although this has been legitimized in one jurisdiction,[24] it raises serious problems of equity and the viability of guidelines as a disparity reduction mechanism. Clearly, to have any effect, the guidelines must include plea bargains, which make up the majority of the criminal caseload in most courts. Given this necessity, what are the implications for guidelines operation? First, designers must recognize that guidelines <u>will</u> affect plea negotiations, and they must plan specifically for this fact. Second, procedures for ensuring the integrity of the guidelines process in plea negotiations must be established.

Typically, incentives for guilty pleas are created through two related mechanisms: explicit or implicit discussion of the sentence to be imposed, and explicit manipulation of the conviction offense (and hence the sentence). Since guidelines constrain judges' sentencing discretion, attorneys' ability to obtain "sentence bargains" is apparently reduced. Incentives to plead guilty might therefore be diminished. Nevertheless, plea negotiations did not seem to decrease during the initial year of guidelines in the two states. In part, this may be because the form but not the function of sentence negotiations changed under the guidelines. Judges rarely sentenced above the recommended range. In addition, the guidelines sentence ranges were generally acceptable to the defense. Thus, pleading guilty with the understanding that the guidelines sentence would be imposed offered two of the same benefits as previous methods of negotiating for sentences: certainty and a reasonable sentence.

Had the guidelines ranges been much more severe than past practice, such a procedure might not have worked. However, with moderate ranges, the incentives to plead guilty were not actually reduced under the guidelines. Whether moderate

sentence ranges are derived empirically, by including plea negotiated cases in the construction sample used to develop the guidelines, or normatively, by developing ranges which are necessarily constrained by external limitations such as prison capacities, it is clear that the ranges must be "reasonable" for a majority of cases.[25]

Reasonable ranges do not necessarily speak to the other approach to inducing pleas: charge negotiations. Increased use of charge negotiations has often been suggested as a likely adaptation to sentencing guidelines. Charge bargaining under guidelines is difficult to control and will certainly have a detrimental effect on sentencing uniformity. As one member of the Minnesota Guidelines Commission noted:

> Some pleas agreements, however, do pose a threat to the Guidelines' ideal in that they move the defendant to a different box in the grid or invoke a sentence less than the presumed sentence. To prohibit all such concessions would be unrealistic. Those who contend that the prosecutor should never reduce a charge presume an infallibility regarding the charged offense. A prosecutor may have incomplete information while making the charge decision.[26]

Given that charge negotiations cannot realistically be prohibited, and yet can subvert the guidelines and maintain disparity, few solutions are apparent. One which has been suggested is to structure prosecutor's charging decisions through some sort of guideline.[27] This approach is certainly worthy of further investigation, although the difficulties of establishing such guidelines would appear to be substantial.

One final alternative response to the question of plea negotiations under the guidelines is to recognize the exigencies of case pressures in the courts and to develop a bifurcated system, in which different sentences are explicitly provided for cases settled by plea and trial. This has been recommended by at least one group of researchers,[28] although the potential legal and ethical questions regarding this approach were acknowledged to be substantial. The principal advantage of structuring guidelines on this basis is that it opens the issue to explicit public and legal debate. However, by establishing two standards for uniformity--one of trials and one for pleas--it mocks the concept of equity and cripples guidelines' ability to achieve true uniformity.

FOOTNOTES

1. See, for example, Kenneth Carlson, Mandatory Sentencing: The Experience of Two States, Policy Brief (Washington, D.C.: U.S. Department of Justice, National Institute of Justice, 1982); William J. Cook, "The 'Bitch' Threatens, but Seldom Bites--A Study of Habitual Criminal Sentencing in Douglas County," Creighton Law Review 8 (1975): 893-919; A.G. Fredericks, "Habitual Offender Prosecution in New Jersey," Criminal Justice Quarterly 83 (Spring 1978); H.R. Glick, "Mandatory Sentencing: The Politics of the New Criminal Justice," Federal Probation 3 (March 1979); Milton Heumann and Colin Loftin, "Mandatory Sentencing and the Abolition of Plea Bargaining: The Michigan Felony Firearm Statute," Law and Society Review 393 (Winter 1979); Malcolm M. Feeley, Court Reform on Trial--Why Simple Solutions Fail (New York: Basic Books, Inc., 1983); Hans Zeisel and Shari Seidman Diamond, "Search for Sentencing Equity: Sentence Review in Massachusetts and Connecticut," American Bar Foundation Research Journal 881 (Fall 1977).

2. Richard F. Sparks, Bridget A. Stecher, Jay S. Albanese and Peggy L. Shelly, with a chapter by Donald M. Barry, Stumbling Toward Justice: Some Overlooked Research and Policy Questions about Statewide Sentencing Guidelines (Newark, NJ: School of Criminal Justice, Rutgers University, 1982); William D. Rich, L. Paul Sutton, Todd R. Clear and Michael J. Saks, Sentencing by Mathematics: An Evaluation of the Early Attempts to Develop and Implement Sentencing Guidelines (Williamsburg, VA: National Center for State Courts, 1982).

3. Many of the items included on this questionnaire were drawn directly from the instrument developed by the National Center for State Courts for their evaluation of the early guidelines projects in Denver, Chicago, Philadelphia, and Phoenix. Copies of the questionnaire are presented in Appendix G.

4. Minnesota Sentencing Guidelines Commission, Preliminary Report on the Development and Impact of the Minnesota Sentencing Guidelines (St. Paul, MN: Minnesota Sentencing Guidelines Commission, 1982).

5. Minnesota Sentencing Guidelines Commission, The Impact of the Minnesota Sentencing Guidelines--Three Year Evaluation (St. Paul, MN: Minnesota Sentencing Guidelines Commission, 1984), pp. 71-86.

FOOTNOTES
(continued)

6. Feeley, Court Reform on Trial, pp. 118-147; Albert W. Alschuler, "Sentencing Reform and Prosecutorial Power: A Critique of Recent Proposals for 'Fixed' and 'Presumptive' Sentencing," University of Pennsylvania Law Review 126 (1978): 550-577.

7. These jurisdictions include Jacksonville, West Palm Beach, Bartow (10th Circuit), and Baltimore.

8. Although this policy change was not formally announced, many respondents indicated that this exemption had been granted by the guidelines project. Jacksonville was the only jurisdiction which reported this change, and respondents generally believed that it was instituted in response to the extreme resistance of Jacksonville participants to the guidelines and their concern that guidelines would impede plea negotiations.

9. In two sites--Montgomery County and the 14th Circuit (Marianna/Panama City)--respondents indicated that guidelines were usually discussed, but were divided on whether or not there was a change in the plea negotiation process. This may reflect that respondents did not consider these discussions as a change, or that the discussions resulted in no net modification to the process.

10. Note again, that although respondents in the 14th Circuit and Montgomery County reported that prosecutors used the guidelines in constructing recommendations, they did not generally agree that a change had taken place in the plea negotiation process. From this we may infer either that use of the guidelines in the prosecutor recommendation was less common in these sites than in others, or that such a recommendation was not viewed by many as a change in the process.

11. One question in our written survey, for example, asked respondents to comment on the appropriateness of the guidelines ranges. In Florida, 71 percent of the prosecutors found the sentences too lenient (n=45), while in Maryland, this figure was 85 percent (n=72).

12. See, for example, Rich et al., Sentencing by Mathematics, pp. 166-191.

13. According to the central project staff, this revised policy actually allowed the guidelines to operate as originally intended. Thus, although the prosecutors viewed this as a means to withdraw support, the practical effect of this move was to make guidelines use more feasible in Prince George's County.

14. Note that Baltimore offers the most ambiguous results concerning possible increases in trials. Independent examination of the responses of judges, prosecutors, and defense attorneys may shed some light on this situation. Judges and defense attorneys generally found no significant increase in the number of trials. Only 10 percent of the judges and 21 percent of the defense counsel indicated greater likelihood of trial. However, prosecutors, who opposed the guidelines with some fervor, believed that trials had increased as a result of guidelines--55 percent noted an increase. It would appear that the ambiguous finding for Baltimore City is less suggestive of an increase in trials than of a strong divergence of opinion between respondent groups.

15. Using data from their construction sample and the scoresheets collected during the test year, the Maryland guidelines project noted that overall, trials increased from 16.2 percent of all cases before guidelines to 25.7 percent during guidelines. However, most of this increase resulted from bench trials. Jury trials--by far the most time consuming procedure--actually decreased from 9.4 percent of all cases to only 7.3 percent.

16. For example, even the Minnesota guidelines call for presumptively stayed sentences in up to 80 percent of the cases. See William E. Falvey, "Defense Perspectives on the Minnesota Sentencing Guidelines," Hamline Law Review 5 (June 1982): 259.

17. Note, for example, that the Minnesota legislature made completion of scoresheets mandatory, while presentence investigations were made discretionary. Minnesota Sentencing Guidelines Commission, Preliminary Report on the Minnesota Sentencing Guidelines, p. 53.

18. Eagerness, given the importance of score bargaining.

19. Johnson v. Levine, 450 Supp. 648 (D. Md. 1978); Nelson v. Collins, 455 F. Supp. 727 (D. Md. 1978); Costello v. Wainwright, 397 F. Supp. 20 (M. D. Fla. 1975).

20. Note that some guidelines systems, such as the statewide guidelines in Minnesota and Washington State, do make prison populations an explicit consideration in that the sentence ranges are designed to maintain population levels within existing capacity. See Minnesota Sentencing Guidelines Commission, Preliminary Report on the Minnesota Sentencing Guidelines; Andrew Von Hirsch, "Constructing Guidelines for Sentencing: The Critical Choices for the Minnesota Sentencing Guidelines Commission," Hamline Law Review 5 (June 1982): 164-216; State of Washington, Sentencing Guidelines

Commission, Report to the Legislature (Olympia, WA: Sentencing Guidelines Commission, 1983).

21. One test for changes in severity due to the guidelines might have been to apply the guidelines scoring to the case data collected by the Florida and Maryland guidelines projects for construction of the guidelines, and then to compare actual sentences to those prescribed by the guidelines. Unfortunately, as noted in Chapter 4, the construction sample data do not include many of the factors considered in the guidelines, especially in Maryland. In addition, these data were often incomplete, even for those factors ultimately used in the guidelines. As a result, we did not undertake this analysis.

22. It is interesting to note that the guidelines project staff observed a similar split in the sites' reactions to the guidelines ranges. They frequently remarked that if the defense thinks the ranges are too high, while the prosecution thinks the ranges are too low, they are probably just about right.

23. Rich et al., Sentencing by Mathematics, p. 162.

24. In Massachusetts, guidelines apply only to cases disposed through trials. See Sandra Shane-Dubow, Alice P. Brown, and Erik Olsen, Sentencing Reform in the United States: History, Content, and Effect (Madison, WI: Wisconsin Center for Public Policy, 1983), p. 255. (Typewritten.)

25. Note that under the Minnesota guidelines, for example, 80 percent of the cases receive presumptively stayed sentences. See William E. Falvey, "Defense Perspectives on the Minnesota Sentencing Guidelines," Hamline Law Review 5 (June 1982): 259.

26. Stephen C. Rathke, "Plea Negotiations Under the Sentencing Guidelines," Hamline Law Review 5 (June 1982): 289.

27. Ibid., pp. 290-291.

28. Rich et al., Sentencing by Mathematics, pp. 211-215.

PART II

Part II of this report provides a detailed examination of compliance (Chapter 7) and the quantitative analyses of sentencing impact (Chapter 8). The findings of these analyses have been cited in earlier chapters; here we present their substance, including discussions of data sources and methodologies. Appendices F, H, I and J supplement the text of these chapters.

CHAPTER 7

JUDICIAL COMPLIANCE WITH SENTENCING GUIDELINES

This chapter focuses on the empirical analysis of judicial compliance with the multijurisdictional sentencing guidelines that were developed and field-tested in Florida and Maryland. Compliance refers to the actual use of the guidelines by judges sentencing convicted offenders during the test period. To what degree did judges in the test jurisdictions cooperate in the use of the guidelines? To what degree did their sentences fall within the appropriate guidelines ranges?

Compliance is a central focus in analyzing the guidelines test because it represents the link between the guidelines as designed in Maryland and Florida and the hoped-for sentencing impacts. Sentences for like cases cannot be made more uniform across different judges and jurisdictions unless judges changed their decisionmaking patterns through guidelines use. If the guidelines were consulted infrequently or only intermittently applied to eligible cases, they cannot be expected to have a significant impact. If the sentence ranges were only rarely followed, again not much impact on sentencing would result. On the other hand, judicial attention to the factors emphasized in the guidelines scoring and to the associated ranges could produce greater consistency in sanctions. Thus, analysis of compliance is essential to an overall understanding of both implementation and impact.

The concept of compliance was introduced in Chapter 5, where the discussion focused on compliance under a voluntary, judicially mandated guidelines system. There, we argued that compliance in such circumstances depends critically on building support for guidelines, accommodating local differences in norms and procedures, and monitoring the resulting patterns of use. This chapter provides the results of a detailed analysis of compliance during the one-year guidelines test. Because these findings are central to the argument of Chapter 5, readers will wish to examine the ways in which they were derived and the substantial variations revealed.

The analysis questions about compliance form two groups:

Procedural

1) Did judges consult the guidelines in reaching sentence decisions?

2) Were scoresheets completed on eligible cases?

Substantive

3) Did sentences fall within the guidelines ranges?

4) Were reasons provided for extra-guidelines sentences?

5) What kinds of reasons were given for extra-guidelines sentences?

To answer these five questions, several kinds of data are available. Information on whether and how judges used the guidelines in reaching sentence decisions was gathered in interviews with Florida and Maryland judges at the end of each state's test year. This kind of evidence is critical to analyzing compliance, since "the mere fact that a particular sentence is within the range specified by the guidelines does not prove that the judge chose the sentence because of the guidelines, especially when the guidelines were designed to accommodate previous sentencing practices."[1]

The second data source for analyzing compliance is the data sets each guidelines project staff built from the scoresheets or worksheets completed on cases sentenced in the test period. Recorded on these scoresheets were information about the offense and offender, the score elements and total, the sentence and (if applicable), the judge's reason for going outside the guidelines.[2] In both states, some effort was made to gather the scoresheets from the test jurisdictions and to monitor their contents. These scoresheet data form the basis for answering questions about substantive compliance, with the important caveat that their quality is affected by the differences in verification and monitoring activities discussed in Chapter 5 and by the local adaptations analyzed in Chapter 6.

The third data source for analyzing compliance is actual court data collected by Abt Associates from jurisdictions participating in the guidelines test. These data, gathered from docket books, case files, and criminal history records, describe all the burglary cases sentenced during the test year. There is information on convicted charge(s), county and circuit, sentencing judge, and sentence rendered (among other items).[3] Because they represent the universe of burglaries, for that one family of offenses they allow us to assess whether scoresheets were indeed completed and filed on all eligible cases. Interview data from judges, prosecution and defense attorneys, and parole and probation officers allow fuller understanding of the patterns found.

In sum, the question of compliance goes to the very core of sentencing guidelines mechanisms--the effort to structure judicial decisionmaking. This chapter offers the details of one important explanation for the limited sentencing impact evidenced in the multijurisdictional guidelines test.

7.1 Judicial Compliance in Florida

This section presents the analysis of compliance in the four Florida test jurisdictions (the 4th, 10th, 14th, and 15th circuits). It is structured according to the questions on procedural and substantive compliance enumerated above.

7.1.1 Judicial Use of Guidelines (Florida)

The difficult question of judges' actual use of the guidelines was addressed earlier, in Chapter 5. There, we saw that there was considerable variation, both within and among the Florida test jurisdictions, in the degree to which judges consulted the guidelines in reaching sentencing decisions. For example, most of the judges in the 4th Circuit (Jacksonville) paid no attention to the guidelines in sentencing the vast majority of cases, which are disposed by plea. In contrast, guidelines were incorporated into the routine of plea negotiations in the 15th Circuit (West Palm Beach), and judges there appear to have used them extensively. In the two rural test jurisdictions, judges reported considerable commitment to and reliance on the guidelines.

Across all these sites, "using the guidelines" had meanings that ranged from simply accepting the recommended sentence, through using the scoring to help weigh the facts of the case or the recommendations of other actors, to sentencing mentally and then validating this judgment by consulting the grids. Under a voluntary guidelines system, all these types of "use" are acceptable forms of compliance.

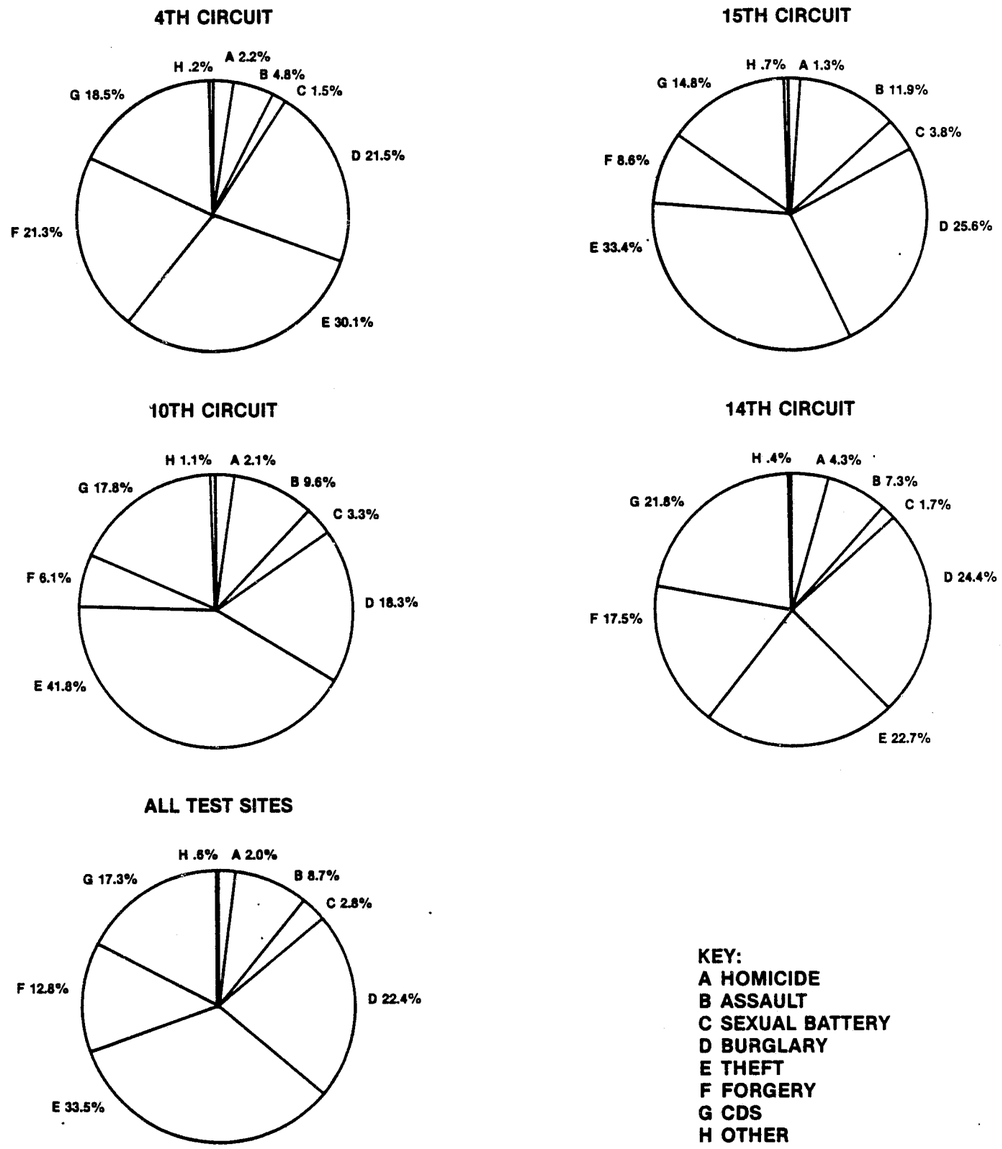
7.1.2 Completion and Filing of Scoresheets (Florida)

The second aspect of procedural compliance concerns the completion of scoresheets on cases covered by the guidelines. In Florida, the guidelines were designed to apply to the majority of felonies (and about 85 percent of the felony caseload).[4] In this section, we analyze the degree to which the primary paperwork of a guidelines system--the completion of scoresheets--was carried out in the four Florida test jurisdictions. One important limitation of this analysis should be noted at the outset. The data set of Florida scoresheets contains only those scoresheets both filled out and transmitted to the project staff in Tallahassee. It is possible that some number of scoresheets were completed in the test sites but not filed with the project. As Chapter 5 noted, the court clerks were somewhat belatedly recruited to the task of copying and batching the scoresheets, and the process was rather cumbersome, with incomplete submission a likely consequence. However, the empirical analysis presented here cannot distinguish between completions and filings. Its approach is to compare the volume of filed scoresheets to caseload figures.

Altogether, about 3,000 guidelines scoresheets were filed for the test year.[5] Figure 7-1 displays the distribution of Florida scoresheet cases according to major categories of felonies. Seven groups of crimes account for over 99 percent of the

FIGURE 7—1

Distribution of Guidelines Scoresheet Cases By Major Offense Group: Florida Test Sites



Data Set: Florida scoresheets.
Base: All cases within guidelines year and with sentence data (N = 2943).

primary offenses on the scoresheets.[6] The two largest groups are burglary (including burglary of occupied or unoccupied dwellings, structures, or conveyances, armed burglary, and possession of burglary tools) and theft (which includes grand and petty theft, stolen property offenses, grand larceny, robbery, and robbery with a deadly weapon). Thefts are the largest category in every circuit except the 14th, where burglary leads. Burglary is second in the other three, accounting for 22.4 percent of the total volume of scoresheets.

