RUTGERS
THE STATE UNIVERSITY
OF NEW JERSEY

School of Criminal Justice
15 Washington Street
Newark, New Jersey 07102

STUMBLING TOWARD JUSTICE

Some Overlooked Research and Policy Questions About Statewide Sentencing Guidelines

by

Richard F. Sparks
Bridget A. Stecher
Jay S. Albanese
Peggy L. Shelly

with a chapter by Donald M. Barry

School of Criminal Justice
Rutgers University
Newark, N.J.

31 March 1982

(Final report of the Evaluation of Statewide Sentencing Guidelines Project supported by Grant No. 78-NI-AX-0147 awarded to the School of Criminal Justice, Rutgers University, by the National Institute of Justice, U.S. Department of Justice, under the Omnibus Crime Control and Safe Streets Act 1968 (as amended). Points of view and opinions stated in this report are those of the authors, and do not necessarily represent the official views or policies of the U.S. Department of Justice.)

Guidelines We Never Finished Reading

Guidelines for Correcting Non-Compliance with Established Task Standards

In recent years, there has been an effort to achieve implementation of measurable objectives for departments and sections of court operation followed by establishment of measurable task objectives for individual employee roles. In many areas, tasks are being monitored on a consistent basis. The problem arises over a plan of action if non-compliance with basic task standards is a chronic problem on the part of the individual. Upon implementation of corrective action, there should be consideration of:

1. The prioritization of the task relevant to the role and the chronicity, seriousness and gestault of the non-compliance in one or more areas.....

Memo from the Summit County, Ohio, Juvenile Court, reprinted in The New Yorker magazine, 22 March 1982, at page 129.

...what rough beast, its hour come round at last,
Slouches toward Bethlehem to be born?

W. B. Yeats, writing about "just deserts" in
The Second Coming.

From Michael Robartes and the Dancer, Dundrum,
England, 1921 (Cuala Press). First published in
The Dial, November 1920, and The Nation, 6 November 1920.

Contents

	Page
List of Tables	ix
List of Figures	xiii
Preface	xiv
Chapter 1: What This Report is About	1
History of the Research Project	2
What Ever Happened to New Jersey	6
The Chronology of Guidelines Development	9
Overview of the Report	16
Notes to Chapter 1	20
Chapter 2: Concern About Variations in Criminal Sentences: A Cyclical History of Reform	26
Discretion and Disparity	26
Early Empirical Studies of Sentencing	29
Other Pressures for Sentencing Reform	32
Reform of the Italian Penal Code	33
Glueck's "Rational Penal Code"	34
Perspectives of Concern About Sentencing	35
The Cyclical Concern for Sentencing Reform	36
Dissatisfaction with Indeterminacy: Determinacy Revisited	37
Sentencing Institutes	38
Sentencing Councils	39
Appellate Review	41
Disagreement About Disparity	43
Discretion and Its Control	44
Notes to Chapter 2	49
Chapter 3: The Concept of Sentencing Guidelines	52
Origins of the Concept	53
Vulgarization of the Concept of Guidelines	60
Guidelines as a Species of Sanctioning Rule	72
The Supposed Empirical Basis of Sentencing Guidelines	82
Description vs. Prescription, Reemphasized	87
The Decision to Incarcerate	88
Widths of Prescribed "Normal" Ranges	91
Controlling Compliance with Guidelines	94
Conclusions about the Concept of Guidelines	96
Notes to Chapter 3	99

Chapter 4: The New Jersey Sentencing Guidelines 111

Description of the New Jersey Guidelines 112
Reconstructing the New Jersey Guidelines Data 122
Description of the Reconstructed New Jersey Sentencing Guidelines Data 125
Offense of Conviction 125
Criminal History Score 125
Amenability to Non-Custodial Supervision 126
Exacerbating Factors 126
Community Background 131
Actions Since Arrest 132
Institution of Confinement 132
Median Sentence in Months 135
Do Configurations Matter? 135
What About Cell Scores? 138
Simplifying Cell Score 140
Combining Offense Categories 151
Conclusions and Implications 154
The Construction of the Guidelines 154
The Content of the Guidelines 157
The Use of the Guidelines 158

Notes to Chapter 4 162

Chapter 5: Prisoners' Perceptions of Sentencing in New Jersey 165

Method 166
Characteristics of the 1979 Sample 167
Questionnaire Design 168
Findings 173
Analysis of Sentencing Perceptions 173
Analysis of Seriousness and Severity Perceptions 178
Analysis of Guidelines Perceptions 188
Discussion 196

Notes to Chapter 5 198

Chapter 6: Sentencing in Massachusetts: The Climate for Sentencing Guidelines 203

Origins of the Massachusetts Guidelines 204
The Legal Structure of Sentencing in Massachusetts 213
Court Jurisdiction and Offense Classification 214
Minimum and Maximum Penalties 217
Appellate Review of Sentence 217
Parole in Massachusetts 218
Institution of Confinement 219
Parole Eligibility for Special Offenses 221
An Overview of Sentencing and Parole: Sentencing Guidelines in a Climate of Reform 222

Notes to Chapter 6 239

Chapter 7: Developing and Implementing the Massachusetts Sentencing Guidelines 256

Development of the Massachusetts Sentencing Guidelines 257
Defining and Collecting the Data 258
Making Sentencing Policy in Guidelines Research 269
Structure of the Massachusetts Sentencing Guidelines 278
The Seriousness of the Current Offense 280
Dangerous Weapon Use 282
Degree of Injury 283
The Seriousness of Prior Offenses 285
Computing the Guidelines Sentence 286
Deviating from the Guidelines 290
Implementing the Massachusetts Sentencing Guidelines 291
Judicial Reception of Sentencing Guidelines 293
The Probation Department and Sentencing Guidelines 300
Using Sentencing Guidelines: The Reactions of Judges, Prosecutors, Probation, and Defense 302

Notes to Chapter 7 314

Chapter 8: The Massachusetts Sentencing Guidelines: A Secondary Analysis 334

A Secondary Analysis of the Massachusetts Sentencing Data 334
Demographic and Background Information 335
Juvenile Criminal History Information 338
Criminal Justice Process and Offense Information 338
Current and Prior Offense Information 341
Information on Injury to the Victim 346
Weapon Use Information 347
Judges' Sentence Decisions and Prosecutors' Recommendations 348
The Decision to Incarcerate 352
Length of Jail and Prison Terms 355
Replicating the Massachusetts Sentencing Guidelines 362
Examining Outliers 380
Examining the Guidelines Ranges 388
Constructing New Massachusetts Sentencing Guidelines 392
Examining Correlations Among Items 394
Finding the Best "Fit" 397
Examining the Guidelines Ranges 404
Summary and Conclusions 408

Notes to Chapter 8 411

Chapter 9: Other Examples of Statewide Sentencing Guidelines 413

The Minnesota Sentencing Guidelines 413
The Pennsylvania Guidelines 439
Sentencing Research and Guidelines in Michigan 458
Summary and Conclusion 489

Notes to Chapter 9 495

Chapter 10: An Empirical Comparison of Guidelines Structures in Two States 512

By Donald M. Barry and Timothy L. Kennedy

Rationale 512
Method of Offense Translation: Current Offense 516
Method of Criminal History Index Construction 516
The In-Out Decision 519
Distributions of Time Incarcerated 523
Sentence Characteristics According to Conviction Offense Type 525
Correlates of Grossly Discrepant Sentences 528
Projected Prison Populations Under Alternative Sentencing Schemes 532
Summary 535

Appendix 10A: The Determination of Offense Equivalence Classes in Minnesota and Massachusetts 537

Chapter 11: How to Evaluate Statewide Sentencing Guidelines 550

Sanctioning Rules Revisited 550
Empirical Research and Sentencing Guidelines 558
Guidelines in their Social and Institutional Context 562
How to Measure Compliance with Guidelines 564
Concluding Unscientific Postscript 573

Notes to Chapter 11 574

Bibliography 577

List of Tables

		Page
3.1	Average Total Time in Months Served Before Release (including jail time)	55
4.1	Overview of the New Jersey Sentencing Guidelines: Attribute Inclusion by Offenses	115
4.2	Criminal History Factor Score	127
4.3	Amenability to Non-Custodial Supervision Factor Score	128
4.4	Exacerbating Factor Score	130
4.5	Community Background Factor Score	133
4.6	Actions Since Arrest Factor Score	134
4.7	Rank Correlation Coefficients (Kendall's τ) Between Orderings of Configurations for Seven Offense Types in the New Jersey Sentencing Guidelines	139
4.8	Regression of Percent of Cases Incarcerated on Cell Score	141
4.9	New Jersey Sentencing Guidelines Correlation of Offender Attribute Scales with the Incarceration Decision	143
4.10	Pearson Correlations Between Factors: Total Sample (N=10,629)	144
4.11	Stepwise Regression of Original and Modified Cell Scores on the Incarceration Decision (N=2,118)	146
4.12	MCR Analysis of Original Cell Score	148
4.13	MCR Analysis of Modified Cell Score No. 1	148
4.14	MCR Analysis of Modified Cell Score No. 2	149
4.15	MCR Analysis of Modified Cell Score No. 3	149
4.16	MCR Analysis of Modified Cell Score No. 4	150
4.17	MCR Analysis of Modified Cell Score No. 5	150
4.18	Regression of the Percent of Cases Incarcerated for Grouped Offenses on Cell Score (N=10,629)	153

4.19	Percent of Cases Incarcerated by the Cell Score for Each Offense (N=10,629)	155
4.20	Percent of Cases Incarcerated by the Criminal History Score for Offense Groups (N=10,629)	156
5.1	Percentage of Responses Indicating Dissatisfaction with Sentencing in New Jersey	175
5.2	Percentage of Responses Indicating Dissatisfaction with Own Sentence	177
5.3	Medians and SAM Values for Offense Descriptions	181
5.4	Medians and SAM Values for Sanctions	184
6.1	Massachusetts Parole Eligibility - Sentence Conversion Rules	220
6.2	Massachusetts Criminal Justice System Staff Interviews in Four Counties: 1979 and 1980	225
6.3	Massachusetts Interviews with Judicial, Legislative and Correctional Personnel: 1979 and 1980	225
7.1	Massachusetts Sentencing Guidelines Sample Selection	263
7.2	County Case Representation in Original and Final Massachusetts Sentencing Guidelines Sample	266
8.1	Sentence Decision by District Attorney Recommendation	351
8.2	Incarceration Decision by County	354
8.3	Time Incarcerated Until Parole, by Type of Trial (Incarcerated Cases only)	359
8.4	Massachusetts Sentencing Guidelines Variable Weights and Predictive Ability	365
8.5	Regression of Massachusetts Sentencing Guidelines Factors on Sentence Length for Two Incarcerated Offender Samples	372
8.6	Regression Analyses of the Massachusetts Sentencing Guidelines Ability to Explain Variation in the Incarceration Decision for Selected Cases	376
8.7	Regression Analyses of the Massachusetts Sentencing Guidelines Ability to Explain Variation in Sentence Lengths for Selected Cases	379

8.8	Characteristics of Cases with Excessive Residual Values	383
8.9	Cases Not Actually Incarcerated, for whom the Guidelines Prescribe Incarceration (N=248)	387
8.10	Distribution of Massachusetts Actual Sentences Within Ranges of the Massachusetts Guidelines Sentence, All Cases (N=1,440)	390
8.11	Percentages of Sentences Above and Below the Massachusetts Guidelines Sentence by Range for All Cases (N=1,440)	393
8.12	Rutgers Secondary Guidelines Analysis of the Decision to Incarcerate (N-1,440)	399
8.13	Rutgers Secondary Guidelines Analysis of the Length of Sentence for Incarcerated Cases, Excluding 47 Outliers, and Life Sentence Cases (N=721)	402
8.14	Distributions of Massachusetts Actual Sentences Within Ranges of the Rutgers Revised Version of the Massachusetts Guidelines Sentence for Incarcerated Cases, Excluding 47 Outliers, and Life Sentences (N=721)	406
8.15	Percentages of Sentences Above and Below the Rutgers Revised Version of the Massachusetts Guidelines Sentence by Range for Incarcerated Cases, Excluding 47 Outliers, and Life Sentences (N=721)	407
9.1	The Minnesota Sentencing Guidelines Matrix	418
9.2	Percentages Imprisoned in Minnesota Court Disposition Data, Classified by Guideline Categories	423
9.3	Mean Times to First Parole in Minnesota Prisoner Sample, Classified by Guideline Categories	426
9.4	Additive Model Fitted to Mid-ranges of Minnesota Matrix (Effects and Residuals in Months)	429
9.5	Multiplicative Model Fitted to Mid-ranges of Minnesota Matrix (Effects and Residuals in Natural Logarithms)	432
9.6	Multiplicative Model Fitted to Mid-ranges of Minnesota Matrix (Effects and Residuals in Months)	433