What was the rate of scoresheet filing in the four Florida test sites? The burglary data gathered by Abt Associates directly from court and Florida Department of Law Enforcement records provide a precise base against which to check the volume of scoresheets. The burglary and scoresheet data cover exactly the same time period and the same set of offenses (all burglaries were eligible for sentencing under the guidelines). In both data sets, differences in case filing procedures among jurisdictions have been adjusted by making the unit of analysis a sentencing event. Most important, the burglary data represent the universe of sentencing events in which sentences were given for burglary as the primary offense. We can use them to measure the rate of scoresheet filing on burglary cases.[7]

Table 7-1 shows this comparison. For all types of burglaries and all test circuits combined, scoresheets were filed for 57 percent of the burglaries. Jurisdictional variations are striking: the rate of scoresheet filing was only about 40 percent in the 4th and 14th Circuits, but it was 60 percent in the 10th and 100 percent in the 15th.[8] Filing was also more complete for the more serious types of burglaries--armed, with assault, or of an occupied dwelling--than for the less serious. In the two largest groups (burglary of an unoccupied dwelling or structure), the rate of filing was only 50 to 60 percent.[9]

In the four test circuits, it appears that over 500 offenders convicted of burglary charges may have been sentenced without direct reference to the sentencing guidelines. Of course, the scoresheet data set may be incomplete because of problems in verification or filing procedures in the circuits; relatively weak preparation and monitoring (see Chapter 5) did little to assure that all prepared scoresheets would be filed.[10] However, we cannot be sure scoresheets were prepared for all or even most of these cases. Without a scoresheet, it is not clear that there is any way for a judge to consult the guideline grids or consider a sentence in terms of the recommended range. Under such circumstances, the guidelines can have had no direct impact on the disposition of these cases. Of course, burglaries constituted only 22 percent of all the scoresheet offenses (see Figure 7-1). However, if the pattern of lower filing for less serious conviction offenses were to hold for other crimes, filing rates of less than 50 percent would be expected for thefts, forgeries, and drug cases as well. These account for 64 percent of all the scoresheets but--by inference--a far greater proportion of the actual sentencing caseflow.

Table 7-1

Comparison of Scoresheet and Total Case Volume for Burglaries: Florida Test Sites

Specific Burglary Offense[a]	4th Circuit (Jacksonville)	15th Circuit (West Palm Beach)	10th Circuit (Bartow)	14th Circuit (Marianna/Panama City)	All Test Sites
Burglary With Assault & Armed Burglary					
Total Volume	10	3	4	1	18
Number of Scoresheets	8	8[b]	4	0	20[b]
Percent with Scoresheets	80.0	100[b]	100	0	100[b]
Burglary of An Occupied Dwelling					
Total Volume	18	6	6	9	39
Number of Scoresheets	16	13[b]	12[b]	8	49[b]
Percent with Scoresheets	88.9	100[b]	100[b]	88.9	100[b]
Burglary of An Unoccupied Dwelling					
Total Volume	168	113	85	27	393
Number of Scoresheets	59	111	56	9	235
Percent with Scoresheets	35.1	98.2	65.9	33.3	59.8
Burglary of An Unoccupied Structure					
Total Volume	352	132	118	99	701
Number of Scoresheets	121	135[b]	54	39	349
Percent with Scoresheets	34.4	100[b]	45.8	39.4	49.8
Possession of Burglary Tools					
Total Volume	3	2	1	0	6
Number of Scoresheets	1	2	1	0	4
Percent with Scoresheets	33.3	100	100	100	66.7
Armed Trespass					
Total Volume	0	0	1	3	4
Number of Scoresheets	0	0	1	1	2
Percent with Scoresheets	100	100	100	33.3	50.0
All Burglaries					
Total Volume	551	256	215	139	1161
Number of Scoresheets	205	269	128	57	659
Percent with Scoresheets	37.2	100	59.5	41.0	56.8[c,d]

Data Sets: Florida scoresheets
Florida burglary data (test jurisdictions only)

Base: For scoresheets, all cases with burglaries as primary offense, within guideline year and with sentence data (N=695), missing=36). For burglary data, universe of cases with burglaries as primary offense (N=1164, missing=3).

Note: The definition of "case" is the same in both data sets; the unit of scoresheet filing and of burglary data collection is the sentencing event.

a. Primary offenses only.
b. See text for discussion.
c. If 36 scoresheet cases missing sentence data are included, the figure rises to 59.7 percent.
d. The Florida Final Report notes that there were "many" instances of confusion in classifying burglaries by crime category (guidelines grid), with the distinction based on whether a burgled dwelling was occupied or not. However, it lists the total number of "wrong form" cases in Categories 3 and 4 as 60. If all these were burglaries and are included in the table total, the figure rises to 61.8 percent.

The substantial discrepancies in scoresheet volumes relative to the full caseload can be understood, at least in part, if we briefly review the procedures for filling out scoresheets. The discussion in Chapter 5 noted that the major problems arose in cases disposed by plea. In the three test jurisdictions (4th, 10th, 14th) where parole and probation personnel were responsible for scoring, the scoresheets could not be prepared in time if sentencing closely followed the plea. The same constraint would prevent preparation of Presentence Investigation (PSI) reports.[11] In the 15th Circuit, however, scoresheets were prepared by prosecuting and defense attorneys, avoiding the time constraint.

It would be expected, then, that scoresheet filings would closely track the preparation of PSIs, except in the 15th Circuit. Figure 7-2 strongly supports this view. In the 4th, 10th, and 14th Circuits, virtually all scoresheets had an identified PSI investigator. Although these data cannot prove it, we can surmise that neither PSIs nor scoresheets were prepared for most cases disposed by plea negotiation. By contrast, over three-fourths of the scoresheets in the 15th Circuit did not have a PSI investigator. This was the only site in which a mechanism was developed to produce scoresheets in plea bargain situations. This was also the site with the least gap between scoresheet and total burglary volume.

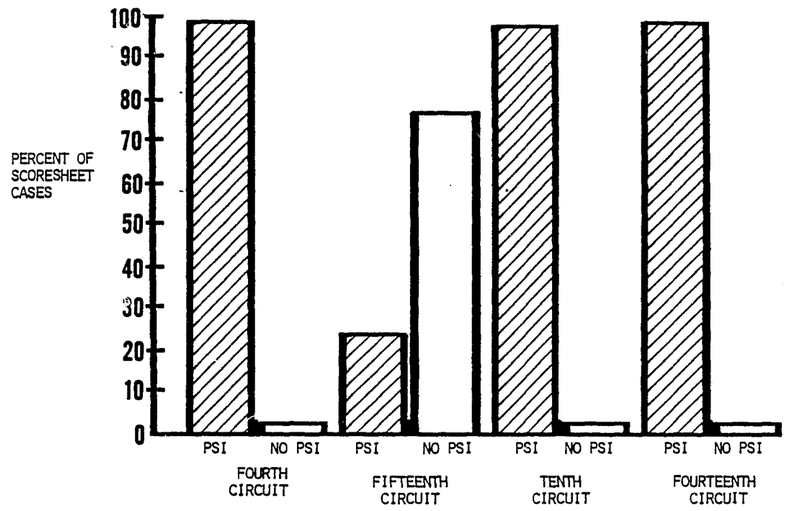
Because there appears to be a substantial shortfall of scoresheets in three circuits, the question arises whether any systematic differences exist between cases receiving scoresheets and those apparently unscored. Systematic differences would amount to a bias in the application of the guidelines during the test year. The issue of bias can be investigated by comparing the burglary and scoresheet data on a matched basis. That is, efforts were made to individually match the cases from these two independent sources, then compare the characteristics of matched with unmatched burglary cases.[12]

A comparison of the sex and race of defendants showed no significant differences between those receiving scoresheets and those not scored. By age, those 18 to 21 were significantly more likely to be scored under the guidelines, while defendants between 21 and 25 were less likely to have scoresheets than those in other age groups. Defendants on restricted status at the time of the offense had the same chance of being scored under the guidelines as those who were not restricted. Some systematic differences did exist between cases with scoresheets and cases without them: scoresheets were more likely to be filled out if there were multiple convicted counts or if the defendant had no prior record. However, there would appear to be little, if any, overall evidence of bias in scoresheet use. The significant differences appear related to PSI preparation, which is most likely for the young, first-time offender.

The individual matching of burglary with scoresheet cases also allows us to examine some elements of the scoring for accuracy. As Chapter 6 described in some detail, score bargaining between defense and prosecution was reported in at least one

FIGURE 7-2

RELATIONSHIP OF SCORESHEETS FILED TO PSI PREPARATION:
FLORIDA TEST SITES



Data Set: Florida scoresheets

Base: All cases within guidelines year and with sentence data (N = 2943, missing = 214)

Note: The scoresheet data were coded to identify the PSI investigator or to indicate that there was one (even if the individual could not be identified). Of the 2097 cases with a PSI, 130 had an unidentified preparer. Only cases where the field was blank are treated as "no PSI prepared."

Florida jurisdiction, and in others there were indications that scores were adjusted to support sentence bargains. The burglary data collected by Abt Associates allow exact checking of three of the score elements: those based on primary offense at conviction, number of counts of primary offense, and defendant's legal status at the time of the offense. Partial checks can be conducted on scoring for two elements: other counts at conviction and prior adult record. No check can be made on the scoring for prior juvenile record, type of weapon, or the defendant's role in the crime (accessory, leader, or neither).[13]

Scoring comparisons were made at three different levels. First, the burglary data were used to simulate scoring on the items for which full information was available, allowing exact checking. Second, total scores were simulated. These combined the full information (exact) items with the partal simulations on prior adult record and other offenses at conviction; zero points were allocated for the items that could not be simulated. Third, the total scores were divided into the categories of the appropriate guidelines grid, to assess any effects of score differences on recommended sentence ranges.

Table 7-2 shows the results of these comparisons. Nearly 80 percent of the cases showed no difference between simulated (burglary data) and actual (scoresheet) scoring on the exact items. The remaining cases were divided evenly between instances in which the simulated scores exceeded the actual and instances where the actual scores were larger. Thus, on the exact items the maximum extent of score bargaining would be 10 percent of the cases. Possible sources of these discrepancies include errors in scoring, data collection errors, and score bargaining. Lack of local or central verification procedures would certainly contribute to scoresheet errors.

With respect to total scores, the results were equal in only a quarter of the cases. This is to be expected, given the absence of some information items in the burglary data set. That half of the cases show actual scores greater than simulated ones is largely a result of these missing elements (weapon, juvenile record, etc.). The 27 percent of cases with simulated scores greater than actual ones indicates the maximum overall extent of score bargaining, although some discrepancies are undoubtedly the result of errors.

When the locations of the scores on the guideline grids are calculated, in only 11 percent of the cases would possible score reductions (whether due to error or bargaining) have affected the guidelines sentence. However, these cases are particularly concentrated in the 15th Circuit. There, 21 percent of the score differences made a difference (downward) in the recommended sentence range; this was significantly more than in the other three sites. Since it was in the 15th Circuit that state and defense attorneys drew up the scoresheet together, the data indicate that, in the process, advantage was taken of the opportunity to score bargain. Thus, the compliance analysis both confirms the reports of score bargaining (see Chapter 6) and estimates the impact of this practice; a full fifth of the West Palm Beach burglary

Table 7-2

Comparison of Simulated With Actual Guidelines Scoring: Florida Test Sites

	Matched Items Only[a]	Total Score[b]	Score Category for Guideline Grid
Percent of Cases With:			
Burglary data and scoresheet results equal	79.3%	25.4%	69.8%
Burglary data results greater than scoresheet results	10.4	26.7	11.2
Scoresheet results greater than burglary results	10.4	48.0	19.0
Total	100.0	100.0	100.0

Data Sets: Florida scoresheets
Florida burglary data (test sites only)

Base: All matches between scoresheets and burglary data (N=474, missing=5).

a. Matched items are those items for which full information was collected on burglary cases: primary offense at conviction, number of counts of primary offense, offender's legal status at time of the offense.

b. Total score includes matched items plus adult prior record, juvenile record, other counts at conviction, type of weapon, offender's role in the crime.

cases may have received lower sentences due to downward shifts in guidelines scores and recommended ranges. The findings on procedural compliance in Florida are thus as follows:

1) Judges' use of the guidelines in reaching sentencing decisions was variable, and there appear to be jurisdictional differences in their patterns of use. The pressure of plea negotiations was the most frequently cited explanation for sentencing without the guidelines. In all but the urban jurisdiction, guidelines use and adherence were reported to be strong.

2) Although weak verification, collection, and monitoring procedures affect the data, it appears that scoresheets were not completed on all eligible cases. To the contrary, there were large numbers of convictions on guidelines-eligible offenses for which no score was computed. The filing rate for burglaries was only 57 percent, and strong jurisdictional differences were observed. However, with the exception of defendants with no prior record or who were convicted on multiple counts (these defendants were more likely to receive scoresheets), there do not appear to be biases in scoresheet filing with respect to offense or offender characteristics. A simulation of scoring for burglaries showed substantial corroboration of scoresheet point allocations; score bargaining which could have lowered recomended sentences cannot have occurred in more than 11 percent of the cases (21 percent in the 15th Circuit).

7.1.3 Comparison of Actual and Guidelines Sentences (Florida)

How did the sentences given on scoresheet cases fall relative to the guidelines? This is the first of the analytic questions dealing with substantive compliance. Again, we must note that the scoresheets offer an incomplete picture of sentencing in at least three of the four jurisdictions during the test year, although the picture is probably not significantly biased in terms of offense and offender characteristics. Still, the scoresheet data can be checked against the appropriate guidelines grid to assess the rate at which judges sentenced within, above, or below the recommended ranges.[14]

To compare sentences to the guideline grids, scoresheet cases were grouped according to the crime category and the total points scored. Readers will recall that there were six Florida grids, corresponding to groups of crimes similar in character and seriousness (see Chapter 4 and Appendix D). Each grid was designed to show particular intervals of point scores and a corresponding range of sentences. (The point

scores reflect a combination of offense and offender characteristics.) For example, a case with a primary offense classified in Category 2 (Aggravated Battery, Aggravated Assault) and with a total score between 176 and 200 points would have a recommended sentence range of 7 to 9 years. A case of a drug offender with between 81 and 100 points scored would have a recommended sentence of 3.5 to 4.5 years under Category 6 (Possession, Sale, Delivery, Importation of a Controlled Substance). While all the guideline grids show midpoint sentence recommendations as well as ranges, for this analysis the upper and lower boundaries of the appropriate range have been used to separate sentences within the guidelines from those outside them.

Figure 7-3 summarizes the distribution of sentences against the guidelines for the four test circuits. Taken together, some 78 percent of all the cases fell within the guidelines; in the 14th Circuit this percentage was higher than in the others, at 84 percent.[15] The remaining scoresheets showed sentences above the guidelines (about 10 percent of the total) or below (about 12 percent). However, the circuits differed significantly in this split. Above-guidelines sentences were rendered in almost 16 percent of the 4th Circuit's scoresheet cases, nearly twice the proportion of any other circuit. Conversely, the 10th and 15th Circuits showed substantially greater proportions of below-guidelines sentences.[16]

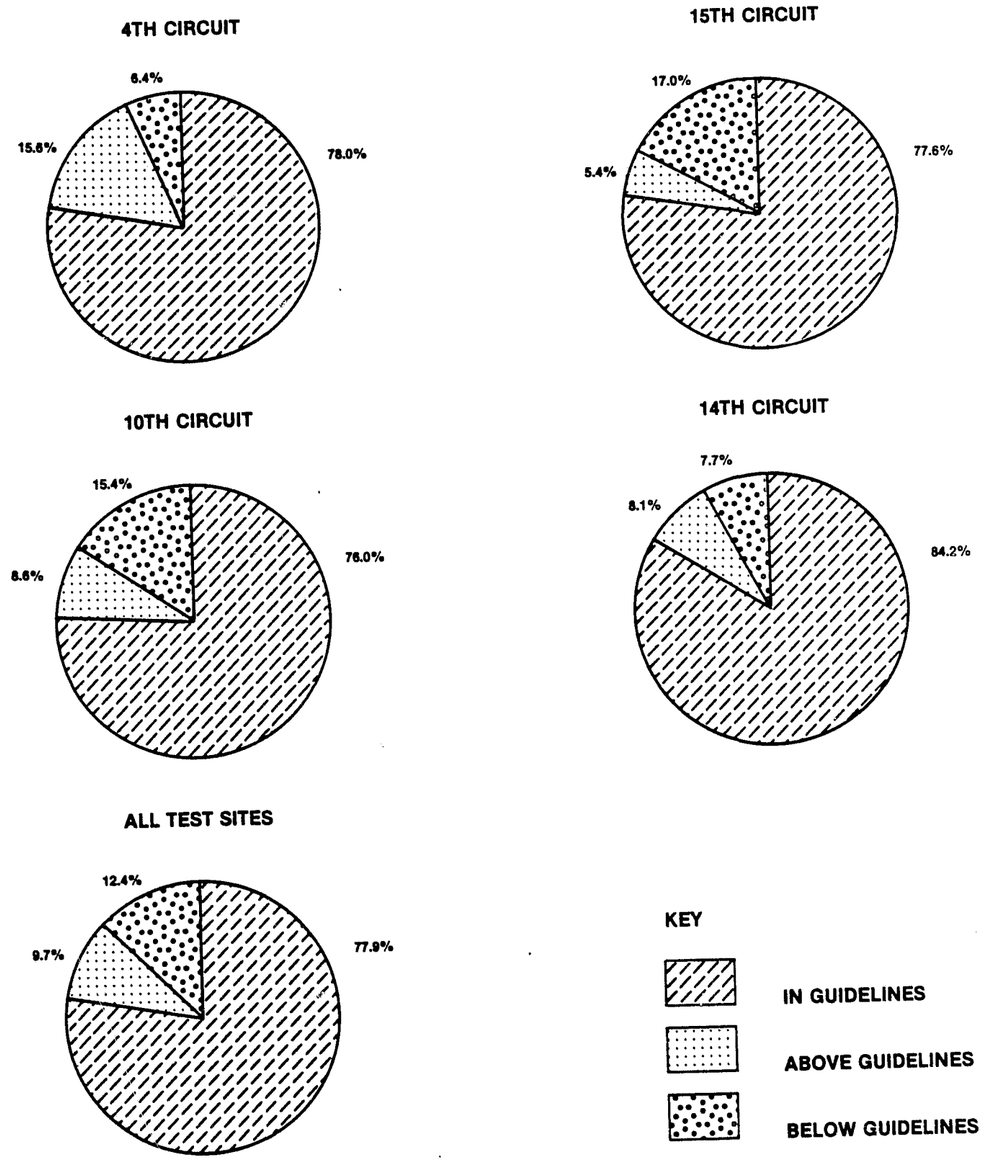
The pattern of sentencing relative to the guidelines also varies greatly by crime category. Table 7-3 shows this pattern by circuit for each of the six different (crime category) grids. Categories 1 and 3 (the former Murder, Manslaughter, Kidnapping, Lewd and Lascivious Assault; and the latter Burglary with Assault, Burglary of an Occupied Dwelling, and Robbery) show very low rates of substantive compliance in terms of sentencing within the guidelines. In contrast to these rates of roughly 50 percent or lower agreement, the other categories show 70 percent or more of the cases being sentenced within the guidelines.

What is particularly interesting about these differences is that many judges, in interviews, cited Category 3 crimes as warranting harsher sentences than those recommended by the guidelines. Several noted local public concern about rising residential burglary. Yet, for Category 3, the rate of agreement with guidelines is <u>not</u> low because a greater proportion of cases receive above-guidelines sentences (the 12 percent figure is close to those for Categories 2 and 4). Rather, there is a much larger proportion of <u>below</u>-guidelines sentences. A very similar pattern obtains for Category 1.

Table 7-3 also indicates that there were significant interjurisdictional differences for all the categories. In general (although not in Categories 2 and 5), the 4th Circuit showed a higher incidence of sentences above the recommended ranges. For Categories 2, 3, and 6, 15th Circuit sentences were more often below the guidelines than those of the other circuits. Since reducing disparity in sentencing between different jurisdictions is a central goal of multijurisdictional guidelines, it is important to note that significant disparities remain in sentencing relative to the guidelines ranges.

FIGURE 7–3

Relationship of Sentences to Recommended Guidelines Ranges: Florida Test Sites



Data Set: Florida scoresheets.
Base: All cases within guidelines year and with sentence data (N = 2943).
Note: Differences in sentence distribution among circuits relative to guidelines are statistically significant ($p \leq .01$).

Table 7-3

Distribution of Sentences Relative to the Guidelines, by Crime Category: Florida Test Sites

Crime Category[a]	Sentence Relative to Guidelines	4th Circuit (Jacksonville)	15th Circuit (West Palm Beach)	10th Circuit (Bartow)	14th Circuit (Marianna/Panama City)	All Test Sites
1	Within Range	29.7%	62.3%	51.1%	53.3%	50.6%*
	Above	32.4	6.6	15.6	20.0	16.5
	Below	37.8	31.2	33.3	26.7	32.9
	(N)	(37)	(61)	(45)	(15)	(158)
2	Within Range	71.7	72.8	79.4	70.6	74.2**
	Above	19.6	6.4	13.2	29.4	12.1
	Below	8.7	20.8	7.4	0	13.7
	(N)	(46)	(125)	(68)	(17)	(256)
3	Within Range	36.5	34.3	41.9	53.9	38.0**
	Above	33.8	2.9	6.8	0	12.6
	Below	29.7	62.8	51.4	46.2	49.4
	(N)	(74)	(102)	(74)	(13)	(263)
4	Within Range	59.5	78.5	73.0	91.8	72.7**
	Above	34.6	6.8	7.8	2.0	15.2
	Below	6.0	14.7	19.1	6.1	12.2
	(N)	(185)	(251)	(115)	(49)	(600)
5	Within Range	91.7	88.6	84.4	87.6	88.7+
	Above	6.2	6.7	8.7	7.9	7.1
	Below	2.1	4.8	6.9	4.5	4.2
	(N)	(436)	(358)	(275)	(89)	(1158)
6	Within Range	92.6	89.1	88.0	92.2	90.4**
	Above	6.8	.6	4.8	5.9	4.3
	Below	.6	10.3	7.2	2.0	5.3
	(N)	(176)	(156)	(125)	(51)	(508)

Data Set: Florida scoresheets

Base: All cases within guidelines year and with sentence data (N=2943, missing=204).

Note: Statistical significance of differences among circuits in sentence distribution (relative to quidelines) tested by X^2:

** $p \leq .01$
* $p \leq .05$
+ $p \leq .10$

a. The six crime categories from the Florida quidelines grids. See Appendix D for crimes included in each.

There are further indications of disparity in the rates of incarceration for scoresheet cases. The incarceration rate measures the proportion of convicted offenders given jail or prison sentences rather than full suspensions or probation. Table 7-4 shows these rates by circuit and crime category. In every category, offenders in the 4th Circuit were most likely to receive incarcerative sentences. In four of the six categories, the 15th Circuit scoresheets showed the lowest incarceration rates.[17]

Of course, some of this apparent interjurisdictional disparity does result from differences in the seriousness of the cases, especially in Category 1. However, as Table 7-5 indicates, the great preponderance of all cases falls within the lowest cell of the appropriate grid. The table shows the proportion of scoresheet cases that fall into each grid cell, on the basis of total score and primary offense. The shaded boxes indicate the modal cell for each grid--that is, the cell to which the largest number of cases belongs. For all the scoresheets combined, nearly three-quarters fell into the lowest cell (lowest point scores and sentences) of the appropriate grid. For Categories 5 and 6, about nine-tenths of the cases were scored in the lowest cell. Only for Category 3 did more cases fall into a higher cell, and the lowest cell ranked second.

Given a scoresheet distribution in which the great majority of cases falls in the least serious guideline range, there is little reason to think that major point differences may account for sentence disparity relative to the guidelines. However, the wide sentence range for the lowest cell (0 to 18 months in all categories, except 0 to 36 months in Category 3) may well be a significant factor. Chapter 4 described the early revisions to these grids, in which the two lowest cells (OUT and 1-18 or 1-36) were combined. These revisions certainly left even more room for sentence disparity within the guidelines ranges.

As Section 7.1.2 showed, it appears that scoresheets were filed on only a subset of all the cases eligible for guidelines sentencing. The rate of filing (as measured for burglaries) varied from a low of 37 percent in the 4th Circuit to 100 percent in the 15th. Using the same matched comparison of scoresheets against the universe of burglary cases, we can assess whether the patterns of sentencing relative to the guidelines are likely to have been the same for the full caseload during the test year as they were for cases with scoresheets filed.

In Figure 7-4, the burglary data set is split according to whether there was a matching scoresheet or not. Using simulated scores, the proportions of sentences within, above, and below the guidelines are calculated for each of these groups separately and for all burglary cases combined. As described in Section 7.1.2, simulated scoring of cases in the burglary data set is only a low approximation of true scoring, due to uncollected data elements such as juvenile record and weapon. Therefore, the proportions of in-guidelines and extra-guidelines sentences for Category 4 in Table 7-3 are likely to be more accurate than those in Figure 7-4. (Category 4 includes most burglary cases and no other offenses.) Here, the important

Table 7-4

Incarceration Rates, by Crime Category and Circuit: Florida Test Sites

Crime Category[a]	4th Circuit (Jacksonville)	15th Circuit (West Palm Beach)	10th Circuit (Bartow)	14th Circuit (Marianna/Panama City)	All Test Sites
1. Murder, Manslaughter, etc.	91.9% (37)	63.9% (61)	82.2% (45)	80.0% (15)	77.2% (158)
2. Aggravated Battery, Aggravated Assault	63.0 (46)	36.8 (125)	36.8 (68)	47.1 (17)	42.2 (256)
3. Burglary of an Occupied Dwelling, Robbery, etc.	94.6 (74)	65.7 (102)	93.2 (74)	69.2 (13)	81.7 (263)
4. Armed Burglary, Burglary of an Unoccupied Dwelling, etc.	75.1 (185)	53.0 (251)	52.2 (115)	57.1 (49)	60.0 (600)
5. Grand Larceny, Theft, Forgery, etc.	39.9 (436)	29.6 (358)	35.3 (275)	34.8 (89)	35.2 (1158)
6. Possession, Sale, Delivery, Importation of a Controlled Substance	34.1 (176)	21.2 (156)	15.2 (125)	25.5 (51)	24.6 (508)
All Categories	53.0 (954)	40.3 (1053)	43.7 (702)	43.2 (234)	45.5 (2943)

Data Set: Florida scoresheets

Base: All cases within guidelines year and with sentence data. (N=2943, missing=204).

Note: Percents represent proportion of convicted offenders in each group given incarcerative sentences. Total number of cases is shown in parentheses.

a. The six crime categories for the Florida guidelines grids.

Table 7-5

Distribution of Guidelines Scoresheet Cases by Crime Category and Location in Grid: Florida Test Sites

	GUIDELINES CRIME CATEGORIES						
GRID CELL[a]	1	2	3	4	5	6	All Categories
1	24.7%	73.8%	15.2%	65.3%	89.1%	92.1%	73.4%
2	3.8	6.6	9.5	17.2	5.4	2.4	7.7
3	9.5	6.6	6.8	5.7	3.1	1.6	4.4
4	7.0	4.7	8.0	4.3	1.4	2.4	3.3
5	5.7	.8	10.7	2.5	.4	.8	2.1
6	12.0	1.2	9.1	1.5	.4	.6	2.1
7	2.5	3.9	19.8	2.0	0	.2	2.7
8	10.8	1.2	9.1	.7	.1	0	1.7
9	2.5	.4	4.9	.3	.1	0	.7
10	5.7	.8	1.1	.3	0	0	.5
11	5.7	0	5.7	0	.1	0	.9
12	10.1	0		0	.1	0	.6
13		0		.2	0	0	.03
14		0				0	0
All Cells	100.	100.	100.	100.	100.	100.	100.
(N)	(158)	(256)	(263)	(600)	(1158)	(508)	(2943)

Data Set: Florida scoresheets

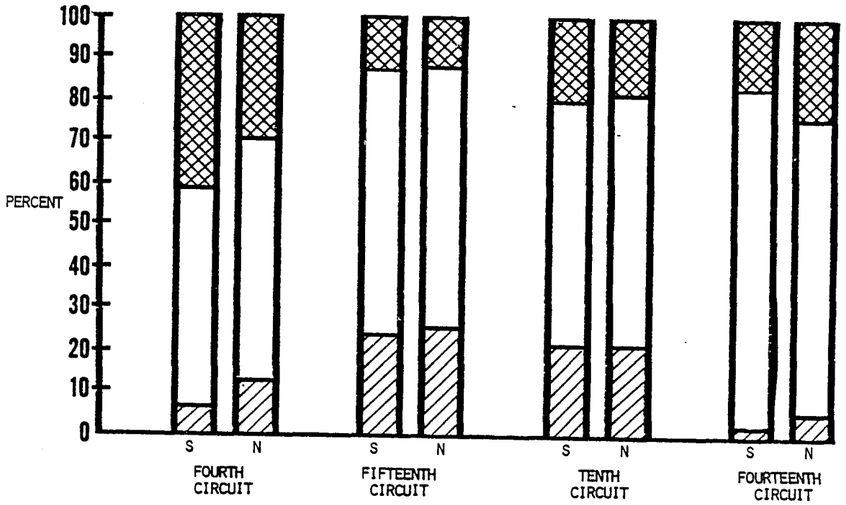
Base: All cases within guidelines year and with sentence data.

Note: Shaded boxes indicate modal grid cell for each crime category.

a. Cells are numbered from top to bottom of each grid, in ascending score order. Cells correspond to different point ranges and sentences across the crime categories. Actual grids are shown in Appendix D.

FIGURE 7-4

DISTRIBUTION OF SENTENCES RELATIVE TO GUIDELINES FOR BURGLARY CASES WITH AND WITHOUT SCORESHEETS: FLORIDA TEST SITES



Data Set: Florida burglary data

Base: All cases with prior record and sentence information (N = 1147, missing = 17)

Notes: Sentences are based on simulated scores only. See Table 8-3 for distribution of sentences based on actual scores.

Differences in sentence distribution between burglary cases with and without scoresheets were statistically significant in the 4th Circuit ($p \leq .01$).

S = With scoresheets
N = Without scoresheets

KEY:

ABOVE GUIDELINES

WITHIN GUIDELINES

BELOW GUIDELINES

comparisons are within the figure, where cases are treated uniformly as to simulating scores.

For the four test sites together, 60 percent of the cases were sentenced within the recommended range based on the simulated score, while 25 percent of the sentences were higher and 15 percent lower. As with the scoresheet data, the circuits varied in these outcomes. The 4th Circuit had the lowest agreement rate and largest proportion of above-guidelines sentences, while the 14th had the highest rate of agreement. In three of the circuits, there was no significant difference between burglary cases with scoresheets and those without in terms of these proportions. However, in the 4th Circuit, the distribution of sentences (based on simulated scores) for cases without scoresheets showed a significantly lower proportion above the guidelines and greater proportions of within- and below-guidelines sentences. In fact, the "true" overall percentage of below-guidelines burglary sentences in the 4th Circuit is nearly double that shown on scoresheet cases (10.2 percent against 5.7 percent). Since we know that cases disposed by plea negotiation were rarely, if ever, scored in the 4th Circuit, we can infer that such negotiations often resulted in sentence bargains below the recommended ranges.

7.1.4 Presence of Reasons for Extra-Guidelines Sentences (Florida)

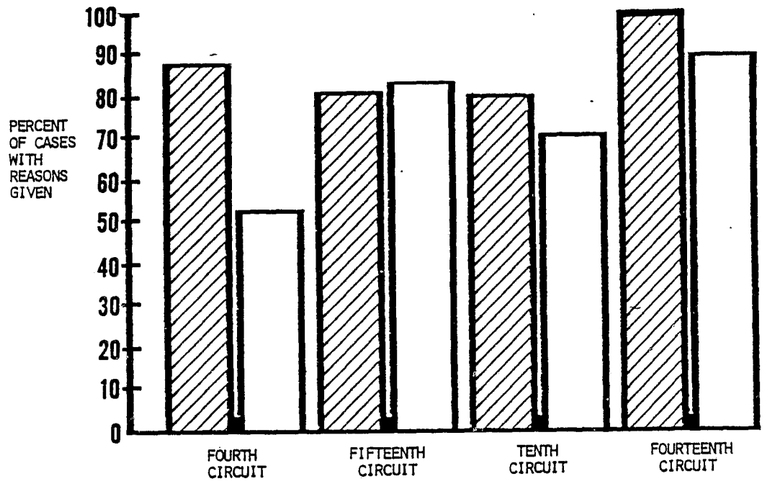
We have seen that a considerable proportion of cases scored under the Florida guidelines (some 22 percent overall) received extra-guidelines sentences.[18] This is not necessarily an unexpected rate of substantive non-compliance. The ranges in the six guideline grids were designed to cover a substantial majority of actual sentences, but it was understood that the remaining cases would receive extra-guidelines sentences because of special circumstances. In fact, guidelines systems have a specific mechanism to accommodate cases with aggravating or mitigating factors. When sentencing such a case outside the recommended range, the judge is asked to state his or her reason for the sentence.

Now we review the evidence on the degree to which Florida judges in the test jurisdictions complied with the requirement to give reasons. (An examination of the nature of their reasons follows, in Section 7.1.5.) The analysis uses the scoresheets collected by the Tallahassee project staff, with the attendant caveat concerning the completeness and quality of this data set. It focuses on the several hundred cases in which extra-guidelines sentences were rendered and checks whether reasons were, in fact, provided for them.

Figure 7-5 summarizes the rate of compliance with respect to judicial reasons. For the four test sites combined, just under 80 percent of the extra-guidelines sentences were accompanied by reasons. This rate varied by circuit, with greatest compliance in the 14th and least in the 10th. The variations were statistically significant. In three of the four circuits, cases receiving above-guidelines sentences were more likely to have the required reasons.[19] The difference is most

FIGURE 7-5

PRESENCE OF JUDICIAL REASONS FOR EXTRA-GUIDELINES SENTENCES: FLORIDA TEST SITES



Data Set: Florida scoresheets

Base: All cases within guidelines year and with extra-guidelines sentences (N = 651)

Note: Differences among circuits in sentence distribution were statistically significant ($p \leq .01$).

KEY:

CASES SENTENCED ABOVE GUIDELINES

CASES SENTENCED BELOW GUIDELINES

striking for the 4th Circuit. Whereas reasons were given for 88 percent of the above-guidelines sentences, they were given for only 53 percent of sentences below the recommended range. We also saw that scoresheets probably represent an underestimate of the incidence of below-guidelines sentences in that circuit. The pattern of missing reasons, particularly for below-guidelines sentences, is more likely to derive from differential handling of plea-negotiated cases resulting in sentence bargains below the recommended range. As we shall see, the evidence on the nature of the reasons given also supports this inference.

7.1.5 Nature of Reasons for Extra-Guidelines Sentences (Florida)

As we have just seen, in over 500 scoresheet cases, Florida judges in the test jurisdictions articulated reasons for extra-guidelines sentences. This section presents an analysis of those reasons, in terms of their substance and whether they were offered for above-guidelines or below-guidelines sentences. Some 87 separate response categories, containing 519 responses, are found in the scoresheet data set. They represent the raw data on which this exercise in classification and interpretation is conducted. The caveat offered in the Florida Final Report on this subject is well-taken and worth quoting here: "...due to the problems associated with classifying and interpreting open-ended, subjective responses, care should be taken in any attempt to make meaningful interpretations or generalizations."[20]

In classifying the reasons, the primary distinction made is between reasons that essentially challenge the guidelines and those citing aggravating or mitigating circumstances. Challenges include reasons that repeat factors already counted in the scoring, as well as those citing plea negotiations, which were clearly meant to be covered by the guidelines. Such challenges raise an important question. On the one hand, they are clear instances of a judge declining to comply with the compromise or consensus represented by the final guidelines. In a guidelines system with review or appeal of sentences, these would be potentially reversible decisions. Under the Minnesota guidelines, for example, plea negotiations have been ruled by the Minnesota Supreme Court to be an inadequate reason for an extra-guidelines sentence.[21] In fact, the Minnesota guidelines include guidance as to factors which may be cited as reasons for sentences outside the recommended ranges. Their allowed factors are highly specific, meant to apply to small numbers of cases.

The Florida sentencing guidelines did not include any specific instructions or prohibitions concerning the reasons judges might give for extra-guidelines sentences. Indeed, as in the other guidelines projects growing out of the work of the Albany Criminal Justice Research Center, the reasons were seen as an important feedback mechanism contributing to periodic review and revision of the guideline ranges. Viewed this way, reasons that challenge the guidelines by citing offense or offender characteristics counted in the scoring are not illegitimate. Rather, they represent objections to the weighting of these characteristics, and a pattern of such objections

can contribute to revising the score weights. Systems that prohibit this kind of reason forego a significant type of feedback. Therefore, while we classify some reasons given by the judges as challenging the guidelines, in this context the challenges (or at least some of them) are both legitimate and potentially informative.

Table 7-6 shows the reasons given by judges for scoresheet cases sentenced below the recommended range. Twenty-two percent of these reasons challenged the guidelines, while 78 percent cited mitigating factors. The preponderance of challenges cited plea negotiations--plea bargains, compromises between state and defense, point or category negotiation. Absence of any prior record accounted for nearly a fourth of the challenges, with offender's role ("only an accessory," "really self-defense") accounting for 10 percent. Finally, in some instances, judges cited "too many points" or "close to lower category," simply rejecting the guidelines sentence.

Mitigating circumstances cited by the judges were rather evenly spread among several groups of factors. Case-related considerations included cooperation of the defendant with the authorities, consistency with a co-defendant's sentence, and evidentiary problems in the state's case. The Youthful Offender Act was invoked in a fifth of the cases with mitigation. Recommendations of other parties, including the police, the state, and the victim, were also important. An offender's circumstances--need for alcohol or drug treatment, presence of a handicap, hardship--and his or her demeanor in court were frequently cited. Mitigating offense characteristics included apparent overcharging, but in 24 of the 33 cases, no specific information accompanied "circumstance of the offense."[22]

The judicial reasons given for above-guidelines sentences are shown in Table 7-7. Again, they are divided into two sets: those challenging the guidelines (55 percent of the total) and those citing aggravating circumstances (45 percent). Note that the proportion of challenging reasons is over twice that observed for below-guidelines sentences.

The largest group of challenging reasons concerns the offender's prior record; while 37 of the 59 instances simply indicated "prior record," others cited extensive recent record or criminal lifestyle. With regard to offense characteristics, the modal reason concerned scoring considered low when several separate cases were involved in a single sentencing.[23] In 22 cases, the judge felt that even more points should be required if the defendant was on parole or probation or out on bond at the time of the offense. Under "plea bargains," instances were cited in which the state dropped charges or did not file them in exchange for a plea.

The reasons concerned with aggravating circumstances show considerable variety. Offense characteristics include the severity of the primary offense (threats to kill the victim, a shot fired) and "real offense" issues--more accurate original charges, jury verdicts for lesser charges than those upon which the judge would have convicted, other crimes not prosecuted by the state. Such reasons bear a relationship

Table 7-6

Reasons Given for Below-Guidelines Sentences: Florida Test Sites

Reasons[a]	Number of Times Cited	Percent of Category	Percent of All Reasons
Challenges to the Guidelines	61	100.0	22.3
1. Plea Negotiations	33	54.1	12.1
2. Prior Record	14	23.0	5.1
3. Role of the Offender	6	9.8	2.2
4. Unexplained Rejection of Guideline Sentence	8	13.1	2.9
Mitigating Circumstances	212	100.0	77.7
1. Recommendations of Other Parties	39	18.4	14.3
2. Youthful Offender Act	45	21.2	16.5
3. Offender Characteristics	38	17.9	13.9
4. Offense Characteristics	33	15.6	12.1
5. Case-Related Considerations	48	22.6	17.6
6. Other Factors	9	4.2	3.3

Data Set: Florida scoresheets

Base: All cases within guidelines year, with below-guidelines sentence and with reason given (N=273).

a. Reason categories are explained in text.

Table 7-7

Reasons Given for Above-Guidelines Sentences: Florida Test Sites

Reasons[a]	Number of Times Cited	Percent of Category	Percent of All Reasons
Challenges to the Guidelines	**135**	**100.0**	**54.9**
1. Offender Prior Record	59	43.7	24.0
2. Offense Characteristics	24	17.7	9.8
3. Offender Legal Status	22	16.3	8.9
4. Plea Negotiations	21	15.6	8.5
5. Unexplained Rejection of Guideline Sentence	9	6.7	3.7
Aggravating Circumstances	**111**	**100.0**	**45.1**
1. Offense Characteristics	30	27.0	12.2
2. Recommendations of Other Parties	29	26.1	11.8
3. Offender Prior Record	20	18.0	8.1
4. Other Offender Characteristics	11	9.9	4.5
5. Other Factors	21	18.9	8.5

Data Set: Florida scoresheets

Base: All cases within guidelines year, with above-guidelines sentence and with reason given (N=246).

a. Reason categories are explained in text.

to those rejecting plea bargains with reduced or dropped charges. Among the recommendations of other parties leading to above-guidelines sentences were legislatively-mandated sentences and the Youthful Offender Act.

Reasons related to the offender's prior record but not challenging the guidelines scoring (i.e., not reflecting elements already in the scoring) concerned arrest history, repeats of an offense, rearrests before plea or sentence, crimes committed while an escapee, and parole or probation violations after only a short length of time. Other offender characteristics, such as demeanor, age, and lack of family supervision, also appear as aggravating circumstances. Finally, among miscellaneous aggravating factors cited were a recent increase in burglaries in the community and consistency with sentences given co-defendants.

Considering all these reasons for extra-guidelines sentences, the greatest limitation of the reason-writing mechanism in actual use appears to be the judges' lack of specificity. "Circumstances of the offense" tells nothing that would help in revising guidelines. "Plea negotiations" without a more specific rationale seems to imply abrogation of judicial decision-making in the face of prosecution-defense compromises. On the other hand, some reasons raise significant issues. Enhanced sentences for cases where negotiation led to reduced or dropped charges must lead to a question of due process, since punishment is based on charges other than those on which the offender was convicted. When jurisdictional differences, like the presence of a crime wave, affect sentencing, they raise anew the question of what sources of disparity are legitimate in a guidelines system.

7.2 Judicial Compliance in Maryland

We have seen that the record of judicial compliance in Florida under the multijurisdictional sentencing guidelines test was rather varied in terms of guidelines utilization, and sufficiently incomplete to support doubts about the possible impact of the guidelines on sentencing. On the other hand, nearly 80 percent of scoresheet sentences fell within the recommended ranges, and about 80 percent of extra-guidelines sentences were accompanied by judicial reasons. In turning to the question of compliance in Maryland, the analysis will again seek to assess the evidence with regard to these main issues:

- judges' use of guidelines in sentencing decisions;
- completion of scoresheets on eligible cases;
- relationship of actual sentences to recommended ranges;
- provision of reasons for extra-guidelines sentences; and

- nature of reasons for extra-guidelines sentences.

Again, a combination of data sources will be brought to bear on these points, as we seek to reach a judgment about procedural and substantive compliance in Maryland, and similarities to or contrasts with the Florida experience.

7.2.1 Judicial Use of Guidelines (Maryland)

The degree to which judges in the four Maryland test jurisdictions used the guidelines and the ways in which they did so were discussed in Chapter 5. Judges in all of the sites reported substantial awareness of the guidelines, but patterns of use varied. Baltimore City judges often used the grids to check their initial independent decisions; this was also reported by some Prince George's County judges. Other Maryland judges used the scoring and sentence ranges as one factor among several in reaching a sentence choice.

A pattern of substantial agreement between sentences and guidelines ranges was noted by many judges--as well as by prosecutors, defenders, and probation officers--during the test year. This pattern was thought to have at least two sources: some feeling of obligation to stay within the guidelines (especially in Prince George's County), and fairly general consonance between judges' own views and the recommended ranges. As we shall see later in this section, actual rates of agreement are perhaps not as high as these perceptions would suggest.

7.2.2 Completion and Filing of Scoresheets (Maryland)

The responses of judges and others concerning guidelines use seem to indicate widespread preparation of scoresheets and attention to them in sentencing. Were scoresheets completed for all eligible cases? Once again, we can use the scoresheet data assembled by the guidelines project, in conjunction with external data, to check the degree to which paperwork procedure was followed in filing scoresheets.[24]

Table 7-8 provides an overview of the case mix in the four participating jurisdictions during the guidelines test year. It shows the distribution of cases for which scoresheets were prepared among the three major offense types corresponding to the guidelines grids.[25] Within these types, cases are divided according to a seriousness ranking established by the Advisory Board of the Maryland guidelines project.[26] The categories cut across the crime types, in the sense that (e.g.) category 3 carries the same level of seriousness whether the crime is against persons or property or is a drug offense.