9.7	"Inside-Out" Plot of Non-Zero Residuals from Table 9.6	436
9.8	Original Pennsylvania Sentencing Guidelines	445
9.9	Revised Pennsylvania Sentencing Guidelines	448
9.10	Multiplicative Model Fitted to Mid-ranges of Pennsylvania Matrix (Effects and Residuals in Months)	453
9.11	Statistically Significant Variables in the "In-Out" Regression Equation for Sex Crimes in Michigan Sentencing Research	464
9.12	Cross-Classification of Predicted Versus Actual Sentences for Sex Crimes in Michigan Sentencing Research	465
9.13	Statistically Significant Variables in the "Length" Regression for Sex Crimes in Michigan Sentencing Research	467
9.14	"Empirical Sentencing Matrix" for Sex Crimes, Michigan Sentencing Research	475
9.15	Comparison of Guidelines Within Each Cell by Statutory Maximum for Sex Crime Groups, Michigan Sentencing Guidelines 1980	483
10.1	The In-Out Decision: Minnesota Guidelines Vs. Massachusetts Guidelines, Version 1	520
10.2	The In-Out Decision: Minnesota Guidelines Vs. Massachusetts Guidelines, Version 2	520
10.3	The In-Out Decision: Minnesota Guidelines Vs. Massachusetts Actual Disposition	521
10.4	The In-Out Decision: Massachusetts Guidelines Version 1 Vs. Massachusetts Guidelines Version 2	521
10.5	The In-Out Decision: Massachusetts Guidelines Version 1 Vs. Massachusetts Actual Disposition	522
10.6	The In-Out Decision: Massachusetts Guidelines Version 2 Vs. Massachusetts Actual Disposition	522
10.7	Sentence Characteristics by Sentencing Scheme and UCR Offense Type	526
10.8	Comparison of Characteristics in the Two Subsamples Exhibiting the Greatest Discrepancies Between Minnesota and Massachusetts Guidelines Sentences	530

List of Figures

		Page
3.1	Hypothetical Frequency Distribution of Lengths of Prison Terms	92
4.1	New Jersey Sentencing Guidelines Breaking and Entering Score Derivation	114
4.2	48 Cell Locator Sheet	117
4.3	48 Cell Guidelines Matrix: Breaking and Entering or Entering	118
4.4	New Jersey Sentencing Guidelines: An Overview for Breaking and Entering: 1 Cell Configuration (Cell Number 7)	120
4.5	New Jersey Sentencing Guidelines: An Overview for Breaking and Entering: 1 Cell Score	123
5.1	Disparity Between Perceived Fair and Going Sentences by Seriousness of the Offense	187
5.2	Prisoners' Perceptions of Guidelines Sentences for Robbery Case	190
8.1	Hierarchical Clustering Scheme Based on Distributions of Sentence Severities in Massachusetts Counties	361
8.2	Predicted Massachusetts Guidelines Sentences Against Residual Values, Excluding Life Sentence Cases (N=768)	369
8.3	Predicted Massachusetts Guidelines Sentences Against Residual Values, Excluding 47 Outliers, and Life Sentence Cases (N=721)	374
10.1	Distributions of Expected Lengths of Incarceration	524
10.2	Projected Prison Population Rates Under Four Sentencing Schemes	534
11.1	Hypothetical Frequency Distributions of Lengths of Prison Terms Before and After Guidelines	566
11.2	Hypothetical Probability Distributions for "In" and "Out" Cases Sentenced After Guidelines	570

Preface

The research described in this report has some attributes which would, in our opinion, have delighted that great lover of logical contradictions and paradoxes, the Reverend Charles Lutwidge Dodgson. That fact itself is something of a paradox, since it seems to us highly likely that the Rev. Dodgson would have been bored to tears, if not to death, by the subject-matter of our research -- unless, of course, he had used his inimitable skill, during a convenient few minutes between brillig and dinner, to turn that subject-matter into something like a Lobster Quadrille or a tale of WASP in a Wig (i.e. a British High Court judge). That apart, the principal paradoxes of our work are as follows. Our research was done both too early, and too late; too expensively, and yet too much on the cheap; too hastily, and yet (inevitably) much too slowly; it was too local in focus, yet too grandiose in scope. The ultimate paradox is this. The principal object of our research is a mere piece of legal technology, which is by definition trivial; yet that bit of technology just happens to have appeared at a time when one of the noblest of human enterprises -- namely Doing Justice -- was in the course of coming back into fashion, for reasons which have precious little to do with either our research or its object; and yet the bit of technology is, as it turns out, well-suited to the noble enterprise! It is rather as if some pimply technocrat had emerged from the bowels of a Swiss pharmaceutical factory, just in time to persuade Socrates that Valium was chemically more efficient than hemlock.

Anyone acquainted with social research will know (though they may not admit it) that it is much more fun to criticize others' research than to do one's own. That is by no means the only reason, however, why this report contains a fair amount of criticism of what other researchers have done. It is, we think, unfortunately true that a great deal of the research that has been done in recent years, with the avowed aim of developing statewide sentencing guidelines, has been very badly done. In the greater scheme of things, this might not ordinarily matter much (after all, who <u>does</u> care who killed Roger Ackroyd?). In this case, however, the misperformances of social researchers <u>are</u> important, because they may -- if not neutralized -- impede the efforts of those who wish to make the sentencing of convicted offenders more rational and fair, and who (rightly) see sentencing guidelines as a way of accomplishing this important goal. Indeed, <u>some</u> of the blunders, made by <u>some</u> of those who have done research aimed at constructing "empirically-based" sentencing guidelines, are potentially even more dangerous: they may lead to sentencing policies that are even less fair, and less rational,

than those now found in most American jurisdictions. We have no evidence that this has yet happened. But if it has not, this is almost certainly just a matter of luck; and where the stakes are high -- as they are, where sentencing policy is concerned -- it is better not to push one's luck too far.

Since "social research" has apparently become anathema in some quarters lately, and since some of our criticisms are aimed at what are pretty clearly examples of "social research", we ought perhaps to emphasize that in our opinion there are many ways in which such research can be of great help to those who wish to eliminate discrimination and disparity in sentencing by introducing sentencing guidelines. We discuss some of these ways in this report. Our primary concern, in fact, is not so much to point out what has been done badly by some researchers in the past, but to suggest some things that can and should be done by researchers in the future -- provided that those things are done well.

This report is the product of work done by far more people than its title page would suggest; and it reflects the cooperation, criticism and support of a still larger number of persons, to whom the authors are deeply indebted. To begin with those listed as authors: it is possible to assign responsibility for initial drafts of particular chapters to particular persons (1, 3, 9 and 11 to Sparks, 4, 6, 7 and 8 to Stecher, 2 to Albanese, 5 to Shelly, 10 to Barry). The final versions, however, are in a very strong sense the product of a joint enterprise; each of us has benefitted from the comments and criticisms of the others, in a great variety of ways that we can no longer identify. In its present form, the report reflects a pleasant (if occasionally hectic) collaboration, over the three years or so in which our research was done.

We are also deeply grateful to many other colleagues who helped us in many ways during our research. Our special thanks must go to Donald Barry, who, in addition to writing Chapter 10 of the report (with the assistance of Timothy Kennedy), served as a member of the project's Advisory Board, and carried out some of the interviews with prison inmates which are reported in Chapter 5. The other members of our Advisory Board were Don M. Gottfredson and Andrew von Hirsch; we are grateful to them, and to the special consultants -- Professors Alfred Blumstein, Herbert Solomon and Todd Clear, and Judge Morris Lasker -- who joined them in reviewing our work, and who provided valuable insights and guidance. For assistance in the interviewing of prison inmates, we should also like to thank Lela Keels, Deborah Koster, Aaron Lewitt, Julia Mueller and Fred Roth; and Sally Manning, for assistance in coding these often exasperating data. Todd Clear and Julia Mueller also deserve thanks for their help with the interviews with judges, prosecutors, defense counsel and others, which we carried out in Massachusetts in 1980 and 1981 (and which are discussed in Chapters 6, 7 and 8). We should also like to thank Alex Greer and Kathleen Hanrahan for their invaluable assistance in the collection and analysis of legal materials relating to sentencing in general, and guidelines in particular.

Many others, not formally associated with our project, generously provided help without which our work could not have been done. Among those involved in developing sentencing guidelines in various states, our special thanks go to Michael Hutner, Amy Craddock and Thomas Marx, in Massachusetts; to Jan Smaby, Dale Parent and Kay Knapp in Minnesota; to Marvin Zalman and his associates in Michigan; and to John Kramer in Pennsylvania. Among administrators in the New Jersey correctional system, we should especially like to acknowledge the help which Michael Power and Howard Beyer gave us, in facilitating our interviews with inmates at Rahway State Prison; and, of course, we are grateful to the inmates themselves for giving us the benefit of their views. It would be invidious in the extreme for us to single out, by name, any of the large number of judges, prosecutors, defense counsel and probation officers in Massachusetts, for the time which they spent, and the information which they gave us, during our two periods of fieldwork there in 1979 and 1980. We must, however, make one exception: we could not have done the work which we did, without the support and help of Judge John J. Ronan, who -- as chairman of the Superior Court's Committee on Probation and Parole -- was largely responsible for overseeing that state's sentencing guidelines. If our report is occasionally less than enthusiastic about some aspects of those guidelines, this is no reflection on Judge Ronan and his colleagues; on the contrary, without their candor and generosity, we would all know much less than we do about the complex business of sentencing reform. (In this respect, the support and guidance which we received in Massachusetts contrasts sharply with what we received in another state somewhat closer to home: see Chapter 1.) The patience of our project monitor, Mr. Jay Merrill, of the National Institute of Justice is gratefully acknowledged.

Following a curious tradition which seems prescribed for the authors of reports like this one, we save our greatest debts until last: these are owed to Carol Kenney and Dorothy Webster, our project secretaries. If we had to be grateful to <u>them</u>, we would be helpless.

In conclusion -- and not by way of acknowledgement, but as a guide to the litigious: though this report is, as noted earlier, the product of a close collaboration between its four authors, sole responsibility for the remarks on pp. 6-9 and pp. 60-72 must be imputed to the senior author. Rank has its privileges; and tenure has its obligations.

Chapter 1: What This Report Is About

> Judges have within their capabilities today the means by which they may sharply curtail, if not virtually eradicate, sentencing disparities in most American jurisdictions.
>
> (Wilkins et al., 1976, p. xi)

In 1974, a group of researchers headed by Professors Don M. Gottfredson and Leslie T. Wilkins began research designed to show the feasibility of developing and implementing empirically-derived guidelines, as a means of structuring judicial discretion and eliminating disparity in sentencing. Gottfredson and Wilkins had previously developed such guidelines for use in connection with parole decision-making; and the U.S. Parole Commission had been using various versions of the Gottfredson-Wilkins guidelines since 1972 (see Gottfredson et al., 1975; Gottfredson, Wilkins and Hoffman, 1978). Extension of the concept of guidelines to judicial decision-making seemed natural; and by the end of their feasibility study, Gottfredson and Wilkins were convinced that this could and should be done. The opening words of their report are quoted above.

Through further research, sentencing guidelines were subsequently developed and implemented in a number of county-level jurisdictions, including the Denver District Court, the Cook County (Chicago) Circuit Court, the Essex County and Superior Courts (Newark, New Jersey), the Maricopa County (Phoenix) Superior Court, and the Philadelphia Court of Common Pleas. In addition, in a number of states research was undertaken, beginning in the late-middle 1970's with a view to developing statewide sentencing guidelines. While not all of these guidelines followed what may be called the Gottfredson-Wilkins model, the majority of them have done so.

The basic concept of the Gottfredson-Wilkins guidelines model is as follows. Decision-makers in the criminal justice system (e.g. parole boards, or judges) are given information about the patterns of decision-making in their jurisdictions in the <u>past</u>; they then use this information to guide their decisions in the <u>future</u>. In the case of parole decision-making, for example, the information might consist of a range of months or years to be served in prison before release; this range is derived from an analysis of time served before release by different sorts of prisoners in the past. In using the guidelines, the parole board may release a prisoner after a term falling within the stipulated range,

without any further justification. Alternatively, the board may "depart from" the guidelines -- setting a term which falls outside the suggested range -- if there are special factors which make this appropriate, though they must state their reasons for such a departure. Similarly, sentencing guidelines can provide a range of years or months to be served by offenders who are to be incarcerated; a sentence within that "normal" range would need no further justification, though reasons would have to be given for any sentence outside that range.

The research described in this report is concerned with the evaluation of statewide sentencing guidelines. How generally has this concept been accepted, in the years since it was first suggested? How have sentencing guidelines been constructed, in those jurisdictions in which they have been developed to date? What do the sentencing guidelines developed in different states look like? Are they similar in form and content, and do they exemplify the concept originally propounded by Gottfredson, Wilkins and their colleagues? More generally, what is distinctive about this technique of controlling judicial discretion, and how do guidelines differ from "presumptive" sentencing laws (such as California's Uniform Determinate Sentencing Law of 1976)? Are guidelines a good thing, or a bad thing? Do they entail any particular philosophy of punishment, such as deterrence or "just deserts" (von Hirsch, 1976)? Still more generally, how should one evaluate sentencing guidelines? What should be the criteria for determining whether they work well or badly, and what procedures should be followed in assessing sentencing guidelines that are now in use, or which may be developed in the future?

It is questions like these with which the present report is concerned. A general overview of the report is presented later in this chapter. Before that, however, we describe in some detail the history of our project, since a knowledge of that history is necessary in order to understand the way in which the project's objectives were originally formulated and why those objectives had to be changed; and why some of the objectives finally decided upon could not, in the event, be accomplished.