There are notable differences among the jurisdictions in the mix of cases sentenced with scoresheets during the test year. While crimes against persons made up 40 percent or more of the case volume in Baltimore City and Prince George's County, they were less than 20 percent of Montgomery's cases and only 10 percent of

Table 7-8

Distribution of Guidelines Scoresheets by Major Offense Group: Maryland Test Sites

Type of Crime and Seriousness Category[a]		Baltimore City	Prince George's County	Montgomery County	Harford County	All Test Sites
Crimes Against Persons*						
Category	1	3.4%	4.1%	1.9%	1.3%	3.2%
	2	18.3	19.0	7.1	3.8	16.1
	3	17.7	10.0	3.8	2.6	13.0
	4	8.3	6.1	5.2	1.3	7.0
	5	.3	.2	.3	0	.3
	6	1.2	.9	.6	1.3	1.0
	Total	49.2	40.3	19.0	10.3	40.6
Number of Cases		(901)	(225)	(69)	(8)	(901)
Crimes Against Property**						
Category	3	4.4%	1.6%	2.5%	7.7%	3.5%
	4	25.7	31.4	36.8	52.6	29.9
	6	5.4	8.8	22.0	9.0	9.1
	Total	35.5	41.8	61.3	69.2	42.5
Number of Cases		(432)	(233)	(223)	(54)	(942)
Drug Crimes**						
Category	3	8.3%	11.7%	12.9%	14.1%	10.1%
	4	4.1	2.7	3.9	3.9	3.7
	6	2.9	3.6	3.0	2.6	3.1
	Total	15.3	17.9	19.8	20.5	16.9
Number of Cases		(186)	(100)	(72)	(16)	(374)
All Scoresheets**		(1217)	(558)	(364)	(78)	(2217)

Data Set: Maryland scoresheets

Base: All scoresheets with disposition dates between June 1, 1981 and February 28, 1982 and sentenced before April 1, 1982 (N=2217, missing=2).

Note: Statistical significance of differences among jurisdictions in crime distribution tested by chi-square:

** $p \leq .01$
* $p \leq .05$

[a] Seriousness categories are those established by Maryland Guidelines Project. They are numbered from most to least serious, and classified across offense types.

Harford's. The latter two counties had a slightly greater proportion of drug crime cases, but most of the difference is made up in the property crimes category. There, seriousness category 4 crimes were the largest group for every site (in fact the largest group overall), but the pattern in Montgomery County--with over a third of the property crimes in the lowest seriousness category--was significantly different from all the other sites.

It is within the group of crimes against property that we can make a more precise check on how scoresheet filings compare to full caseflow through the courts.[27] Using the Maryland burglary data to represent the universe of burglary cases sentenced during the guidelines year, we can make a comparison matched in time period, range of offense (since all burglaries were guidelines offenses), and unit of analysis. The burglary data represent sentencing events (a single judge and offender on a single day); so, by and large, do the scoresheet data.[28] Figure 7-6 arrays the scoresheet counts against the total burglary case volume.[29]

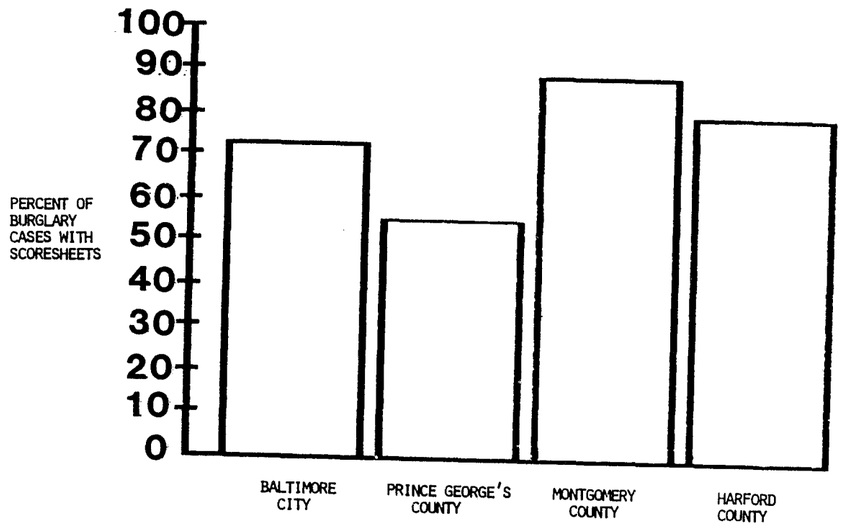
Overall, there appear to be scoresheets for 70 percent of all the burglary cases, with some variation among jurisdictions.[30] Montgomery County's compliance rate on filing was the highest, while Prince George's County showed a notable shortfall of scoresheets (nearly half). Baltimore City filing was also low, except for the middle-level burglaries which raised the overall rate for that site closer to the rate for Harford County.

Chapter 5 examined the monitoring procedures used by the Maryland project staff to track scoresheet filing against court dockets. The project estimated a filing rate of 84.5 percent overall, somewhat higher than our 70 percent figure. However, the data set of scoresheets used in this analysis also reflects the project's editing for "fatal errors." Scoresheets with certain kinds of errors were never entered into the data base. While it was estimated that this affected 10 percent of the cases received in Annapolis, no separate tally of the errors was maintained. It seems likely that some or all of the discrepancy between the 70 percent figure based on the burglary data and the project's 85 percent figure reflects a higher rate of loss from errors.

Although not as dramatic as the gap between scoresheets and case volume in Florida, a gap does appear to exist in Maryland as well (resulting from both non-filing and errors). There was considerable evidence to point to plea-negotiated cases in Florida as the ones most likely omitted from guidelines scoring. Was this also true in Maryland? The scoresheet data indicate both type of disposition and whether a PSI was available, so that this possibility can be checked directly. Examining these items (shown in Table 7-9) reveals that a great proportion of the scoresheets are for plea-negotiated cases, and that these often had PSIs prepared. Certainly, this was less likely in Baltimore City and Prince George's County than elsewhere, but even in those jurisdictions there were scoresheets filed for plea cases without PSIs. As we saw, this was very rarely done in three of the four Florida test circuits. Thus, the origin of the Maryland discrepancy in scoresheet filing is not as clear.[31]

FIGURE 7-6

COMPARISON OF SCORESHEET FILING WITH TOTAL BURGLARY CASE VOLUME: MARYLAND TEST SITES



Data Sets: Maryland Scoresheets
Maryland Burglary Data (test jurisdictions only)

Base: All cases disposed (conviction or plea) on or after September 1, 1981 and sentenced before April 1, 1982.

Note: For detail by type of burglary, see Appendix J, Table J-3.

Table 7-9

Incidence of Plea Negotiations and PSI Preparation Among Scoresheet Cases: Maryland Test Sites

	Baltimore City	Prince George's County	Montgomery County	Harford County	All Test Sites
Crimes Against Persons					
Percent disposed by plea	65.4	49.8	60.9	62.5	61.2
Percent with PSI available	66.4	80.9	97.1	100.0	72.7
Percent of pleas with PSI available	50.8	66.1	95.2	100.0	57.7
	(599)	(225)	(69)	(8)	(901)
Crimes Against Property					
Percent disposed by plea	79.2	68.2	77.6	63.0	75.2
Percent with PSI available	54.2	73.4	92.8	94.4	70.4
Percent of pleas with PSI available	43.3	62.3	91.3	97.1	61.9
	(432)	(233)	(223)	(54)	(942)
Drug Crimes					
Percent disposed by plea	74.7	71.0	81.9	56.3	74.3
Percent with PSI available	50.5	71.0	93.1	93.8	66.0
Percent of pleas with PSI available	36.7	59.2	91.5	100.0	56.1
	(186)	(100)	(72)	(16)	(374)

Data Set: Maryland scoresheets

Base: All scoresheets with disposition dates between June 1, 1981 and February 28, 1982 and sentenced before April 1, 1982 (N=2219, missing=2).

Despite the greater simplicity of Maryland's guidelines scoring relative to Florida's,[32] the burglary data collected by Abt Associates were not suitable for a scoring simulation of the type described above for Florida. Therefore, we cannot assess the quality or accuracy of scoresheet completion, except internally to the scoresheets. We note that, whether from scoring or from data entry errors, there were 128 cases of the 2,217 total (5.8 percent) with discrepancies between the point elements and the summary offense and offender scores. As we shall see, there were also errors in choice of guideline range. For ongoing guidelines systems, accuracy of scoresheet usage is likely to be of great importance to the issues of review, appeal, or reversible error. The Maryland project staff did edit the scoresheets for certain errors and missing information (as discussed in Chapter 5), and such efforts will be of even greater significance where guidelines become a regular feature of court operation.

7.2.3 Comparison of Actual and Guideline Sentences (Maryland)

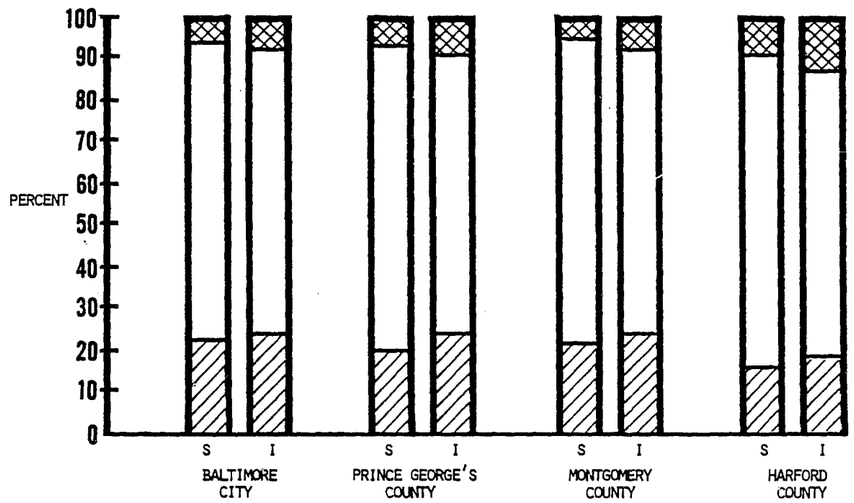
We have seen that the evidence on procedural compliance in Maryland indicates a fair degree of judicial use of guidelines but some shortfall on completion and filing of scoresheets. (This shortfall, while less dramatic than that for two of the Florida sites, was particularly marked for Prince George's County.) We turn now to the issues of substantive compliance, and first to the question of how the sentences recorded for scoresheet cases fall relative to the recommended ranges on the guidelines grids. The Maryland guidelines involved three grids (or sentencing matrices) --one each for crimes against persons, for property offenses, and for drug offenses. Scoring of crimes against persons involved computing an offender score (with adult and juvenile criminal record, prior convictions on same offense, adult parole/probation violations, and status at time of offense as the elements), and an offense score (related to seriousness, victim vulnerability and injury, and weapons use). For property and drug offenses, only the offender score was computed; depending on the level of offense seriousness, the same score carried different recommended ranges.

Taking all the scoresheets together, Figure 7-7 shows how sentences were distributed against the guidelines ranges. The scoresheet data set contained a variable added by the project staff to indicate whether this sentence was within, above, or below the range; this classification is shown in the figure by the label "according to scoresheets." Parallel measures are also presented in Figure 7-7, based on an independent check of the correct ranges. The check used the offender score and the offense score (for crimes against persons) or offense group (for drugs and property crimes) to link the scoresheet cases to the sentencing grids.[33]

While the independent check shows slightly lower proportions of within-guidelines sentences (68 percent overall compared to 72 percent), the patterns do not differ. Almost three times as many cases were sentenced below the guidelines as above, except in Harford County where the balance was more even but the number of

FIGURE 7-7

RELATIONSHIP OF SENTENCES TO GUIDELINES RANGES:
MARYLAND TEST SITES



KEY:

ABOVE GUIDELINES

WITHIN GUIDELINES

BELOW GUIDELINES

Data Set: Maryland scoresheets

Base: All scoresheets with disposition dates between June 1, 1981 and Feb. 28, 1982 and sentenced before April 1, 1982 (N = 2217, missing = 2)

Note: S = According to scoresheets

I = By independent check

See text for explanation.

cases very small. There was virtually no variation among the four sites in the rate of extra-guidelines sentences. Similarly, jurisdictions did not vary in the relationship of sentences to ranges by the three major crime types.[34]

We have seen that across the four Maryland test jurisdictions, rates of agreement between actual and recommended sentences were quite uniform and in the 67 to 68 percent range (according to the independent check). Note that these rates were substantially lower than those in Florida (76 to 84 percent), even though the Maryland ranges in each cell are typically quite wide and there is a great deal of overlap from cell to cell (see Chapter 4). The finding is also surprising in light of perceived high levels of agreement, reported in interviews, between judges' sentences and the guidelines. Yet it appears that there was less substantive compliance in Maryland than in Florida with respect to sentencing within the guidelines.[35] Of course, the guidelines system by no means required that judges stay within the recommended limits for sentencing, so that full conclusions on substantive compliance require examining use of the reasons mechanism as well.

7.2.4 Presence of Reasons for Extra-Guidelines Sentences (Maryland)

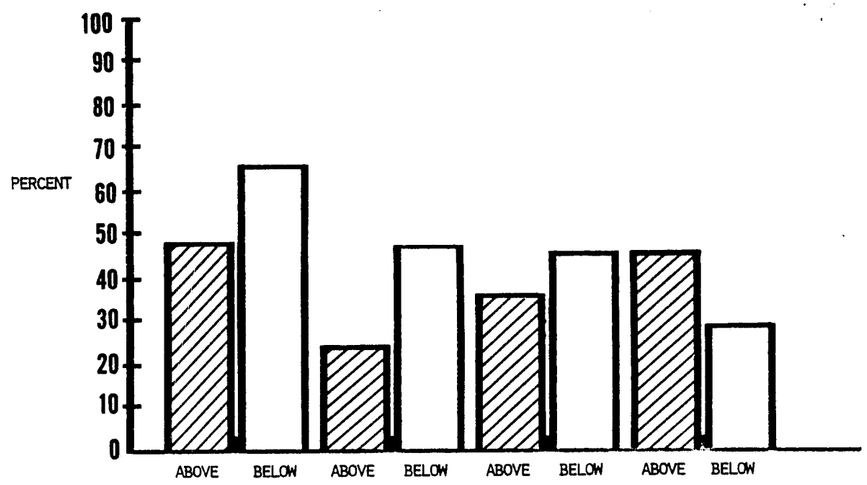
We have seen that the judges in the four Maryland jurisdictions rendered more than 700 extra-guidelines sentences during the test year. Under the guidelines model promoted in the multijurisdictional field test, such sentences are <u>not</u> non-compliant if they are accompanied by judicially articulated reasons. In the Maryland system, as in Florida's, such reasons were expected. To what degree were they actually provided?

Figure 7-8 shows, for all the scoresheets, the proportion of above- and below-guidelines sentences with reasons present. Taken all together, just over 50 percent of cases with reasons required actually had them. This proportion ranged from 36 percent for Harford County to 60 percent for Baltimore City. Reasons were present less frequently for above-guidelines than for below-guidelines sentences, except in Harford County. The difference was particularly strong in Prince George's County, where less than a quarter of the above-guidelines sentences were accompanied by reasons, compared to nearly half the below-guidelines sentences. This pattern is in striking contrast to the Florida findings, which show a much greater tendency to provide reasons for the harsher sentences than for the lower ones (see Figure 7-5).

Although there is some variation by type of offense,[36] the overall rate of compliance with the requirement to give reasons is very low. Even taking into account the re-classification of cases based on an independent check against the grids (see Figure 7-7), there is a major shortfall that cannot be explained except by recognizing that Maryland judges often ignored the reasons mechanism. By contrast, Florida judges provided reasons in nearly 80 percent of the extra-guidelines scoresheet cases; as previously noted, they also sentenced within the guidelines ranges more frequently.

FIGURE 7-8

PRESENCE OF JUDICIAL REASONS FOR EXTRA-GUIDELINES SENTENCES: MARYLAND TEST SITES



Data Set: Maryland scoresheets

Base: All cases with extra-guidelines sentences and with disposition dates between June 1, 1981 and February 28, 1982 and sentencing dates before April 1, 1982 (N=712).

7.2.5 Nature of Reasons for Extra-Guidelines Sentences (Maryland)

What kinds of reasons did the Maryland judges provide for the extra-guidelines sentences they rendered on scoresheet cases during the test year? As we have discussed before, sentences outside the guidelines ranges are expected under most guidelines systems, for cases with aggravating or mitigating circumstances relative to the factors considered in the scoring. Under those circumstances, judges are expected to record the factor(s) that they believed warranted a sentence outside the recommended range. It is also possible for guidelines systems to allow or accept reasons that challenge the weights given to the factors contributing to the guidelines scores. Not all systems permit such reasons, but in the multijurisdictional field test neither of the Advisory Boards limited the types of reasons judges could use to justify departures from the guidelines.

In this section, we analyze the content of the reasons Maryland judges did provide when sentencing outside the recommended ranges. The Maryland scoresheet data set contains 218 separate response codes, many of them copied verbatim from scoresheets. Little categorization was done by the staff, so that there is substantial room for classifying and interpreting the responses. Because there are many ambiguities in wording, and since these responses are analyzed apart from the other information on the scoresheets, a note of caution is appropriate. It is probably best to take the results discussed here as illustrative of the potential value of reasons as a feedback mechanism for revising guidelines over time.

As did the analysis of the reasons given by the Florida judges for extra-guidelines sentences, the discussion here separates reasons associated with above-guidelines sentences from those explaining sentences below the recommended ranges. Also, the distinction is made again between reasons that challenge the guidelines (by citing factors already considered in the scoring or in the system's scope) and those that concern aggravating or mitigating circumstances. It may well be that the challenging reasons are the more direct feedback on the guidelines, in that they comment on aspects of the operating system. On the other hand, an aggravating or mitigating factor cited often enough could suggest an addition to (or other revision of) the scoring.

Table 7-10 presents a classification of the reasons provided by Maryland judges for below-guidelines sentences. Of the total, 44 percent of the reasons challenge the guidelines and 56 percent cite mitigating factors.[37] Plea negotiations represent the largest group of responses; they account for over three-fourths of the challenges. Offender characteristics -- most notably "first-time offenders" and other aspects of prior record -- are the other sizeable source of challenging reasons.

Concerning the offender characteristics that represent over a third of the mitigating factors for sentencing, two themes emerge: the potential for offender rehabilitation and special offender needs. Employment record (which was eliminated

CONTINUED

3 OF 4

Table 7-10

Reasons Given for Below-Guidelines Sentences: Maryland Test Sites

Reasons[a]	Number of Times Cited	Percent of Category	Percent of All Reasons
Challenges to the Guidelines	340	100.0	44.4
1. Plea Negotiations	268	78.8	35.0
2. Offender Characteristics	48	14.1	6.3
3. Offense Characteristics	13	3.8	1.7
5. Unexplained Rejection of Guideline Sentence	11	3.2	1.4
Mitigating Circumstances	426	100.0	55.6
1. Offender Characteristics	168	39.4	21.9
2. Case-Related Considerations	142	33.3	18.5
3. Offense Characteristics	68	16.0	8.9
4. Recommendations of Other Parties	33	7.8	4.3
5. Other Factors	15	3.5	2.0

Data Set: Maryland scoresheets

Base: All cases with reason given for below-guidelines sentence (N=551).

Note: Each case could have two reasons entered in the data set.

a. Reasons categories explained in text.

from the Maryland guidelines scoring after the first three months) is frequently mentioned, as is the defendant's youth or need for a therapeutic program rather than imprisonment. Case-related considerations often involve weaknesses in the state's case or victims reluctant to prosecute or testify. In many instances, a defendant's cooperation with the state, or the fact of restitution already made, is cited in reducing the sentence. Cases in which defendant participation in the crime was minimal, or in which the offense was less serious than the charge would indicate, form the majority of offense-related mitigations. However, "circumstances of the case," a commonly given reason, is singularly uninformative. Finally, the recommendations of Parole and Probation, as well as those of the victim, are taken into account in some below-guidelines sentences.

Turning to reasons given for above-guidelines sentences, Table 7-11 shows them to be evenly divided between those challenging some aspect of the guidelines and those offering distinct aggravating factors. The most frequent challenge, again, was plea negotiations; this time, the judges were rejecting pled-down charges. Among offender characteristics, prior record -- either repeat offenses of that same nature or length of record -- was most frequently cited; offender legal status also appeared. Offense characteristics were heavily weighted toward victim injury. In addition, several judges simply indicated "guidelines too low" or "should be higher," without further explanation.

Reasons related to offense characteristics other than those counted in the guidelines scoring included the severity of the offense ("victim terrorized," "extremely vicious crime") and "real offense" considerations. These latter entail other counts that were dropped, pending accusations against the defendant, and one dispute with a jury decision on a prior conviction. One reason alleges that "the defendant was probably distributing, not just possessing, drugs." Of course, the legitimacy of enhanced sentences given for untried offenses and unproven facts may be questioned. Judges cited dangerousness and incorrigibility, as well as the defendant's primary role in a crime, among aggravating offender characteristics. A substantial number of other factors -- often vague, as in "circumstances of the case" -- also were offered to explain above-guidelines sentences.

The observation made with respect to reasons articulated by Florida judges also applies here. Lack of specificity will hamper or preclude the use of reasons to revise guidelines for the future. Where guideline systems incorporate some form of review or appeal on sentences, vague reasons are also likely to increase the number of appeals of extra-guidelines sentences.

In meetings at the end of the test year to modify the Maryland guidelines for continuing use, the quality of reasons was raised as an issue. Consideration was given to including sample reasons and a list of invalid reasons in the new guidelines manual. Among the reasons proposed as invalid were "plea bargain without reasons," "background of offender without specifics," "circumstances of the case without

Table 7-11

Reasons Given for Above-Guidelines Sentences: Maryland Test Sites

Reasons[a]	Number of Times Cited	Percent of Category	Percent of All Reasons
Challenges to the Guidelines	90	100.0	50.0
1. Plea Negotiations	49	54.4	27.2
2. Offender Characteristics	29	32.2	16.1
3. Offense Characteristics	8	8.9	4.4
4. Unexplained Rejection of Guidelines Sentence	4	4.4	2.2
Aggravating Circumstances	90	100.0	50.0
1. Offense Characteristics	40	44.4	22.2
2. Offender Characteristics	22	24.4	12.2
3. Other Factors	28	31.1	15.6

Data Set: Maryland scoresheets

Base: All cases with reason given for above-guidelines sentence (N=114).

Note: Each case could have two reasons entered in the data set.

a. Reasons categories explained in text.

specifics." In addition, the list included a number of factors ruled out because they were already covered in the guidelines. Thus, the Maryland Advisory Board and project staff are clearly aware of the issues raised here with respect to the substance of judges' reasons and the implications not only for using them as a feedback mechanism, but also for strengthening the guidelines as a system for increasing uniformity of sentencing.

7.3 Summary of Compliance Findings

This chapter has provided the detailed answers to a series of questions concerning procedural and substantive compliance during the multijurisdictional sentencing guidelines field test. The concept of compliance covers a range of issues from the paperwork required under a guidelines system -- completing scoresheets for each eligible case, providing judicial reasons for extra-guidelines sentences -- to the ways that judges use guidelines in decision-making and the patterns of their sentences relative to the recommended ranges.

Much of the analysis presented in this chapter is dependent on the scoresheet data sets constructed by the Florida and Maryland guidelines project staffs. As Chapter 5 described, the two projects took rather distinct approaches to the monitoring of scoresheet completion and to central data collection; they also handled errors somewhat differently. These practices inevitably had some consequences for the reliability of the data sets and of the conclusions that can be drawn from analyzing the scoresheets. However, we do not believe that our conclusions about compliance would be fundamentally different, had project monitoring been more thorough but the guidelines systems remained voluntary.

Table 7-12 provides a capsule summary of the findings about compliance. Points of similarity between the Florida and Maryland results include:

- General awareness of the guidelines on the part of judges in the test sites; and
- A tendency for the judges to give very vague reasons for extra-guidelines sentences, reducing the value of reasons as a feedback mechanism for guidelines revision.

The two states' compliance experiences also differed in some notable ways:

- There was far more variation among the Florida sites than among the Maryland ones, with respect to judicial use, filing rates, the relationship of actual to recommended sentences, and the provision of reasons;

Table 7-12

SUMMARY OF COMPLIANCE FINDINGS
FOR FLORIDA AND MARYLAND

Compliance Issue	Florida Findings	Maryland Findings
Did judges consult the guidelines in sentencing?[a]	General awareness of guidelines but sharp differences in use. Strong jurisdictional variation; minimal use in urban court, related to high volume of plea negotiated cases.	General awareness and use of guidelines, though variations in their role in the sentencing decision. Reports of over-compliance with recommended ranges.
Were scoresheets completed on eligible cases?[b]	Scoresheets were filed on only about 60 percent of all burglary cases; there were striking jurisdictional differences.	Scoresheets were available for just under 70 percent of all burglary cases, with some variation among circuits.
Did sentences fall within the guidelines ranges?	Scoresheet sentences were within guidelines for 78 percent of all cases (from 76 to 84 percent by circuit). There were significant differences in the direction of extra-guidelines sentences, by jurisdiction and crime category. Incarceration rates also varied widely.	Scoresheet sentences were within guidelines for 68-72 percent of all cases, with virtually no site variations in overall agreement or in the direction of extra-guidelines sentences.
Were reasons provided for extra-guidelines sentences?	Nearly 80 percent of scoresheets with extra-guidelines sentences gave reasons. There were significant site differences (with a range of 74 to 95 percent).	Just over 50 percent of the cases requiring reasons actually had reasons on the scoresheet; there was some variation by site and type of crime.
What kinds of reasons were given for extra-guidelines sentences?	Reasons given for above-guidelines sentences often challenged the guidelines (by citing factors already in the scoring), but mitigating circumstances were provided for most below-guidelines sentences. Reasons were frequently lacking in specificity.	A substantial proportion of reasons for both aggravated and mitigated sentences challenged by guidelines (by citing factors already counted in the scoring). Reasons were often too vague to provide useful feedback for guidelines revision.

[a]Primary discussion in Chapter 5.

[b]Findings subject to the limitations of the scoresheet data sets collected by the Florida and Maryland projects. See discussion of data quality in Chapter 7 and of monitoring in Chapter 5.

- The scoresheet filing rate of 70 percent in Maryland was considerably higher than the rate of 57 percent in Florida;

- A lower proportion of scoresheet sentences fell within the guidelines ranges in Maryland than in Florida, despite the width and overlap of the Maryland ranges;

- Reasons accompanied a substantially higher proportion of cases with extra-guidelines sentences in Florida than in Maryland; and

- Reasons given by Maryland judges were somewhat more likely to challenge the guidelines -- that is, to cite factors already taken into account in the offender or offense scores.

The implications of these similarities and differences have been discussed throughout this report. In essence, they raise questions about the meaning and viability of a guidelines system that is voluntary in terms of participation and without rules or limitations with respect to the degree of deviation from the ranges or the reasons for doing so. This is why the question of mandate must be one of the central issues in designing guidelines systems to bring about real change in sentencing.

FOOTNOTES

1. William D. Rich, L. Paul Sutton, Todd R. Clear and Michael J. Saks, Sentencing by Mathematics: An Evaluation of the Early Attempts to Develop and Implement Sentencing Guidelines (Williamsburg, VA: National Center for State Courts, 1982), p. 91.

2. For the Florida scoresheets see Appendix D, for the Maryland scoresheets Appendix E.

3. The burglary data are described fully in Appendix F. They are also utilized in the sentencing impact analysis (Chapter 8).

4. Guidelines coverage is discussed in Chapter 4.

5. The Final Report of the Florida Multijurisdictional Sentencing Guidelines Project indicates that, out of a total of 3,327 scoresheets, 170 scoresheets received in Tallahassee were for non-guidelines offenses or used the wrong grid. These cases were omitted from the scoresheet data set provided to Abt Associates. In addition, 214 forms are excluded from this compliance analysis because they lack sentence data (204) or because sentencing occurred after the test year (10).

6. Where multiple offenses were sentenced at the same court appearance, the most serious one was listed as "primary".

7. An attempt to use published Florida court statistics on criminal dispositions to assess the rate of scoresheet filing for all eligible offenses did not prove fruitful. Based on 1981 figures published in the Florida Judicial System Statistical and Program Activity Report: 1980 and 1981 (Tallahassee: Office of the State Courts Administrator, June 1982), the overall filing rate would be 41 percent and would range from 33 to 59 percent by jurisdiction. However, in addition to differences in the time period (calendar versus guidelines year) and in how cases are counted from circuit to circuit, there may be variations in caseload composition that affect the proportion of cases eligible for the guidelines. Finally, the Office of the State Courts Administrator has undertaken but not yet completed a program to assist the circuits in improving the quality and uniformity of the statistics they report.

FOOTNOTES
(continued)

8. Indeed, more scoresheets than burglary cases appear in the 15th. The most probable explanation is that the scoresheets were not consolidated to match sentencing events in all cases. While consolidation has been checked in the burglary data, the scoresheet data set is missing enough identifiers (especially sentencing date and defendant's date of birth) to make our check less than fully effective in identifying unconsolidated scoresheets. This could account for some or all of the 13-case discrepancy.

9. There are evident discrepancies in some of the Table 7-1 cell counts, with scoresheets outnumbering burglary cases. Aside from the consolidation question (previous footnote), these may be accounted for by the following rule used in collecting the burglary data. Where records did not clearly identify whether the burgled dwelling was occupied or whether the burglary was of an unoccupied dwelling vs. another structure (in each instance, they share a statute citation), the less serious primary offense was assumed.

10. Because it was thought that lack of preparation might have reduced filing in the early months of guidelines use, an analysis of filing rates was done for different segments of the test year: months 1 through 3, 4 through 9, and 10 through 12. The rates for months 4 through 9 of the year (the 6-month period after start-up and before any reduced effort in anticipation of the year's end) were virtually the same as those shown in Table 7-1. The rates for the first three months were higher than for the later periods, except in the 15th Circuit; the rates for the last three months were lower. Rather than showing an improvement as problems were resolved, the rates seem to indicate a continuing decline in filing as the year wore on.

11. PSI preparation is only mandatory for first-time offenders and is frequently waived for others.

12. Matching was performed on the basis of four information items: circuit, judge, sentencing date, and defendant's date of birth. (Neither personal nor docket identifiers were available on the scoresheets as alternatives.) When 1,164 burglary cases were matched against 695 burglary scoresheets, a match rate of 68.2 percent of the scoresheets resulted (N=474). The analysis presented here contrasts the characteristics of the 474 burglary cases with scoresheets to those of the 690 burglary cases without scoresheets.

13. The burglary data collection form, sources, and methodology are reviewed in Appendix F.

14. Of the entire scoresheet data set constructed by the Florida project staff, some 204 of the 3,157 cases (6.5 percent) were missing information on the sentence actually given. These cases could not, of course, be used for analysis.

15. The results presented in Figure 7-3 and subsequent tables and figures differ, sometimes substantially, from those presented in the Florida project's Final Report. One reason for these differences is clear. In their tabulations, the Florida project staff excluded scoresheet cases missing judges' reasons. Since all such cases fall outside the guidelines, excluding them necessarily raises the apparent rate of agreement between actual sentences and the grids. (A possible further source of discrepancy is the unexplained gap of 54 cases between the tables and the technical appendix (Appendix G) describing the project's sample. See Multijurisdictional Sentencing Guidelines Project: Final Report (Tallahassee, FL: Office of the State Courts Administrator, July 1982), p. G-2.)

16. There may be errors in score addition in the scoresheet data that could slightly affect this analysis. In 35 cases (1.2 percent), the point total did not match the sum of the point elements. (Such discrepancies could also be due to errors in data entry.) For 20 of the 35 cases, the re-summed score fell in a different guideline grid cell than the original; for 14, this changed the sentence from "in" to "outside" the guidelines. However, since a very small number of cases is thus affected, the original point totals are used throughout this analysis.

17. Incarceration rates by guideline cell are shown in Appendix J, Table J-1.

18. See Figure 7-3.

19. This pattern also holds by crime category; for every grid, above-guidelines sentences were more often accompanied by a reason than below-guidelines sentences. See Appendix J, Table J-2.

    One other finding about the presence or absence of reasons should be noted. In a number of the scoresheet cases, reasons were given even though the sentence appears to fall within the guidelines. These cases amount to .7 percent of all the within-guidelines cases. Assuming that the data set contains correctly-entered primary offense identifiers and point scores, it would follow that errors were made by judges or other personnel in using the grids for these cases. Of course, the same kind of error may account for the absence of reasons in some portion of the extra-guidelines cases.

20. Florida Final Report, p. E-1.

21. Preliminary Report on the Development and Impact of the Minnesota Sentencing Guidelines (St. Paul, MN: Minnesota Sentencing Guidelines Commission, 1982), p. 55.

22. This assumes that the coders of the reasons did not discard additional information. Since the categories in the scoresheet data set are so detailed, it seems a safe assumption.

23. The score elements corresponding to this reason give added points for multiple counts of the primary offense and for all additional offenses at conviction.

24. As in the Florida analysis, we cannot assess the degree to which scoresheets may have been prepared but not filed with the guidelines project staff. However, as the discussions on error-checking and monitoring by the Maryland project showed (Chapter 5), there is reason to believe this was not likely to occur.

25. Here and throughout this section, data actually refer to the first nine months of the guidelines test year. The data set of multiple-count scoresheet cases provided to Abt Associates by the Maryland guidelines project staff was not complete for the full year, and since these cases differed significantly from single-count scoresheet cases, a bias would have been introduced by simply combining the files. Therefore, the analytic sample used here covers cases with convictions before March 1, 1982 and sentencing before April 1, 1982. (The guidelines year began June 1, 1981.) This adjustment, and the exclusion of cases with critical variables missing, reduce the sample from 2,431 to 2,217 (8.8 percent).

26. Maryland Sentencing Guidelines Manual (Annapolis, MD: Administrative Office of the Courts, June 1981), Appendix A.

27. A comparison of scoresheet volume with published total criminal case termination statistics did not prove valid. Striking differences in how the sites count cases, as well as inability to separate cases originating at the circuit level (and thus covered by guidelines) from those prayed for jury trial from district court, were the main confounding issues. Using figures from the Annual Report of the Maryland Judiciary: Statistical Abstract 1981-1982 (Annapolis, MD: Administrative Office of the Courts), scoresheet filing rates would range from 12 to 29 percent by site and average 24 percent overall.

28. The scoresheet data set shows a small number of separate cases that might have been appropriately consolidated, but the impact on this analysis would be minimal. By not doing so, we err on the side of possibly overstating the rate of scoresheet filing.

29. The data comparison spans the period September, 1981 through March, 1982 (months 4 through 10 of the guidelines test year). It is thus likely that any start-up effects have been excluded.

30. Appendix Table J-3 shows details by type of burglary. Nine separate types of burglary are grouped into three levels of seriousness (corresponding to Categories 4 through 6 in the Advisory Board's ranking).

31. Nevertheless, plea negotiation is still the most likely reason for missing scoresheets in Baltimore City and Prince George's County. It is often estimated that pleas account for 85 to 90 percent of all felony dispositions in the former and 85 percent in the latter. If we estimate the number of additional plea negotiated cases that might be expected in Baltimore City and Prince George's County based on the 85 percent estimate (by inflating the figures for percent disposed by plea in Table 7-9), we find that about 150 more such cases might have been disposed in the former jurisdiction and around 60 in the latter. This would increase the total volume of scoresheets in these sites by 10 to 11 percent, over the entire guidelines year.

32. For Maryland scoresheets, see Appendix E.

33. According to this check, 129 of the 2,217 cases (5.8 percent) belonged in a different category (i.e., within, above, or below the guidelines range) than that indicated by the project's variable. Where correcting errors of score addition would improve the match, this was done before the final results were generated.

34. See Appendix J, Table J-4.

35. However, the guidelines grids themselves appear to be better calibrated for the range of offenses and offender characteristics encountered during the test year. There was not the marked concentration of cases in any single portion of the grids which was noted for Florida (see Table 7-5).

36. Details by the three major groups of crimes are provided in Appendix J, Table J-5.

37. Note that, to include as many reasons as possible in this analysis, the time restrictions that define the analytic sample were relaxed. With a sample of 3,553 rather than 2,217, a total of 946 reasons can be examined instead of 523. (This case base stretches beyond the guidelines test year for single-count scoresheets but ends before the full year for multiple-count cases.)

CHAPTER 8

THE IMPACT OF THE GUIDELINES ON SENTENCING

The most fundamental goal of sentencing guidelines is to increase uniformity of sanctions for like offenses and offenders, both among individual judges and across different jurisdictions. In this chapter, we examine directly whether the multijurisdictional guidelines developed and implemented in Florida and Maryland changed sentencing patterns. The primary question is whether disparity was reduced in the participating sites during the test year. A secondary question concerns change in sentence severity, though the guidelines were not established in order to alter prevailing severity of sanctions. Still, because the severity of criminal sanctions has become a prominent public concern in the past decade, we also analyze whether there were significant changes in the level of sentences received by similarly situated offenders during the guidelines test.

In the abstract, sentencing guidelines should reduce unwarranted disparity in sanctioning. The explicit sentence ranges for specific combinations of offender characteristics and convicted charges should ensure that similar sentences are imposed (in the absence of aggravating or mitigating factors). In reality, sentencing guidelines should reduce disparity if:

- the guidelines' ranges are narrower than prevailing practice, and the scoring includes all the important case and offender characteristics;
- there are no significant changes in charging or plea negotiating, and
- there is substantive and procedural compliance on the part of judges and other actors.

Thus, our expectations about changes in sentencing should be informed and tempered by what we already know about development, implementation, system adaptation, and compliance during the guidelines test.

With respect to guidelines development, two facts bear on the question of reduced sentence disparity. We noted in Chapter 4 that the Maryland grids had very wide and overlapping ranges,[1] which would permit considerable variation even if sentences adhered to the ranges. Also, we have seen that the greatest majority of Florida cases fell into the least serious guideline cell, with recommended sentence ranges of 0 to 18 months or 0 to 36 months, depending on crime category. Thus, even

guidelines sentences were minimally constrained for these cases. A number of findings concerning implementation and system adaptation might also affect sentence uniformity during the guidelines test year. For one thing, the early months were certainly a shake-down period with respect to procedures. Second, the narrow judicial representation on the Florida Advisory Board and the limited training effort there did little to enlist support for the guidelines effort. By contrast, Maryland's judicial representation and the training approach taken there produced broader familiarity with and support for the test. Third, there was strong judicial resistance in Florida's largest site, and prosecutorial resistance in both states' suburban jurisdictions as well as in Maryland's urban site. In West Palm Beach, the announcement of a "no charge bargaining" policy led to reduced pleas, increased trials, and a combination of below-guidelines sentences and score bargaining to relieve caseload pressures. In Prince George's County, defense reaction to the shift in prosecutor sentence recommendations ultimately brought about withdrawal of the State from negotiations and greater cooperation between judge and defense. All these system changes make it much more difficult to predict whether sentencing patterns would change and in what direction.

Finally, it seems clear that the record of compliance under the test's voluntary, judicial guidelines could influence sentencing impact. Overall, compliance was lowest where differences between the guidelines and local practice were greatest. For some 40 percent of Florida's eligible burglary cases and 30 percent of Maryland's, scoresheets were not filed (or there were "fatal" errors); the proportion reached 60 percent in two Florida sites, and nearly half in one Maryland jurisdiction. It is likely that these cases without scoresheets were sentenced without reference to the guidelines. Further, a substantial share of guidelines cases (22 percent in Florida, 30 percent in Maryland) received sentences outside the recommended ranges. There were no standards for how far such sentences could vary from the upper and lower boundaries of the range. The potential impact of guidelines on disparity might well be reduced by both these types of non-compliance.

Of course, observation of a single year's data in an experimental situation cannot tell us with certainty what the long-run effects of an actual program would be. In particular, we cannot predict future revisions and adaptation in a way that lets us adjust for them statistically in estimating sentence impacts. Therefore, the analysis presented here must be understood as limited in time and place. There may well be more information about the future of guidelines--or how to make guidelines work in the future--and more indication of their value in the preceding analyses of development, implementation, adaptation and compliance than in the examination of sentencing impact.

8.1 Introduction to the Impact Analysis and Its Findings

This section briefly introduces our approach to analyzing sentencing change under the multijurisdictional field test. It also summarizes the findings (which are presented in more detail in the remaining chapter sections and in supporting appendices).[2]

Two approaches were taken to the quantitative analysis of sentencing change. The first, which relied on actual court data from the participating sites, is the focus of this chapter. A second analysis based on a sentencing simulation exercise completed by judges in Maryland and Florida was also conducted. The use of hypothetical (simulated) cases was intended to supplement the analysis of actual court data and narrow the range of possible interpretations of sentencing change. In examining actual sentencing decisions, we cannot control for all the operating variables in judges' sentencing decisions; there are always differences among cases in real life that could account for discrepancies in sentences among judges. The use of identical case descriptions, however, theoretically permits complete control of all offense, offender, and case processing variables--that is, all potentially warranted case factors that may influence choice of sentence--and this control, in turn, may allow more accurate assessment of the contribution of judicial differences to sentencing decisions. But there are also drawbacks to the simulation technique. Whether the cases are representative of the usual business of the court, or whether sentences rendered in hypothetical cases provide an adequate approximation of what would happen in the courtroom are two of the most serious challenges to the validity of this approach. A full discussion of this methodology, and the results of our analysis, are presented in Appendix H.

Our main examination of changes in sentencing patterns relies on actual court data from Maryland and Florida. The data were independently collected by Abt Associates from docket books and other sources. In contrast to the data from the guidelines scoresheets, which were examined in the compliance analysis in Chapter 7, the quality and accuracy of actual court data are not affected by problems in implementation or compliance. These data do not reflect, e.g., changes in the "facts" of the case to support a score bargain or the omission of many cases from scoresheet preparation.

A second feature of this analysis is that it does not involve comparing sentences to the guidelines ranges. Instead, it relies on time contrasts and geographical contrasts to assess whether there were changes in sentencing patterns that may have resulted from the implementation of guidelines. The time contrast involve comparing data from before the guidelines went into use with data from the test year. This is a classic pre-post research design, where the intervention is the introduction of the multijurisdictional guidelines. We have also developed geographical contrasts, collecting data from court jurisdictions in Maryland and Florida that were not part of the guidelines test. These jurisdictions, termed the "comparison sites," are

included so that we can better separate observed changes related to the guidelines from changes affecting the state courts more widely.

The comparison sites were selected to match as closely as possible the characteristics of the participating sites. The matching considered population size and mix, urban/suburban/rural nature of the area, and for Florida the court structure of the jurisdiction (number of counties in the circuit). Figures 8-1 and 8-2 show the locations of the test and comparison sites. Table 8-1 presents key characteristics for each matched pair.

The resulting matches are by no means exact. In Maryland, the combined population size in the suburban test sites (Prince George's and Montgomery Counties) is over three times as great as that of the single suburban comparison site (Anne Arundel County). The reverse is true in Florida: the population in the suburban test site (15th Circuit) is only about half that in the comparison site (17th Circuit). Maryland's Baltimore City is far more urban in character than Baltimore County, but there is no other large city in the entire state. The rural panhandle comparison site in Florida (2nd Circuit), while close in size to its test counterpart, contains the state capitol.

The comparison sites are clearly not a perfect control for other factors that might produce changes in sentencing patterns. However, their inclusion does allow us to assess whether there may be exogenous factors producing shifts over time in sentences, and thus provides an important check on statistical statements about changes in the test sites over time.

Since the primary question about guidelines impact concerns sentencing uniformity, the analysis focuses on a measure of how much sentences vary, and whether the amount of variation changed with the introduction of guidelines. Some variation in sanctions results from differences in salient offender and offense characteristics; this is termed "warranted" variation. It is recognized by guidelines and, indeed, is built into their scoring system. Under guidelines, variation in sentencing would be expected to persist to the extent that it is the result of legitimate, or warranted, factors differentiating cases for scoring. However, the components of variation that are unwarranted or essentially random should be reduced. In the framework of the multijurisdictional test, unwarranted components would include factors such as offender race and sex, as well as the results of differences in local practice or in individual judicial views beyond the leeway of the guidelines ranges.