History of the Research Project

Our project initiated as a result of a solicitation early in May 1978, from what is now known as the National Institute of Justice (NIJ),[1] for proposals to carry out an evaluation of the New Jersey statewide sentencing guidelines. Research aimed at developing sentencing guidelines had been conducted, under the auspices of the New Jersey Administrative Office of

the Courts, since 1975; and it was expected that the guidelines themselves would be implemented some time in 1979. New Jersey's would then be the first statewide guidelines in use. Our preliminary proposal (dated 30 May 1978) was accepted by NIJ, and it was expected that we would begin our evaluation of the New Jersey guidelines in October 1978. Our project was to have had three main components. First, we wanted to document the process of designing and implementing guidelines in New Jersey, describing any problems which arose in the course of construction and implementation, and the steps taken to overcome those problems. This narrative part of our project was to have been based in part on documentary sources, but also on pre-guidelines and post-guidelines interviews with judges, prosecutors, defense counsel, court administrators and others, in four selected counties within the state. We also intended to carry out systematic observations of sentencing before and after the introduction of guidelines, in those four counties. Second, we intended to conduct a large number of detailed statistical analyses of sentencing practice in New Jersey, and to make a comparison of pre-guidelines and post-guidelines sentencing in the four selected counties and (so far as possible) in the state as a whole. Third, we intended to carry out a number of smaller studies of other topics, including the effects of the guidelines on case processing and prison populations in New Jersey, and on offenders' perceptions of the fairness and justice of their sentences before and after the guidelines were in use. [2]

Since it was known at the time of our proposal that a number of other states were considering the introduction of sentencing guidelines, we thought that the New Jersey experience could have important implications for those states, and that an analysis of the construction, implementation and use of the New Jersey guidelines could provide valuable information to those states in which other forms of "determinate" or "presumptive" sentencing were under consideration. [3] There are, after all, certain problems inherent in the notion of statewide (as opposed to county-level) sentencing guidelines. Even in a small state like New Jersey, it seemed likely that we would find considerable variation between counties, not only in attitudes to crime and punishment but in crime patterns, organization of the criminal justice system and so on. How would a single set of statewide sentencing guidelines cope with these variations? It was to examine questions of that kind that the NIJ solicitation requested (and we proposed) to look in depth and detail at the experience of using guidelines in four counties, which would be chosen to reflect differences in crime patterns and the other things just mentioned. [4] Neither we nor (it seems) NIJ had any a priori views as to whether or not there

should be variations in sentencing between different counties (e.g. between urban and rural ones, or between those in the industrial north of New Jersey and the largely rural southern part of the state). The very idea of developing statewide sentencing guidelines, however, seems to imply that there should not be such differences. How would this work out, in practice? It was that empirical question, and not the question of policy, with which we were concerned. Even if it were to be accepted that there should be a single set of guidelines used throughout the state, it remains true that the criminal justice system in most American states is largely administered at a county or city level. What effect would this have, on the implementation of statewide sentencing guidelines?

In addition, we had a number of broader theoretical concerns which we hoped that our three-pronged research strategy would address. For example, we were interested in the conceptual foundations of the sentencing guidelines model, as developed in New Jersey and elsewhere. What did those advocating guidelines seek to accomplish? What factors should or should not be included in a set of guidelines -- and why? Do different methods of constructing guidelines make a difference? How might guidelines affect, or be affected by, new legislation or other techniques for controlling judicial discretion? What are the implications of attempting to control decision-making at one point in the criminal justice system -- to wit, the courts -- and how is this likely to affect other aspects of the criminal justice process? Do guidelines seem to be a useful technique for reducing disparity in sentencing, and can they help to provide for the development of a clear and consistent sentencing policy? We proposed to try to answer these questions by studying New Jersey's guidelines; but our choice -- more precisely, NIJ's choice -- of New Jersey was dictated solely by the fact that that state seemed likely to be the first to implement statewide sentencing guidelines. Though the solicitation, and thus our proposal, referred to an "evaluation" of the New Jersey sentencing guidelines, this term was clearly not intended in the cold-blooded utilitarian sense in which it is often used in the field of social research; we were not aiming to prepare a report card that would either castigate or commend the New Jersey guidelines project or those involved in it. It is necessary to emphasize this point since, as we shall see, the term "evaluation" appears to have been interpreted in a somewhat different sense from that in which we (and NIJ) intended it, by those responsible for the New Jersey guidelines.

At any rate, we were informed that our original proposal had been successful, in a letter dated 1 August 1978. At

about the same time, however, the terms of the original solicitation were reconsidered by NIJ. Though it was clear that New Jersey appeared to be closer than any other state to implementing statewide sentencing guidelines, it was evident that there was considerable interest and activity in a number of other states -- perhaps more than had been appreciated at the time of the original solicitation. Something was known about efforts underway in some of these states; but little if anything was known about others. It thus seemed reasonable to begin research on statewide sentencing guidelines by conducting a national survey which would assess the status of efforts in different states, so as to enable NIJ to plan its overall evaluation strategy as well as to provide information on guidelines development to interested parties throughout the country.[5] (Such a survey, it was thought, could also provide NIJ with information necessary for planning, not only with respect to "evaluation" -- whatever that might be taken to mean -- but also with respect to requests for technical assistance or funding for guidelines development or monitoring.[6]

Accordingly, in a series of discussions with NIJ during the month of August 1978, the project envisaged in our original proposal was modified in certain respects. As a result, we proposed to carry out a national survey which would assess the extent of interest and activity in developing statewide sentencing guidelines throughout the country. From this overview, a small number of states would be selected for a more detailed description of current developments: initially it was assumed that there would be about seven states in which this more detailed assessment would be worthwhile, though, as will be seen in a later chapter of this report, the number eventually proved to be smaller than that. A more detailed description and analysis of sentencing guidelines in one state would still be carried out; and at the time of our revised proposal, we still assumed that that state would be New Jersey. This more detailed study would make use of pre- and post-guidelines interviews and systematic observation, as we had originally proposed; and it would also involve "before and after" statistical analyses of sentencing practices in a small number of selected counties, as well as (to the extent possible) in the state as a whole. Finally, the smaller sub-studies envisaged in our original proposal would be carried out, in the state chosen for detailed study. Our modified proposal, therefore, assumed a national focus, rather than being concerned with just one state; in recognition of this, the title of the project was changed from "An Evaluation of the New Jersey Sentencing Guidelines" to "The Evaluation of Statewide Sentencing Guidelines."

What Ever Happened to New Jersey?