To measure this remaining variation and whether it decreased during the guidelines test, we use a statistical technique called multiple regression, which controls for known differences between cases. With this technique, we can examine changes in sentencing, once the effects of warranted offense and offender characteristics have been removed. This is done by using the residuals from the regression equations. The residuals represent the remaining differences in sentence outcome

FIGURE 8–1

Pairing of Florida Test and Comparison Sites

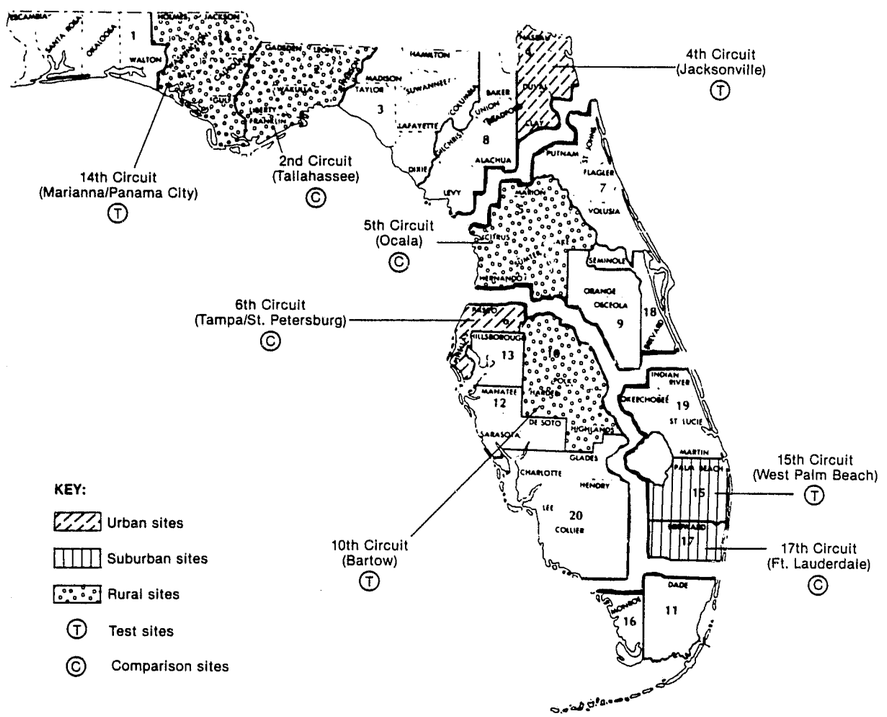


FIGURE 8—2

Pairing of Maryland Test and Comparison Sites

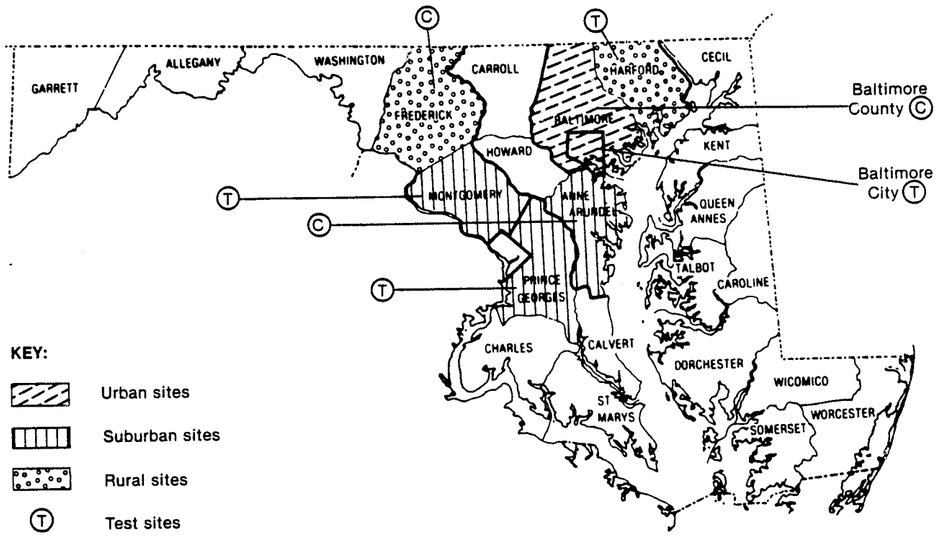


Table 8-1

Pairing of Test and Comparison Sites in Maryland and Florida

	Test Sites	Comparison Sites
MARYLAND		
Urban:	Baltimore City	Baltimore County
Pop. size:[a]	786,775	655,615
# Judges:[b]	23	12
Suburban:	Montgomery and Prince Georges Counties	Anne Arundel County
Pop. size:	1,244,124	370,775
# Judges:	25	9
Rural:	Harford County	Frederick County
Pop. size:	145,930	114,792
# Judges:	4	2
FLORIDA		
Urban:	4th Circuit (Jacksonville)	6th Circuit (Tampa-St. Petersburg)
Pop. size:	670,949	922,174
# Counties:[c]	3	2
# Judges:[d]	25	28
Suburban:	15th Circuit (West Palm)	17th Circuit (Ft. Lauderdale)
Pop. size:	576,863	1,018,200
# Counties:	1	1
# Judges:	22	40
Rural:	14th Circuit (Marianna)	2nd Circuit (Tallahassee)
Pop. size:	186,078	223,731
# Counties:	6	6
# Judges:	5	8
Rural:	10th Circuit (Bartow)	5th Circuit (Ocala)
Pop. size:	388,557	350,802
# Counties:	3	5
# Judges:	14	9

[a] Population figures are from 1980 census data.

[b] The number of judges as of June 1982, Annual Report of the Maryland Judiciary 1981-1982, Administrative Office of the Courts.

[c] The number of counties is only relevant to Florida, since each site in Maryland is composed of one county.

[d] The number of judges is an approximate number of Circuit Court judges in Florida as of May 1982.

once measured differences in the cases (e.g., in offender prior record) are taken into account.

Regression models are particularly useful for assessing changes in the uniformity of sentencing. They provide an estimate of the residual standard deviation, i.e., the standard deviation of the error term in the model (the variation remaining after controlling for the case characteristics). The standard deviation is a commonly used measure of the variation, or spread, of a set of values around the average value. The residual standard deviation is interpretable as the standard deviation for cases with a given set of characteristics. It represents the amount of sentencing variation that is left after the effects of known or measured factors are removed. Thus, in principle, it serves as a natural measure of uniformity (actually of non-uniformity), with the caveat that other warranted factors which we have not taken into account may be responsible for some part of this residual variation.[3]

This approach to analyzing changes in sentence variation will be applied to the actual court data: namely, data on all the burglary cases sentenced in the test and comparison sites during the guidelines year and the preceding twelve months. The decision to focus exclusively on burglaries was based on three considerations:

- the desire to examine closely a relatively homogeneous family of offenses, so that a shared set of factors might be expected to explain much sentencing variation;

- the need to choose a family of frequently occurring offenses, so that the analysis would be based on substantial numbers of cases from every site; and

- the goal of collecting data on all the cases of this offense group, rather than a sample, so that general statements could be made, and so that the compliance analysis could be conducted.

The burglary data offer several advantages for studying whether changes in sentencing practices were brought about by the guidelines. First, the burglary data are based on actual case experience and may therefore be regarded as valid indicators of actual sentencing practice. Second, the burglary data were collected unobtrusively; the data gathering did not affect the sentencing process, since routine case records generated in the ordinary course of legal system operation were used. Third, by restricting attention to burglary cases, we are dealing with a relatively homogeneous class of crimes, so that there are fewer sources of variation in sentencing outcome among cases. This should allow a closer focus on residual, unwarranted variation and whether it was reduced under the guidelines test.

Against these advantages of the burglary data, two important limitations should be recognized. First, we cannot be certain that the conclusions we reach hold for all other types of crime. Burglary is widely viewed as a typical offense in terms of the factors relevant to sentencing and the patterns of case processing. Burglaries constituted nearly a quarter of all guidelines offenses that were scored in Florida and Maryland. However, both sets of guidelines treated crimes against persons somewhat distinctly from property crimes, through additional scoring factors and/or separate grids. To the extent that sanctions for crimes against persons, particularly violent ones, are typically affected by different judicial considerations or case handling, results of the burglary analysis can only speculatively be extended to them.

The second significant limitation concerns the information items collected on each case. They include the most important factors identified in the sentencing literature (e.g., type of offense, number of counts, additional convicted offenses, and offender's prior record, status, age, race and sex), but they by no means cover all the factors relevant to the case or available to the judge when sentencing. Nor do they include all the specific factors scored under the Florida and Maryland guidelines, some of which could not be obtained from court records. As a result, the residuals which are our measure of sentencing variation, and from which the effects of the case characteristics we do know have been removed, clearly still contain some variation due to legitimate factors not known to us. Thus, the residuals cannot be viewed simply as unwarranted variation. Nevertheless, if the guidelines test succeeded in increasing sentencing uniformity by reducing unwarranted variation, the residuals will reflect this reduction relative to the comparison site data and to the year prior to guidelines use.

Findings on Sentencing Impacts

On the basis of this analysis of burglary data, our main findings on sentencing impacts are these:

In Florida:

- A great deal of the variation in sentences was not accounted for by the offense and offender characteristics cited as most important in the literature, although these factors did prove influential.

- The variability of sentences actually increased over time. That is, sentences were less uniform in both the test and comparison sites during the test year than during the year before. Thus, guidelines clearly did not reduce unwarranted sentence disparity.

- There were substantial differences in sentence severity across the sites, both in the level of the sanctions and in how they changed over time. In the urban test site, already known for its heavy sentences before the guidelines effort, severity actually increased more than in the other sites.

- Overall, site differences persisted--and were even amplified--despite the guidelines effort. The four test sites grew further apart. The Florida guidelines, as designed and implemented in the field test, did not begin to mitigate differences in local sentencing practice.

Maryland

- The data on burglary cases sentenced during the guidelines year show a statistically significant reduction in sentence variation in the test sites over time, but not in the comparison sites. Changes in sentencing in Baltimore City are the key to the reduction (which is not observed in the other three sites singly or as a group).

- Sites differed in typical levels of sentences and in severity changes over time. There continued to be substantial differences among jurisdictions, most notably in severity (but also in uniformity) of sanctions.

- The guidelines' impact on sentencing thus appears to be largely within one jurisdiction, rather than multijurisdictional. However, such an impact in a large, urban site suggests that guidelines have a potential for sentencing change and that how they are implemented can make an important difference.

Further discussion of these findings, their limitations and meaning, is found in the balance of this chapter, which also presents details of the analysis.

8.2 Sentencing Impacts in Florida

This section provides a more detailed account of sentencing changes in Florida during the guidelines test. It first describes the characteristics of burglary cases in the test and comparison jurisdictions and the sentences rendered on them in the year prior to guidelines introduction as well as the test year. Contrasts in sentence uniformity and severity, controlling for case differences, are then explored.

Typical sentences for burglaries covered a considerable range in the test sites in the pre-guidelines period, as Figure 8-3 illustrates. The figure shows the values of the average net sentence (net incarcerative time)[4] for each individual site, for all test sites together, and for all comparison sites together. These average sentences combine cases with differing characteristics (e.g., types of burglary or offender prior records), although the differences are narrowed by concentrating on burglary cases only. Comparing the guidelines period to the pre-implementation period, the mean sentence for burglary cases in the test sites increased slightly, from 21.0 to 24.3 months. However, the comparison site sentences changed little, increasing only from 18.5 to 19.7 months.

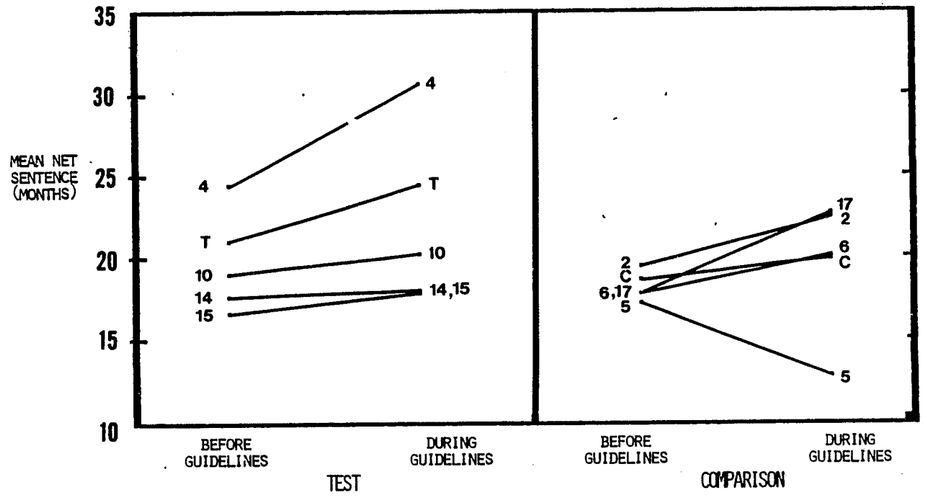
Guidelines are primarily concerned with the uniformity of sentencing among like cases, among judges, and among jurisdictions. Statistically, we measure sentencing variability and say that decreasing variability means increasing uniformity. The indicator of variability--still considering all burglary cases combined--is the standard deviation of the net sentence, a measure of how much individual sentences vary from the mean.[5] Standard deviations of the burglary sentences for individual sites and grouped test and comparison sites are shown in Figure 8-4. The means were shown in Figure 8-3. The patterns for the grouped test and comparison sites are similar. Variation has increased somewhat for both sets of sites, from 34.4 to 37.7 months for the test sites, and from 33.2 to 40.1 months for the comparison sites. There are striking changes in variability for some of the latter, while changes are more moderate for individual test sites.

However, these figures still reflect important case differences considered legitimate sources of sentence disparity. Changes over time in sentencing variation or severity could result from a differing mix of cases between sites, or a shift in the case mix. Such changes must be sorted out from the possible effects of introducing guidelines. Thus, the main purpose of examining case characteristics is to explore the possibility that observed changes in sentencing are attributable to changes in the types of burglary cases between the two periods. We can also see whether offender and offense characteristics were alike initially between the test and comparison sites, and whether there were changes between the guidelines test period and the prior year. We therefore look at a series of characteristics gathered for all the burglary cases sentenced in these sites. The characteristics include the race and sex of the offender, a measure of the seriousness of prior record, offender legal status at the time of the offense, seriousness of the primary offense at conviction, and existence of multiple convicted counts or subsidiary offenses.[6]

Table 8-2 displays these offender and offense variables for the pooled test sites and comparison sites in the pre-guidelines and test periods. The race and sex composition of the offenders as a group, and their prior records, are quite similar, as are the characteristics of the offenses being sentenced. The test and comparison sites are least alike in the proportion of cases with multiple convicted counts.

FIGURE 8-3

CHANGES IN NET SENTENCES FOR BURGLARY:
FLORIDA TEST AND COMPARISON SITES



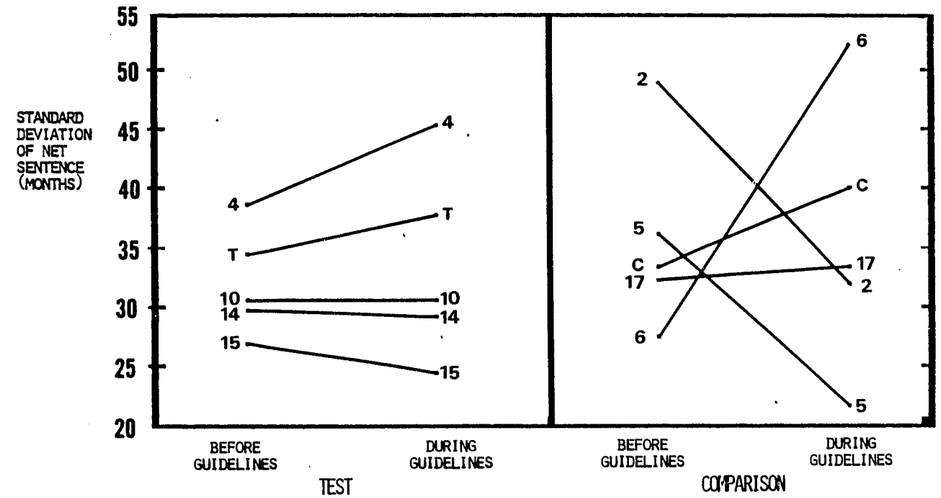
Key: 4 = 4th Circuit (Jacksonville)
15 = 15th Circuit (W. Palm Beach)
10 = 10th Circuit (Bartow)
14 = 14th Circuit (Marianna/Panama City)
T = Test Sites Combined
6 = 6th Circuit (Tampa/St. Petersburg)
17 = 17th Circuit (Ft. Lauderdale)
5 = 5th Circuit (Ocala)
2 = 2nd Circuit (Tallahassee)
C = Comparison Sites Combined

Data Set: Florida burglary data

Base: All cases with no missing values for any of the offender characteristics, offense characteristics, race, or sex.

FIGURE 8-4

CHANGES IN NET BURGLARY SENTENCE VARIATION:
FLORIDA TEST AND COMPARISON SITES



Key: 4 = 4th Circuit (Jacksonville)
15 = 15th Circuit (W. Palm Beach)
10 = 10th Circuit (Bartow)
14 = 14th Circuit (Marianna/Panama City)
T = Test Sites Combined
6 = 6th Circuit (Tampa/St. Petersburg)
17 = 17th Circuit (Ft. Lauderdale)
5 = 5th Circuit (Ocala)
2 = 2nd Circuit (Tallahassee)
C = Comparison Sites Combined

Data Set: Florida burglary data

Base: All cases with no missing values for any of the offender characteristics, offense characteristics, race, or sex.

Table 8-2

Case Characteristics: Florida Burglary Data

Percent of Cases With:	Test Sites		Comparison Sites	
	Pre-Guidelines Period	Guidelines Period	Pre-Guidelines Period	Guidelines Period
Black offender	35.4%	37.8%	30.8%	33.3%
Male offender	97.8	98.2	97.5	96.7
Multiple convicted counts	11.4	12.6	14.7	18.4
Additional (lesser) convicted offenses	29.6	39.2	29.6	35.4
Offender on restricted status	14.0	19.4	12.9	12.1
Average Offense Seriousness[a]	1.36	1.39	1.45	1.40
Seriousness of Highest Prior Adult Conviction[a]	.95	.96	.82	.72
Number of Cases	(1069)	(1147)	(1271)	(1560)

Data Set: Florida burglary data.

Base: All cases with no missing values for any of the offender characteristics, offense characteristics, race, or sex (N=5047).

Note: Case characteristics for each test and comparison site are shown in Appendix I, Tables I-2 and I-3. Variables are defined in Table I-1.

a. Seriousness is measured by statutory degree, as follows:

Current Offense	Prior Adult Conviction
3 = first degree felony	4 = first degree felony
2 = second degree felony	3 = second degree felony
1 = third degree felony	2 = third degree felony
	1 = misdemeanor
	0 = no prior adult conviction

Between the two time periods, the case profiles for the two groups of sites remained quite stable. The primary exception was a growing proportion of cases with lesser convicted offenses. There is less consistency at the individual site level; striking changes in the proportion of cases with lesser convicted offenses or the proportion with offenders on restricted status occurred in the 10th, 14th and 15th (test) circuits and in the 5th (comparison) circuit.[7] These changes suggest that controlling for case characteristics is critical to determining any effects of guidelines on sentencing.

We therefore turn to a regression analysis, which enables us to examine the remaining variation in sentences once the effects of the specific offense and offender characteristics are removed. In this analysis, we want to control as much as possible for warranted variation in sanctions, so that changes in unwarranted variation can be detected. It is unwarranted variation that should be reduced by guidelines use.

Table 8-3 presents the regression models for the two groups of sites before and during the guidelines test. Five case characteristics are entered as independent variables, and all five are shown to have significant impacts on net sentence. (We do not include race and sex here, because their effects on sentence are not legitimate. Their effects are discussed below.) Overall, these five factors explain between 20 and 25 percent of the variation in sentence among cases. The sizeable remaining variation results in part from warranted, but unmeasured, case differences, and in part from unwarranted sources. In Table 8-3, the measure of variation, the residual standard deviation "s," increases both in the test sites and in the comparison sites between the two time periods. This means that, taking into account known differences among cases, sentences were more varied during the guidelines year than before, in both the test and comparison sites. There was no reduction that could have resulted from the introduction of the guidelines.

Two of the most frequently identified factors contributing to disparity in sanctions are the race and sex of the offender.[8] Information on these factors was gathered for the burglary cases in order to enable a test of their influence on sentence variation. Inclusion of race and sex in the regression model showed that they were not significant influences on sentencing in the pooled test sites, either before or during the test period, but were important in the pooled comparison sites. They accounted for an additional 10 percent of sentence variation in the latter sites during the guidelines year.

Returning to the main equations (which exclude race and sex), the changes in residual sentencing variation between the two time periods are illustrated in Figure 8-5, which show the values for each site, as well as for the two pooled groups. Among the test sites, variability increased in the urban jurisdiction (Jacksonville) and in the pooled group. The slight reductions for the remaining test sites were not statistically significant.[9] Among the comparison sites, sentencing variation also increased in the urban jurisdiction (Tampa/St. Petersburg), while it decreased significantly in the

Table 8-3

Regression Equations of the Sentencing Decision:
Florida Burglary Data

Dependent Variable = Net Sentence

Independent Variables	Regression Coefficients: Test Sites, Pre-Guidelines Period	Test Sites, Guidelines Period	Comparison Sites, Pre-Guidelines Period	Comparison Sites, Guidelines Period
Multiple convicted counts	11.65**	15.21**	8.97**	9.48**
Additional (lesser) convicted offenses	6.37**	2.74	11.04**	10.41**
Offender on restricted status	15.80**	18.44**	20.23**	23.13**
Offense seriousness	16.04**	22.80**	17.16**	22.41**
Seriousness of highest prior adult conviction	8.16**	8.85**	6.50**	5.84**
Intercept	-13.99	-22.57	-18.76	-24.07
R^2	.211	.256	.253	.207
s	30.6	32.6	28.8	35.8
Number of cases	1069	1147	1271	1560

Data Set: Florida burglary data.

Base: All cases with no missing values for any of the offender characteristics, offense characteristics, race, or sex.

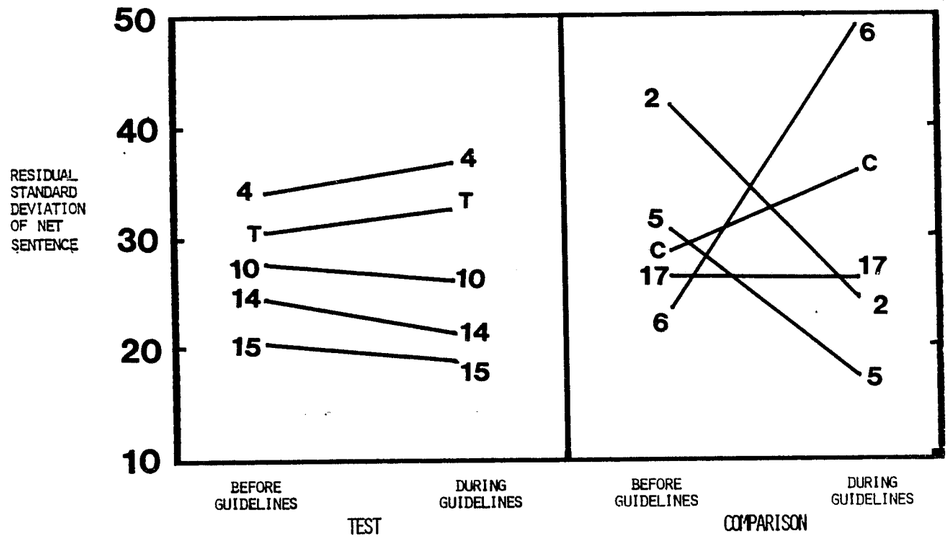
Note: Dependent variable is the net sentence (incarcerative time less suspended time) in months. Regression equations for each test and comparison site are shown in Appendix I, Tables I-4 and I-5.

Statistical Significance:

** $p \leq .01$
* $p \leq .05$
+ $p \leq .10$

FIGURE 8-5

CHANGES IN RESIDUAL VARIATION
IN NET BURGLARY SENTENCES:
FLORIDA TEST AND COMPARISON SITES



Key: 4 = 4th Circuit (Jacksonville)
15 = 15th Circuit (W. Palm Beach)
10 = 10th Circuit (Bartow)
14 = 14th Circuit (Marianna/Panama City)
T = Test Sites Combined
6 = 6th Circuit (Tampa/St. Petersburg)
17 = 17th Circuit (Ft. Lauderdale)
5 = 5th Circuit (Ocala)
2 = 2nd Circuit (Tallahassee)
C = Comparison Sites Combined

Data Set: Florida burglary data

Base: All cases with no missing values for any of the offender characteristics, offense characteristics, race, or sex.

Note: Residual standard deviations are from equations shown in Table 8-3 and Appendix I, Tables I-4 and I-5.

second and fifth Circuits. These contrasts make it clear that the introduction of the guidelines did not reduce unwarranted disparity in the test sites during the one-year test.

The regression models suggest that no real changes in overall sentencing disparity occurred in the test sites over time, or relative to the comparison sites. However, although disparity was not reduced, it is possible that other shifts in sentencing did occur. In particular, we are interested in changes in the severity of sanctions. To investigate this issue, we have constructed three prototypical burglary cases, ranging in characteristics from relatively minor to very serious. The definitions of these three cases appear in Table 8-4. Case A is the least serious; the offender has no felony record and was convicted of a single count of a third-degree felony. Case C is the most serious; the offender not only has a prior felony record, but was on parole or probation when the crime occurred, and the conviction includes multiple charges and subsidiary offenses. Race and sex are not controlled for in these cases.

These examples can be used to see whether changes in sentencing severity occurred for a case with the same measured characteristics before and during the test year. For these three cases, the characteristics from Table 8-4 have been substituted into the regression equation for each of the Florida sites, as well as into the equations corresponding to the pooled test and comparison sites. The results, predicted sentences, are displayed graphically in Figure 8-6 (A-C). For the least serious case (A), there are decreases in the predicted sanction in all the test sites except Jacksonville, and for the pooled test group as well. The comparison sites change little in severity, except for the reduced sanction in the 5th Circuit (Ocala). For the second case (B), there is a striking increase in severity in the urban test site, slight decreases in the other test sites, but an increase for the pooled group due to Jacksonville's influence. With the exception of the 2nd Circuit (Tallahassee), the comparison sites showed increases. Again, in the most serious case (C), the urban test site's sentence increased sharply, from 8 to 11 years. The other sites showed a mixture of more moderate increases and decreases.

These examples serve to illustrate a second finding about Florida sentencing practices during the guidelines test. Patterns of change showed little consistency across sites; there were different levels of predicted sanctions, and they moved in opposite directions. Further, the urban test site, known for severe sentences before the guidelines period, appeared to increase in sentence severity for more serious cases. This trend, as well as site differences, contribute to the slight rise in residual sentencing variability we observed.

Perhaps the most striking finding, in light of the multijurisdictional nature of the guidelines experiment, is the persistence and even amplification of site differences. While the individual test sites showed no statistically significant changes in residual variation, the test sites as a group showed slightly increased variation. It appears that the guidelines effort did not begin to reconcile the differences in local

Table 8-4

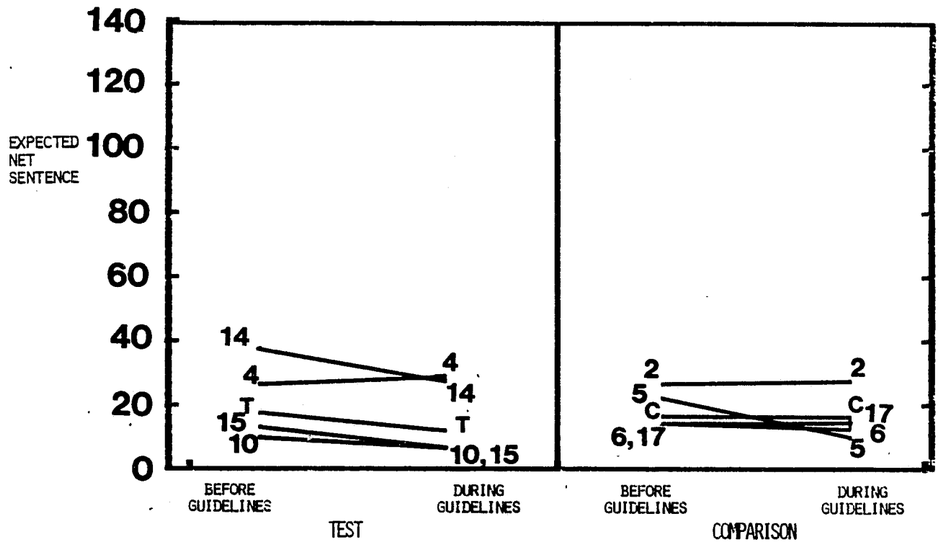
Definitions of Example Cases for Illustrating Sentencing Changes in Burglary Data

Case Characteristics	Case A	Case B	Case C
Multiple Convicted Counts	No	No	Yes
Offense Seriousness	Low	Moderate	Moderate
Additional (Lesser) Convicted Offenses	Yes	No	Yes
Offender on Restricted Status	No	Yes	Yes
Seriousness of Highest Prior Adult Conviction	Misdemeanor	Misdemeanor	2nd Degree Felony

FIGURE 8-6A

CHANGES IN EXPECTED NET BURGLARY SENTENCES FOR EXAMPLE CASES: FLORIDA TEST AND COMPARISON SITES

(CASE A)



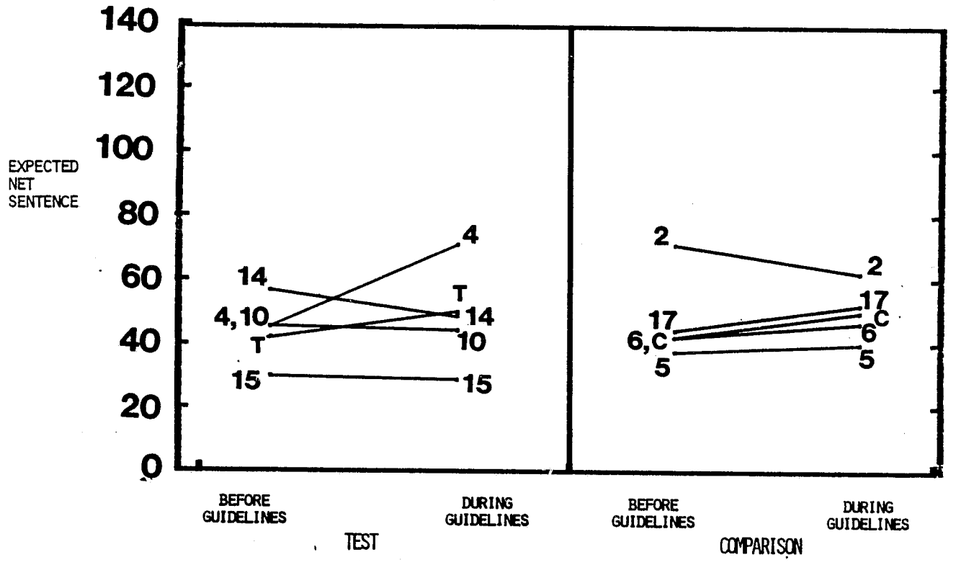
Key:
- 4 = 4th Circuit (Jacksonville)
- 15 = 15th Circuit (W. Palm Beach)
- 10 = 10th Circuit (Bartow)
- 14 = 14th Circuit (Marianna/Panama City)
- T = Test Sites Combined
- 6 = 6th Circuit (Tampa/St. Petersburg)
- 17 = 17th Circuit (Ft. Lauderdale)
- 5 = 5th Circuit (Ocala)
- 2 = 2nd Circuit (Tallahassee)
- C = Comparison Sites Combined

Note: Expected values of net sentence are based on case characteristics specified in Table 8-4 and equations shown in Table 8-3 and Appendix I, Tables I-4 and I-5.

FIGURE 8-6B

CHANGES IN EXPECTED NET BURGLARY SENTENCES FOR EXAMPLE CASES: FLORIDA TEST AND COMPARISON SITES

(CASE B)



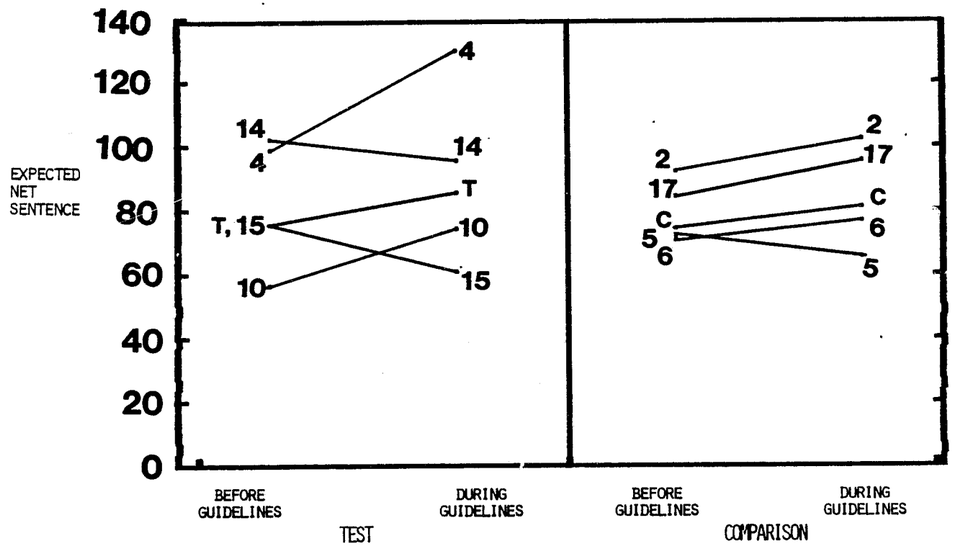
Key:
- 4 = 4th Circuit (Jacksonville)
- 15 = 15th Circuit (W. Palm Beach)
- 10 = 10th Circuit (Bartow)
- 14 = 14th Circuit (Marianna/Panama City)
- T = Test Sites Combined
- 6 = 6th Circuit (Tampa/St. Petersburg)
- 17 = 17th Circuit (Ft. Lauderdale)
- 5 = 5th Circuit (Ocala)
- 2 = 2nd Circuit (Tallahassee)
- C = Comparison Sites Combined

Note: Expected values of net sentence are based on case characteristics specified in Table 8-4 and equations shown in Table 8-3 and Appendix I, Tables I-4 and I-5.

FIGURE 8-6C

CHANGES IN EXPECTED NET BURGLARY SENTENCES FOR EXAMPLE CASES: FLORIDA TEST AND COMPARISON SITES

(CASE C)



Key:
- 4 = 4th Circuit (Jacksonville)
- 15 = 15th Circuit (W. Palm Beach)
- 10 = 10th Circuit (Bartow)
- 14 = 14th Circuit (Marianna/Panama City)
- T = Test Sites Combined
- 6 = 6th Circuit (Tampa/St. Petersburg)
- 17 = 17th Circuit (Ft. Lauderdale)
- 5 = 5th Circuit (Ocala)
- 2 = 2nd Circuit (Tallahassee)
- C = Comparison Sites Combined

Note: Expected values of net sentence are based on case characteristics specified in Table 8-4 and equations shown in Table 8-3 and Appendix I, Tables I-4 and I-5.

sentencing practices previously noted. From what we have documented about Florida guidelines implementation, system adaptation, and compliance elsewhere in this report, these results should not be surprising.

8.3 Sentencing Impacts in Maryland

We turn now to the analysis of sentencing outcomes for actual burglary cases in Maryland. The analysis takes much the same form as that for Florida:

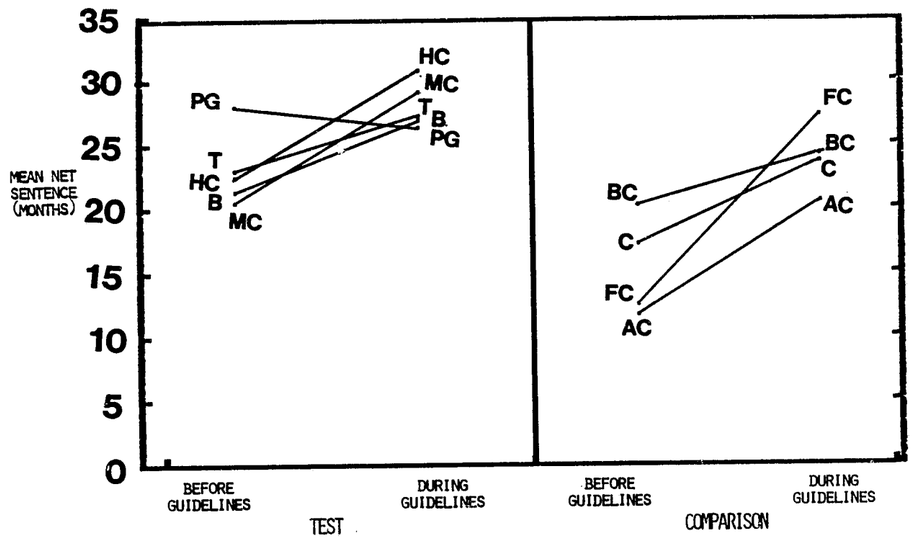
- a description of net burglary sentences (incarcerative time minus suspended time) as rendered, in terms of variability and severity;
- examination of differences in case characteristics among the four test and three comparison sites;
- application of statistical techniques to control for case differences, in order to focus on residual variation in sentencing; and
- discussion of the patterns of residual variation, to assess the effect of the guidelines test on uniformity of sanctions.

The data to be analyzed cover two nine-month periods of court operation, from September 1980 through May 1981 (the period immediately preceding the introduction of the guidelines) and from September 1981 to May 1982 (the last nine months of the test year). Truncation from full-year data was required to adjust the burglary data for guidelines eligibility.[10] Because of possible seasonality in burglaries, the months of June through August were also removed from the data from the pre-guidelines period. By removing data covering the first three months of the guidelines test, guidelines eligibility for burglary cases was assured.[11] It is also likely that some start-up effects in the data were diminished or eliminated.[12]

Figure 8-7 shows the values of the mean net sentence for each individual Maryland site, as well as for all test sites together and all comparison sites together. The data are displayed for the pre-guidelines period and the guidelines test year. Note that while the test sites as a group started out higher in mean net sentence than the comparison sites, the comparison sites appear to have gained considerably between the two time periods. The mean sentence for the pooled comparison site cases increased from 17.1 to 23.6 months, while the pooled test site case sentences went from 23.1 to 27.4 months on average. In addition, the average sentences of the test sites appeared to converge toward each other somewhat.

FIGURE 8-7

CHANGES IN NET SENTENCES FOR BURGLARY:
MARYLAND TEST AND COMPARISON SITES



Key:

B = Baltimore City
PG = Prince George's County
MC = Montgomery County
HC = Harford County
T = Test Sites Combined

BC = Baltimore County
AC = Anne Arundel County
FC = Frederick County
C = Comparison Sites Combined

Data Set: **Maryland burglary data**, 9 month subset.

Base: **All cases with no missing values for any of the offender characteristics, offense characteristics, race, or sex.**

There was substantial variability associated with these average sentences. Overall, for the pooled test and comparison sites, Figure 8-8 indicates that the variability was slightly less during the guidelines year than before. For individual jurisdictions, the pattern was mixed; there were reductions in net sentence variation in the three largest (Baltimore City and Prince George's and Baltimore counties), but increases in the four other, smaller sites.

However, these figures do not directly address the issue of sentence uniformity, since they show the effects of case differences as well as the impacts of inter-judge, interjurisdictional and other differences. Table 8-5 provides a profile of important case characteristics for the two groups of sites before and during the guidelines test. Examining these factors, which are usually taken to be sources of warranted variation in sanctions, can suggest their roles in the patterns just described.

The case characteristics examined here include the offender's race, sex, prior record, and legal status at the time of the offense. Offense factors include the seriousness of the primary charge at conviction, the number of counts of this charge, and whether there were also subsidiary convicted offenses.[13]

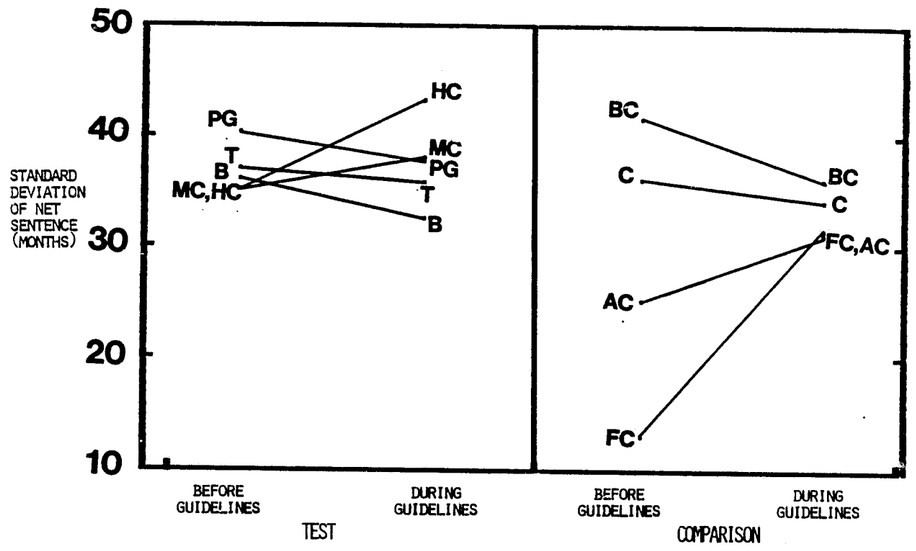
Table 8-5 indicates that there is great similarity between the test and comparison sites in case characteristics. Both groups of sites had a heavy predominance of male offenders. In general, offenders did not tend to have serious prior records. The average seriousness of the burglary was also little different between time periods or groups of sites. Only two distinctions stand out. Racial composition in the test sites was very different from the comparison sites. Similarly, the proportion of offenders on restricted status was much higher in the test than in the comparison sites. Note also that during the year of guidelines use, the proportion of offenders on restricted status more than doubled in the test sites.

The main purpose of examining case characteristics is to explore the possibility that observed changes in sentencing are attributable to changes in the types of burglary cases during the two periods, rather than to the introduction of guidelines. We observed that average sentences apparently increased in both test and comparison sites. The data on case characteristics provide little apparent explanation of these trends. Case characteristics seem rather stable between the time periods of interest. Thus, it is likely that the guidelines were introduced during a period of generally increasing sentence severity, at least for burglary cases.

The key test of guidelines impact in this analysis is whether the residual variation in sentencing--the remaining differences in sanctions once the effects of case differences are removed--was reduced in the test sites relative to the comparison sites. Using the statistical technique of multiple regression, we can control for five case characteristics just described (excluding race and sex) and measure any change in residual variation. Table 8-6 shows the resulting regression equations for the pooled test and comparison sites. The five offender and offenses characteristics

FIGURE 8-8

CHANGES IN NET BURGLARY SENTENCE VARIATION:
MARYLAND TEST AND COMPARISON SITES



Key:

B = Baltimore City
PG = Prince George's County
MC = Montgomery County
HC = Harford County
T = Test Sites Combined
BC = Baltimore County
AC = Anne Arundel County
FC = Frederick County
C = Comparison Sites Combined

Data Set: Maryland burglary data, 9 month subset.

Base: All cases with no missing values for any of the offender characteristics, offense characteristics, race, or sex.

Table 8-5

Case Characteristics: Maryland Burglary Data

Percent of Cases With:	Test Sites		Comparison Sites	
	Pre-Guidelines Period	Guidelines Period	Pre-Guidelines Period	Guidelines Period
Black offender	57.7%	57.6%	22.8%	25.2%
Male offender	97.7	97.6	96.5	97.8
Multiple convicted counts	9.6	12.6	9.8	11.7
Additional (lesser) convicted offenses	18.2	17.2	20.9	19.0
Offender on restricted status	4.7	11.6	8.2	8.8
Average Offense Seriousness[a]	2.01	2.13	2.09	2.10
Seriousness of Highest Prior Adult Conviction[a]	.94	1.01	1.15	.85
Number of Cases	(771)	(778)	(316)	(274)

Data Set: Maryland burglary data, 9-month subset.

Base: All cases with no missing values for any of the offender characteristics, offense characteristics, race, or sex.

Note: Case characteristics for each test and comparison site are shown in Appendix I, Tables I-6 and I-7. Variables are defined in Table I-1.

a. Seriousness measure is explained in Appendix F (Maryland Data Collection).

Table 8-6

Regression Equations of the Sentencing Decision:
Maryland Burglary Data

Independent Variables	Regression Coefficients: Test Sites, Pre-Guidelines Period	Test Sites, Guidelines Period	Comparison Sites, Pre-Guidelines Period	Comparison Sites, Guidelines Period
Multiple convicted counts	26.99**	26.84**	11.12+	21.97**
Additional (lesser) convicted offenses	11.93**	6.80*	12.30**	6.73
Offender on restricted status	27.91**	18.11**	37.34**	30.96**
Offense seriousness	13.22**	11.04**	13.37**	8.37**
Seriousness of highest prior adult conviction	6.94**	11.47**	9.70**	6.75**
Intercept	-16.07	-14.35	-28.69	-6.27
R^2	.174	.267	.244	.1891
s	33.65	30.67	31.60	31.05
Number of cases	771	778	316	274

Data Set: Maryland burglary data, 9-month subset.

Base: All cases with no missing values for any of the offender characteristics, offense characteristics, race, or sex.

Note: Regression equations for each test and comparison site are shown in Appendix I, Tables I-8 and I-9. Dependent variable is net sentence.

Statistical significance:

** $p \leq .01$
* $p \leq .05$
+ $p \leq .10$

("independent variables") were all significantly related to sentencing differences (with the exception of the presence of lesser convicted offenses in the comparison sites during the guidelines period). Note that the way each factor influenced sentencing differed over time in the test sites. Offenders with serious prior convictions received more severe sentences during the test year, while convictions for multiple offenses were treated about the same. Offense seriousness, restricted status, and lesser conviction offenses played a smaller role in sentencing during the guidelines period.