We began our research, under the terms of our modified proposal, on 4 October 1978. Shortly after this time, however, we had to modify our research design in an important respect, since it became impossible for us to use New Jersey as the site of our intensive "case study". This was so, for two reasons. The first is that the New Jersey sentencing guidelines were "implemented" -- in the sense of being given to judges and other persons in the criminal justice system in New Jersey -- on 23 October 1978. As noted in the preceding section, we had considered it important to obtain a detailed picture of sentencing before the introduction of guidelines in our "case study" state, in order to be able to assess the impact of those guidelines, not only on patterns of sentencing but on plea negotiations, charging practices, collection of information on offenders, and the like. Some data on these and other matters could no doubt have been obtained through retrospective interviews. But for a variety of reasons, this is far from being a satisfactory research technique (see Sparks, 1981b, for a discussion of the problems of retrospective interviewing); and this unexpectedly early implementation thus dealt a severe blow to our intended "before and after" design. When we submitted our original proposal, and also at the time of our revised proposal, it seemed likely that we would have between six and nine months to carry out pre-guidelines interviews and systematic observations in the four selected counties in New Jersey, and to interview other persons involved in the development of the guidelines, and in the criminal justice system in general, throughout the state. In the event, only three weeks elapsed between our grant award, and the sudden announcement of the introduction of the New Jersey guidelines. As may be imagined, this sudden implementation was a considerable surprise to us.

The second, and far more important, reason why our case study could not be based in New Jersey is that we were denied access to the data set from which the New Jersey guidelines were constructed. In addition, the Acting Administrative Director of the Courts at the time (the Hon. Arthur J. Simpson) made it clear that he would not give his support to our plans to interview judges in the state about their use of the guidelines (or anything else). The reasons for this denial remain unclear to us. To begin with, we had never had any reason to suppose that we would be summarily denied these research facilities. One of us (Sparks) had been involved as a consultant to the New Jersey guidelines project two years previously; cordial relations had been maintained between our project staff and the New Jersey project director (John P. McCarthy, Esq.) since that time; and the original solicitation

from NIJ had been for an evaluation of the New Jersey guidelines and not of statewide guidelines in general.[7] Yet in a letter dated 3 November 1978, we were informed by Judge Simpson that we would not be given the facilities needed to carry out our original research design. Simpson's letter stated that

> ...we all believe it would be best if you save the evaluation of New Jersey until toward the end of your work. You might proceed with the other aspects and by the time you get to New Jersey we will have meaningful results following our recent implementation. We also expect to have additional results, if we can obtain the requisite continuation funding to proceed in the many other areas of work that we perceive as critically necessary. Your support of our efforts will be appreciated, and you are assured of our cooperation in connection with your evaluations."[8]

It seemed clear to us that this response rested on a misconception of our intentions, and an evidently skimmed reading of our research design. (So far as we are aware, Judge Simpson had by that time been sent a copy of our revised proposal). Any "evaluation" of New Jersey's guidelines would necessarily have had to wait for some time, as we stated in our proposed research design, in order that we could obtain sufficient post-guidelines data; in any case, as explained earlier, our intention was not to evaluate in a report-card sense, but to explore quite general problems inherent in the concept of statewide sentencing guidelines. These would probably arise first in New Jersey, because that state was first in the field; but they could well be expected to appear in other states as well. In short, we had no wish to undertake anything which might have been regarded as a premature evaluation of the New Jersey guidelines.

Nor is it easy to see how our request for a copy of the data from which the New Jersey guidelines were developed could have led to such a premature evaluation. This data set -- consisting of over 15,000 cases, with over 800 variables per case -- was and is a potentially rich resource, not only for our project but for researchers interested in sentencing generally. It would have given us a very complete and detailed picture of sentencing in New Jersey before the implementation of the guidelines; it would also have enabled us to explore a number of conceptual and statistical problems surrounding the development of guidelines, in a way that other data then available to us would not. At most, however, the New Jersey construction data would have provided us with a "baseline" against which sentencing patterns in New Jersey after the implementation of guidelines could be compared. The

data by themselves would not have made possible an evaluation of the impact of the New Jersey guidelines -- though they might have shown whether or not guidelines were really needed in that state, to control sentencing disparity.

In an effort to clarify the reasons behind our request, and to explain more fully the nature of our project, senior staff of our project met with Simpson and McCarthy early in January 1979.[9] At this meeting we stressed again the importance of our having access to adequate data on sentencing in New Jersey before the implementation of guidelines, so that we could study the guidelines' effects. In addition, since questions of confidentiality and privacy had been raised by Simpson as a reason for not giving us their data, we went to some lengths to explain the procedures for the protection of human subjects in research which bound not only ourselves but the School of Criminal Justice and Rutgers University generally. Finally, we offered any help which we could give to Simpson and his colleagues, by way of technical assistance in connection with other research which they might wish to do in the future.[10]

None of this was to any avail. At one point, we were informed by Simpson that he had requested an opinion from the Attorney General of New Jersey, as to the release of information from the data files, "in view of the federal and state laws as to security and privacy".[11] If such a request was made, we were never informed of its outcome.

Subsequently, late in 1979, Judge Simpson was replaced as Administrator of the New Jersey Courts, by Mr. Robert Lipscher. Though it was by this time far too late for us to use New Jersey as the "case study" state in our research, we made several further efforts to obtain a copy of the New Jersey guidelines construction data, in order to carry out some comparative statistical analyses of the kind described in later chapters (particularly chapter 10) of this report. These request were also unsuccessful; we were again denied access to the data, allegedly on the ground that "confidentiality and privacy" could not be protected if the data were released to us. It is difficult indeed for us to resist the conclusion that the confidentiality and privacy involved were those of judges in New Jersey -- not of convicted offenders.

However this may be, it had become plain to us by about March 1979 that we would not be able to carry out our detailed case study in New Jersey, as we had originally planned. Accordingly, we began contacting those involved in developing statewide sentencing guidelines in other states, with a view to seeing which of these states might provide the most

suitable location for our intensive case study. The timing of developmental efforts in those states left us with only two choices: Massachusetts and Minnesota. In each state, a project aimed at developing statewide guidelines was underway; in neither state would the guidelines be implemented before we could carry out a study of pre-guidelines sentencing, in accordance with our original design; in both states, the directors of the guidelines projects (Dr. Michael Hutner in Massachusetts, and Dr. Dale Parent in Minnesota) had already given us generous amounts of information about their work, and had made helpful offers of support for our project should we decide to locate our case study in their state. Our final choice of Massachusetts as a primary case study site was based in large part on budgetary considerations; our design called for spending several extended periods of time interviewing and observing in the chosen state, and we would not have been able to conduct such detailed field work if we had chosen Minnesota. (We did, however, make site visits not only to Minnesota but to Michigan, where an empirical study of sentencing was being carried out under the direction of Dr. Marvin Zalman. The Minnesota guidelines, and those recently implemented on a trial basis in Michigan, furnish an illuminating contrast to the Massachusetts guidelines; they are discussed in Chapter 9 of this report.)