Together, these five case characteristics accounted for 17 to 27 percent of the variation in net sentence. Over time, the five factors explained an increasing proportion of the variation in test site sentences (up from 17 to 27 percent), and a decreasing proportion of comparison site sanctions (down from 24 to 19 percent). In all instances, however, a great deal of sentencing variation remained unexplained.

Conceptually, this unexplained variation is of two kinds: unwarranted and warranted. Sources of unwarranted variation could include race, sex, and (in the terms of the guidelines test) interjudge and interjurisdictional differences. The residual standard deviations ("s") in Table 8-6 include all the unwarranted sentencing variation. But they may also contain some warranted variation, due to case differences legitimately affecting the sanctions but unknown to us. In particular, there are three Maryland guidelines scoring factors (juvenile delinquency, prior conviction for the same offense, and prior adult parole or probation violations) for which we have no data. Since the purpose of guidelines scoring was to make sentences reflect differences in these factors, such differences are warranted or legitimate. But it is not analytically possible, within the limits of the available data, to remove this warranted variation and measure the changes in unwarranted variation only.[14] As a consequence, conclusions must be somewhat tentative.

Nevertheless, it appears that there was a statistically significant reduction in residual variation in the test sites over time but not in the comparison sites. The magnitude of the reduction was modest, about nine percent, but it is reasonable to conclude that it resulted from the introduction of guidelines. Looking at the individual site equations[15] makes it clear that changes in sentencing in Baltimore City are the key to the reduction. Neither the other test sites individually, nor the three of them pooled (omitting Baltimore), show significant reductions in the residual standard deviation. That is, if the Baltimore cases are left out of the equation in Table 8-6, the result shows no significant change in residual sentence variation.

This likely guidelines impact in Baltimore City is a positive finding. If we consider previous evidence about implementation, several factors probably contributed:

- the substantial representation of Baltimore judges on the Advisory Board;

- the leadership of the Chairman of the Advisory Board, a Baltimore judge;

- the extent of the training effort;

- the simplicity of the scoring system, which, combined with greater self-awareness in sentencing, may have facilitated "mental scoring;" and

- the rallying of judicial support for the guidelines in the face of opposition from the Baltimore State's Attorney.

Figure 8-9 is a graphical display of the changes in residual sentence variation from the regression analysis. It shows the reductions for Baltimore City (from 33.2 to 28.5 months) and for the test sites combined (from 33.7 to 30.7 months). No other decreases were statistically significant. The comparison sites as a group showed no change in residual variation, though there were increases for Anne Arundel and Frederick counties. These basic results were maintained even when variables representing the race and sex of the offender were added to the equations. Race and sex accounted for only a very small portion of the residual variation.[16]

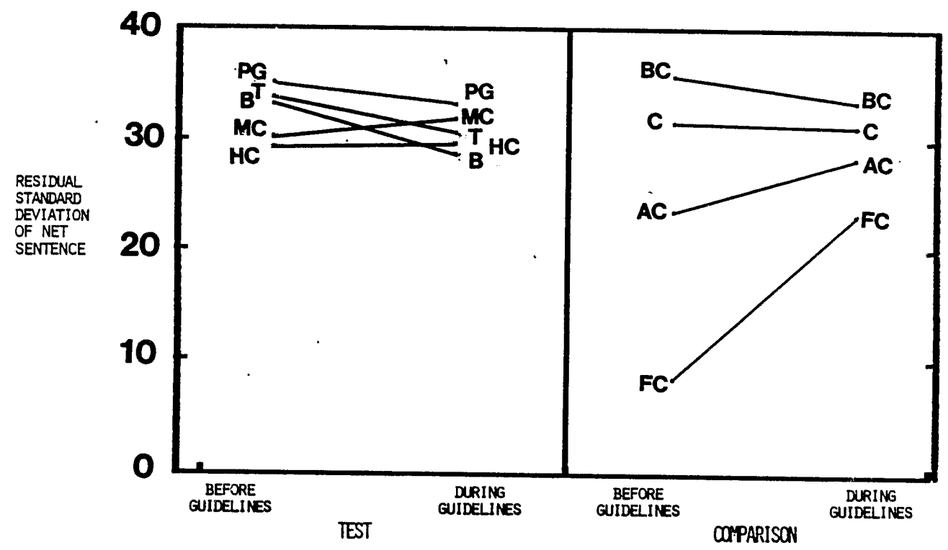
We have examined changes in sentencing uniformity and found increased uniformity in Baltimore City and (as a result) in all four test sites combined. But the regression analysis also provides information about other kinds of changes in sentencing--notably, differences in the weight given to specific offense and offender characteristics and in the severity of sentences for particular cases.

We are interested in changing levels of sentence severity, both in relation to guidelines use and because of potential effects on prison and jail populations. But, when several co-efficients are shifting at once, as they are in the regression models, the overall impact on sentencing levels is not clear. The sentences predicted by the models may increase for cases with some particular sets of characteristics (independent variable values), while for other cases, with different characteristics, they may decrease.

To see what is happening to sentence severity across a range of situations, we again use example cases with a representative range and combination of values for the case characteristics. (Definitions of the cases were shown in Table 8-4.) For each example case, the corresponding independent variable values can be substituted into the regression equation for any site or group of sites to yield an expected level of sentence, given these case characteristics in that site. The results of such a substitution and calculation are displayed graphically in Figure 8-10 (A-C). The least serious case (Case A) shows a mixed pattern: increased sentences in Prince George's, Frederick, and Baltimore counties, the opposite in Montgomery, Anne Arundel, and

FIGURE 8-9

CHANGES IN RESIDUAL VARIATION
IN BURGLARY SENTENCES:
MARYLAND TEST AND COMPARISON SITES



Key:
B = Baltimore City
PG = Prince George's County
MC = Montgomery County
HC = Harford County
T = Test Sites Combined
BC = Baltimore County
AC = Anne Arundel County
FC = Frederick County
C = Comparison Sites Combined

Data Set: Maryland burglary data, 9 month subset.

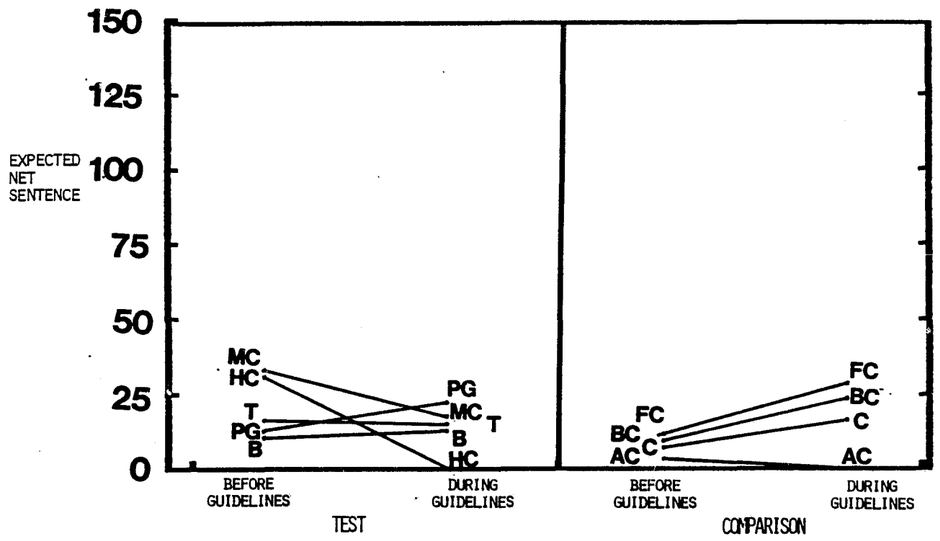
Base: All cases with no missing values for any of the offender characteristics, offense characteristics, race, or sex.

Note: Residual standard deviations are from equations shown in Table 8-6 and Appendix I, Tables I-8 and I-9

FIGURE 8-10A

CHANGES IN EXPECTED NET BURGLARY SENTENCES FOR EXAMPLE CASES: MARYLAND TEST AND COMPARISON SITES

(CASE A)



Key:

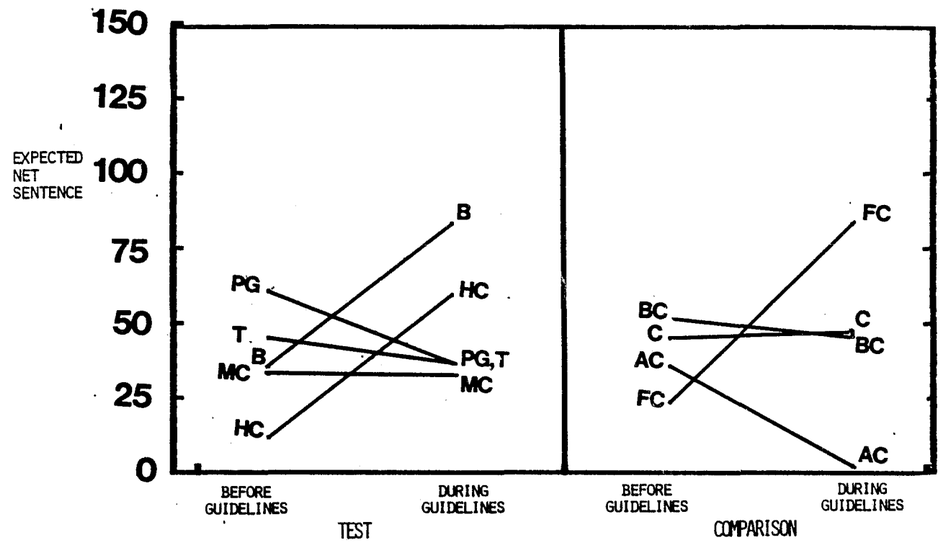
B = Baltimore City
PG = Prince George's County
MC = Montgomery County
HC = Harford County
T = Test Sites Combined
BC = Baltimore County
AC = Anne Arundel County
FC = Frederick County
C = Comparison Sites Combined

Note: Expected values of net sentence are based on case characteristics specified in Table 8-4 and equations shown in Table 8-6 and Appendix I, Tables I-8 and I-9.

FIGURE 8-10B

CHANGES IN EXPECTED NET BURGLARY SENTENCES FOR EXAMPLE CASES: MARYLAND TEST AND COMPARISON SITES

(CASE B)



Key:

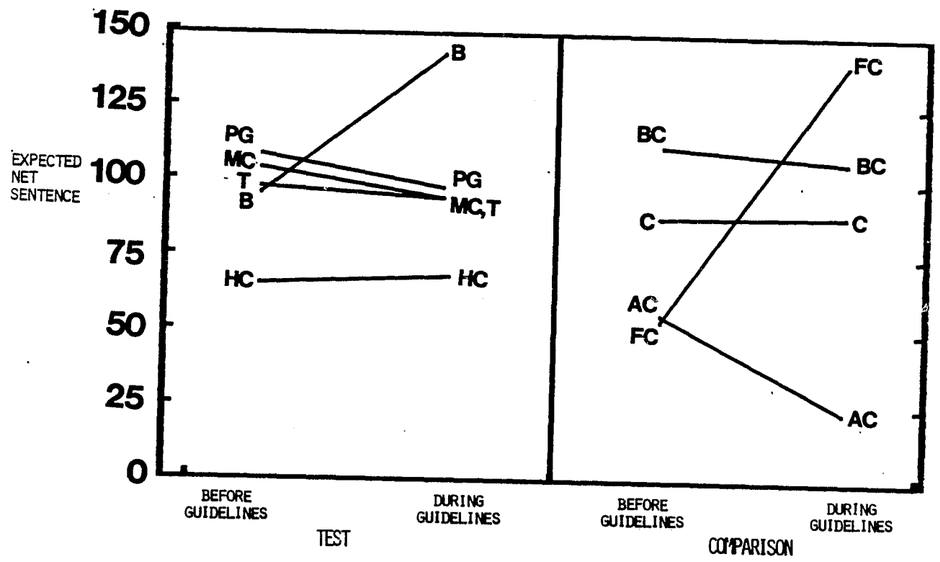
B = Baltimore City
PG = Prince George's County
MC = Montgomery County
HC = Harford County
T = Test Sites Combined
BC = Baltimore County
AC = Anne Arundel County
FC = Frederick County
C = Comparison Sites Combined

Note: Expected values of net sentence are based on case characteristics specified in Table 8-4 and equations shown in Table 8-6 and Appendix I, Tables I-8 and I-9.

FIGURE 8-10C

CHANGES IN EXPECTED NET BURGLARY SENTENCES FOR EXAMPLE CASES: MARYLAND TEST AND COMPARISON SITES

(CASE C)



Key:

B = Baltimore City
PG = Prince George's County
MC = Montgomery County
HC = Harford County
T = Test Sites Combined

BC = Baltimore County
AC = Anne Arundel County
FC = Frederick County
C = Comparison Sites Combined

Note: Expected values of net sentence are based on case characteristics specified in Table 8-4 and equations shown in Table 8-6 and Appendix I, Tables I-8 and I-9.

Harford. The net result is virtually no change for the test sites combined, but a doubled predicted sentence (from 7 to 16 months) for the pooled comparison sites.

The sites contrast more strongly in predicted sentence level for the two more serious cases (B and C). An offender with no prior felony record who committed a moderately serious burglary (e.g., a daytime housebreaking) while on probation would have received a 35-month sentence in Baltimore City before guidelines but an 83-month sentence during the test year; the sanction would also have increased greatly in Harford County (from 12 to 60 months) and Frederick County (from 23 to 84 months). However, severity decreased considerably in Prince George's County and slightly in Baltimore County. Overall, sentences for the two pooled groups of sites changed little and were quite similar. For the most serious case, involving (e.g.) multiple counts of nighttime housebreaking by a paroled burglar with a felony record, the sanctions also increased very sharply in Baltimore City and Frederick. Yet, again, for the pooled test sites and comparison sites, the sanctions were similar and changed little over time.

To sum up, we have seen two kinds of changes in sentencing patterns associated with the period of the multijurisdictional guidelines test in Maryland. It appears that use of the guidelines brought about reduced variability of sentencing in Baltimore City and, as a result, in the four tests sites combined. There was no such reduction in the comparison jurisdictions. At the same time, there continued to be striking differences among jurisdictions, most notably in severity of sanctions. For the same case, not only did sites differ in initial predicted sentence, but the sentences also changed in opposite directions and to varying degrees. Thus, the important question of whether guidelines can reduce variability among jurisdictions has not been answered in Maryland, where the changes were largely within one jurisdiction.

However, the reduction in sentence disparity in a large, urban site offers hope that guidelines can be a viable reform in similar, difficult settings. We have noted the aspects of Maryland's approach that probably contributed to this impact; there was a clear contrast with Florida's approach to judicial representation and training, and a clear contrast in results. While we cannot be certain that these efforts (and not other contrasts in local practices, personnel, and political climate), were the decisive factor, we do believe they made a difference.

FOOTNOTES

1. Section 4.4.1 discusses more broadly the implications of the test experience in guidelines development for sentencing disparity.

2. Supplementary data tables for the impact analysis are presented in Appendix I.

3. Note that this is not the typical use of regression analysis, in which the focus is on explanation (especially the size, sign and significance of the coefficients).

4. The net sentence is calculated as the prison or jail term less suspended time, always in months. A number of alternative forms of the dependent variable were tested, including a scaled version and a binary one (IN/OUT, incarcerative sentence or not). A logarithmic functional form was also tried, to reduce any impact of outlying cases in the regression analysis. There were no appreciable differences from the net sentence results, so the basic variable (with its easily comprehensible units) is used here for presentation.

5. Although the squared standard deviation, or variance, is often used to measure variation, the standard deviation is more easily interpretable, since it is in the same units as the variable itself.

6. Seriousness is measured by the statutory degree defined in the Florida criminal code. However, to convert statutory degree to a numeric variable for purposes of statistical analysis, the most serious felonies (first degree) have been coded as "3," moderate felonies (second degree) as "2," and the least serious felonies (third degree) as "1."

7. Data for individual sites are displayed in Appendix I, Tables I-1 and I-2.

8. Research on race as a source of disparity is briefly reviewed in Chapter 2.

9. Significance was assessed by F-tests calculated on ratios of squared residual standard deviations.

10. Guidelines eligibility was limited to cases with convictions from June 1, 1981 on, while the burglary data collected covered cases sentenced in that period. Examination of the timing relationship between conviction and sentencing (in the scoresheet data) suggested that the overwhelming majority of sentencings took place within three months of convictions.

FOOTNOTES
(continued)

11. Indeed, a substantial number of cases convicted and sentenced in those three months were lost from the analysis.

12. For comparability, the regression analysis was also conducted for Florida using a nine-months observation period. The results with respect to sentencing uniformity in the test sites were unchanged.

13. These are the same seven factors used to describe the Florida burglary cases in Section 8.2. However, a caveat with respect to measurement of two of these characteristics (prior record and seriousness of the primary charge) stems from the nature of the Maryland criminal code. Article 27 is well-known for its incompleteness and lack of structure; the influence of this problem on guidelines development was noted in Chapter 4. To analyze the burglary data, in the absence of a degree structure, and given the continued charging of common law burglaries, it was necessary to develop a seriousness ranking, in order to identify the primary (most serious) offense when several charges were sentenced in one sentencing event, and to assign a seriousness level to the offender's prior adult convictions. The ranking was based on broad groupings of criminal offenses by statutory maximum sentence and by felony/misdemeanor classification (when available). The ranking is shown in Appendix F. The ranking of burglary offenses matches that established by the Advisory Board of the guidelines project for purposes of grid development and prior record scoring. The reader should be aware, however, that the measurement may be imperfect and certainly does not have the same external validity as the Florida measures based directly on that state's criminal code.

14. To assess the impact of omitting these three guidelines factors, regression models were tested on the scoresheet data, which contain all the scoring components as well as the sentence. This analysis of burglary scoresheets produced three very interesting findings. First, prior convictions for the same offense and parole violations were strongly correlated with adult record, so that some of their effect on sentencing is already captured in the adult record's coefficient. Second, juvenile delinquency--which was not correlated with the other scoring factors--made a major difference in explanatory power only in Prince George's County. Third, the equation for Baltimore City was considerably stronger with the scoresheet data using only the same five case characteristics. This proved to be a result of sample bias: the omission of a substantial number of cases, with lower average net sentence, from scoresheet filing, but their inclusion in the burglary data set. Because these cases are present in the burglary data set, the burglary analysis results are more accurate and reliable.

15. See Appendix Tables I-8 and I-9.

FOOTNOTES
(continued)

16. Black offenders received the same sentences as whites, all else equal, in the comparison sites both before and during the test period. In the test sites (where a far higher proportion of offenders were black), blacks received an average sentence six months longer than whites with the same other measured personal and case characteristics. This six-month differential existed before and during the guidelines test. In both groups of sites, male and female offenders were treated the same before the test period, but females were favored (all else equal) during the test period.

CONCLUSIONS

CONCLUSIONS

Taken as a whole, the multijurisdictional test in Florida and Maryland demonstrated that it is possible to develop and implement judicial guidelines. However, the test also gives rise to several cautions about the implementation of sentence reforms in general and the use of judicial, empirical, voluntary guidelines in particular. Both states, for example, experienced pockets of significant resistance which limited (or in the case of the urban site in Florida, obviated) the success of the test. Though increased support building with the judiciary might have alleviated some of this resistance, the expectation that judicial, empirical guidelines would be readily accepted by judges led the projects to under-emphasize such activities. In addition, the focus on judicial participation led to a limited role for defense and prosecution in the development of guidelines and minimal support building activities with these groups. Had defense and prosecution been more involved and aware, it is possible that their resistance might have been minimized.

Using even the most conservative estimates, it appears that judges did not comply with the requirement to consider the guidelines for as many as 15 percent of the eligible burglary cases in one state and 43 percent in the other. The states also fell significantly short of full compliance with the requirement to provide reasons for extra-guidelines sentences--50 percent of the cases were missing reasons in one state, while 20 percent were missing in the second. While some--perhaps many--of these problems were no doubt due to the fact that we examined only the first year of guidelines operation, such difficulties also hint at far more significant and fundamental concerns.

Under the guidelines model implemented in Florida and Maryland, there were no means available to enforce compliance with the guidelines requirements on the part of the judiciary, the defense, or the prosecution. This meant that judges could sentence without even considering the guidelines, and that the defense and prosecution could subvert the guidelines through plea negotiations, or even worse, score bargains. When judges considered the guidelines and decided to impose a sentence outside the recommended range, there was no way to ensure that reasons were provided, that the reasons were valid, or that the sentence remained proportional to the crime. Finally, mistakes in scoring which led to differences in sentence outcome could not be redressed under this system. Given these shortcomings, guidelines cannot offer much greater assurance than other sentencing schemes that sentences will be based on a common standard and that like offenders will be treated similarly.

The fundamental source of these shortcomings does not appear to lie with the guidelines' judicial origins or empirical design, but rather with their voluntary implementation. While certainly not a panacea for all the ills of implementation, it appears that mandatory guidelines would relieve many of the problems noted in the multijurisdictional test. Coupled with a strong monitoring function, such measures as appellate review of sentences and a binding implementation mandate can foster compliance, promote accuracy of use, and provide a mechanism for redress of errors.[1]

However, mandatory guidelines cannot be cast as anything else than the explicit sentencing policy of the state. This, in turn, raises important questions concerning guidelines development. For example, the viability of judicial development and implementation must be seriously questioned. On constitutional grounds, it would appear that mandatory statewide sentencing policy may be established only by the legislature; on practical grounds, it seems unlikely that key groups would acknowledge the legitimacy of mandatory judicial guidelines or that the implementing authority of the judiciary would carry as much force with them as a legislative mandate.

Moreover, if we acknowledge mandatory guidelines as policy, the empirical development of guidelines must also be questioned. Rather than basing new standards on the practices of the past and allowing only the judiciary a determining say in their development, guidelines can and should be developed through a normative, consensus-building process which allows for the input of all parties affected by sentencing policy. It is in this respect that mandatory guidelines offer the greatest advance over previous sentencing reforms. By including all affected groups in a rational policy debate, there is greater opportunity to establish policies which recognize sentencing not as an isolated decision, but as a decision with enormous implications for all components of the criminal justice system.

FOOTNOTES

1. The Minnesota sentencing guidelines are instructive in this regard. See, for example, the Preliminary Report on the Development and Impact of the Minnesota Sentencing Guidelines (St. Paul, MN: Minnesota Sentencing Guidelines Commission, 1982); and Kay A. Knapp, "What sentencing reform in Minnesota has not accomplished," 68 Judicature (1984): 181-189.

END