In its final form, therefore, our research involved (1) a national survey of the state of development of statewide sentencing guidelines; (2) a detailed case study of sentencing and sentencing guidelines in Massachusetts; (3) less detailed studies of guidelines developed and/or implemented in three other states (Minnesota, Michigan and Pennsylvania); (4) sub-studies of inmates' attitudes to sentencing, and (5) a simulation exercise involving data obtained from Minnesota and Massachusetts, and those two states' guidelines.

The Chronology of Guidelines Development

Both in our original proposal to study just the New Jersey sentencing guidelines, and in our modified proposal -- with its national focus coupled with a case study (based in New Jersey or elsewhere) -- we had a number of topics in view. The first of these, both logically and chronologically, concerned the conceptualization and construction of sentencing guidelines. What did the guidelines model adopted in any particular jurisdiction seek to accomplish -- and how did the various models of sentencing guidelines differ? How were the guidelines developed -- and what effect might the method of construction be expected to have on sentencing practices? Second, we were interested in what we called the "interactive effects" of the guidelines on the rest of the criminal justice system. For example, what effect would the introduction of

sentencing guidelines have on the processes of charging and plea negotiation? What might be the impact on prison populations -- and how might this in turn affect the ways in which the guidelines were used (or misused) by judges in practice? Finally, we were interested in what may be called impact outcomes of sentencing guidelines. Would sentencing guidelines actually result in a reduction in judicial discretion, disparity and/or variation in judicial sentencing? If so, how would this come about? Would sentencing guidelines provide for the development of a clear and consistent sentencing policy? Would they be a useful tool for training the judiciary in sentencing? Would they reduce offenders' perceived dissatisfaction with the sentences imposed on them?

It is plain that the last few of these topics require a comparison of sentencing, charging, etc., before the introduction of sentencing guidelines, with those practices after the guidelines come into effect. The conceptualization of a state's sentencing guidelines could to some extent be studied retrospectively; the construction of those guidelines could be investigated through concurrent observation, as well as by secondary data analysis; the process of implementaion could be studied by watching it take place. We had hoped, however, to find out something about whether or not sentencing guidelines made a difference, in the real world; for this purpose a "before and after" comparison was clearly necessary.

It is important to emphasize that we had never expected that we would be able, within the time constraints imposed by the original NIJ solicitation, to make a final assessment (if there is such a thing) of the impacts of sentencing guidelines, even in our chosen case-study jurisdiction. The development and implementation of statewide sentencing guidelines necessarily involves much more than the statewide distribution of a couple of pieces of paper. The trial, conviction and sentencing of offenders takes place in the context of a fairly complex social system; and it is by now a commonplace of "systems analysis" that one cannot tinker with one bit of such a system, without risking further consequences for other parts of the system, which in turn may have further consequences for still other parts of the system, which in turn.... (For illuminating discussions of these complex inter-systemic effects, see Heumann and Loftin, 1979; Clear, DeIlio and Lubitz, 1979.)

To give but one example: suppose that sentencing guidelines are introduced in a particular state or other jurisdiction; and that all judges in that jurisdiction are thoroughly and immediately briefed on this fact; and that all of them -- exhibiting, perhaps some rather un-judicial docility -- begin immediately to try sincerely to use the

guidelines in the way their developers intended. (Even this bit of institutional change is likely to take some time -- even if all of the judges are very fast learners, and the guidelines are of kindergarten simplicity.) Prosecutors, defense counsel, probation officers -- and even offenders -- may in time come to learn about the guidelines; as a result they may well modify their behavior in a variety of ways, which in turn may affect case flow, institutional populations, charging and bargaining relationships -- perhaps even crime patterns. These things, in turn, may affect the kinds of cases appearing before the courts, which in turn may lead to changes in the way in which judges use the guidelines; and so it goes. It is not necessary to postulate that the consequences of introducing sentencing guidelines are literally endless; in time, no doubt, such perturbed social systems usually "settle down" -- either to a new rhythm and tempo, or to something like the one they had before. The point is that such a "working-out" of change in a complex social system does take time -- and that a definitive assessment of the impact of such a change cannot be made until a reasonable amount of time has elapsed.[12]

Even if everything had gone according to plan, the terms of the original NIJ solicitation did not provide for enough time to make such a definitive assessment. Even if (as we originally thought likely) the New Jersey guidelines had been introduced around the middle of 1979, we would not have been able to obtain more than about a year's worth of post-guidelines data, from which we could only have estimated the most immediate and very short-run effects of that state's sentencing guidelines. In the real worlds of criminal justice systems, the amount of time typically required for the mere diffusion of information about change, let alone the consequent adaptation of different actors in the system to that change, would clearly have been much longer than our initial two-year grant period would have allowed. We expected to be able to say something about the short-run impact of sentencing guidelines in at least one state; but we had not expected to be able to say much about the more variable long-term system reaction to guidelines.

Even after our original proposal was revised, however, and after it became clear that New Jersey would not serve as the site of our intensive case study, we still had reason to believe that we would have the opportunity to obtain at least some post-guidelines data in some state -- not only statistical data on sentencing patterns after the guidelines were introduced, but also interview and observational data on the sentencing process itself. We thus expected that, within the time allocated for our original grant, we would be able substantially to accomplish the major objectives envisaged in our revised proposal.

In the event, this was not to be. In Massachusetts -- which we eventually chose, for reasons explained in the preceding section, as the site of our intensive case study -- the sentencing guidelines developed under the sponsorship of the Superior Court were presented to the judges in November 1979; but they were only implemented on a voluntary (and apparently rather small-scale) basis, until April of 1980. At the time, they were revised somewhat; and a further trial period, in which the guidelines still had an essentially voluntary character (i.e. they were not prescribed by a Rule of the Superior Court, or anything analogous) began. By the end of our initial grant period -- 30 September 1980 -- we could only have obtained data on post-guidelines sentencing practices covering a period of three months or so;[13] and that would clearly have been far too short a period of time to permit us to say anything at all about even the short-run impact of the Massachusetts guidelines.[14] (We were however able to obtain some qualitative data on the perceived effects of those guidelines on the Massachusetts criminal justice system as we discuss in more detail in Chapter 7.) The situation in Minnesota proved to be similar. The guidelines developed by that state's sentencing commission were presented to the state legislature on 1 January 1980; barring legislative revision or veto of the guidelines, they were to take effect on 1 May 1980. Again, given the exigencies which inevitably surround the collection and analysis of empirical social data, we would not have been able to study effectively more than about three months of the post-guidelines experience in Minnesota. That is not enough time in which to say anything of interest, let alone importance, about complex institutional change.[15]

We must emphasize that our recounting of these chronological facts is in no way intended as a criticism of those responsible for the hard work of developing and implementing sentencing guidelines, in either Massachusetts or Minnesota. On the contrary: as we try to make clear in our discussion of those efforts in later chapters of this report,[16] in each state an impressive amount of empirical research and analysis was carried out, under severe and not always anticipated time constraints, and no doubt with less by way of resources than those responsible for the projects would ideally have liked. The fact remains that -- after the unanticipated refusal of the New Jersey authorities to cooperate with our research -- there were only two states, namely Massachusetts and Minnesota, in which we could possibly have carried out a project even remotely resembling the one envisaged in our original proposal; and in both of those states, the tempos with which sentencing guidelines were developed and implemented did not permit us to collect sufficient post-guidelines data to make an adequate assessment even of the short-term impact of the guidelines.

This problem of developmental tempo -- which turned out to be andante or even largo, rather than vivace or even allegro --became apparent to us early in 1980. It was discussed at a meeting with our project's advisory committee[17] and special consultants[18] at the end of February 1980, at which we outlined what we then saw as the need for continued research on the evaluation of statewide sentencing guidelines, after the expiration of our original grant. As a result of these discussions, a proposal for the funding of a second phase of our project was submitted by us to NIJ in April 1980. At that time there was still only one state (New Jersey) in which statewide sentencing guidelines had in any sense been implemented; the Massachusetts guidelines were being tried out on an experimental basis; and Minnesota's guidelines would not come into effect for another month. Neither the effects of guidelines on sentencing behavior, nor the adaptations of those states' criminal justice systems to the guidelines, could possibly have been established in the six months then remaining on our original grant.

It seemed clear to us that, even at the end of the further two years, many questions about sentencing guidelines would still remain to be answered. Nonetheless, we felt that at the end of that time we would have a much better grasp of the variety of sentencing guidelines schemes being implemented throughout the country; more importantly, we expected that in at least two states (Massachusetts and Minnesota) we would have had time to observe the impact of sentencing guidelines on actual decision-making, after the transitional phase which necessarily accompanies any change or social policy or action. We therefore proposed to NIJ that our research be continued until the end of October 1982, at approximately the same level of effort that had characterized our first two years. During that second two-year period, we intended to concentrate on three specific objectives. First, we proposed to conduct further detailed studies of the implementation and impact of guidelines in Massachusetts and Minnesota; as we make clear in later chapters,[19] those two states provide many interesting and (we think) important theoretical contrasts, which deserve further study. In addition, we proposed to continue our national survey of the development and implementation of statewide sentencing guidelines.[20] Finally, we proposed to carry out some experimental research on the use made by judges and others of guidelines in making sentencing decisions.

The last of these was admittedly a new objective; but it is one which, we think, follows naturally from the essentially descriptive research which we carried out in the first phase of our project, and which we describe in later chapters of this report. Once a set of sentencing guidelines has "settled

down" within the institutional structure of a given jurisdiction's criminal justice system, there are many questions about judicial and prosecutorial behavior that can and should be studied, not only to assess the impact of guidelines in that jurisdiction, but to clarify more general theoretical questions about decision-making. How (for example) do judges decide what sentence within a stipulated guidelines range to impose? How do they decide whether or not a particular case is "typical" of its kind, and so should receive a sentence within the guidelines range, rather than being the subject of a "departure" and a sentence outside that range? To what extent are judges constrained (or, at any rate, likely to feel constrained) by guidelines of different kinds? We expected to obtain some information on these topics through observation of, and interviews with, judges and other criminal justice system personnel in Massachusetts and Minnesota. But -- given the differences in form and content of those states' guidelines, as well as differences in the organization of their criminal justice systems, crime rates, and so on, we also felt that experimental evidence -- using decision-making "games" of the kind used in a similar context by Wilkins and Chandler (1965), and Gottfredson, Wilkins and Hoffman (1978) -- would be useful.[21]

In a subsequent addendum (dated 21 July 1980) to our phase two concept paper, we proposed something of a shift of emphasis between the three objectives just mentioned: this shift would have involved rather greater concentration on the Massachusetts-Minnesota comparison and the experimental studies of the use of guidelines, and correspondingly less emphasis on the continuation of the national survey.[22] In particular, the Massachusetts-Minnesota comparison seemed -- and still seems -- to us to have considerable importance; in many ways, the two states, and their efforts at guidelines development, can be seen as the endpoints of a continuum, or of several continua. We return to this comparison in later chapters.[23]

We firmly believe that the continuation which we proposed to NIJ would have enabled us to build on the research carried out in our project's first two years and described in this report, so as to accomplish a far greater contribution to knowledge than we were able (in our opinion) to provide in our first two years. As matters stand, there are many questions which we have not had the opportunity to address -- let alone to try to answer. Good social research -- whether or not it concerns the evaluation (in some sense) of social change -- builds incrementally, at a greater-than-linear rate; and there is always a danger that short-run projects, once discontinued, will simply run into the sand.[24]

However that may be, our proposed phase two research was not funded by NIJ.[25] Necessarily, therefore, we were unable fully to accomplish all of the objectives which we should like to have accomplished, and which were set out in our original proposal and its addenda. Many questions about the evaluation of statewide sentencing guidelines in general -- and about the guidelines developed in Massachusetts, Minnesota, and the other states described in this report -- remain to be addressed. We believe, however, that we have accomplished many of the tasks which we initially set ourselves, and have taken significant steps toward accomplishing several others. Specifically, we have been able to describe the development and implementation of sentencing guidelines in two states (Massachusetts and Minnesota); as noted earlier, we believe that a comparison of these two states' experiences is instructive and important for other states which may be contemplating the introduction of this method of controlling judicial discretion in sentencing. In addition, we have carried out a number of secondary analyses of the guidelines developed -- not only in Massachusetts and Minnesota, but also in New Jersey, Pennsylvania and Michigan -- using methods which seem to us to be of general applicability, and to provide a basis for comparing the structures of guidelines which may be developed in other states. These analyses, we believe, take us well on the way to the specification of a general model for the evaluation of statewide sentencing guidelines -- which was one of the principal objectives stipulated in the original NIJ solicitation for the evaluation of the New Jersey guidelines. We have, moreover, been able to complete the two sub-studies (concerned with prisoners' perceptions of fairness in sentencing, and the simulation of the effects of guidelines given different offender populations) which we originally proposed to do.

It may seem paradoxical to claim that we are able to specify a "general model" for the evaluation of statewide sentencing guidelines, given that we have not been able to observe the impact of any statewide guidelines system for a sufficient period of time. But this paradox is more apparent than real. What the expression "general model" in this context means is, roughly, a set of questions, and the procedures by which those questions may be answered, to make it possible to characterize, accurately and fully, what happens if a set of statewide sentencing guidelines is developed and implemented. In order to formulate those questions, and set out those procedures, it is not necessary actually to have studied the long-term (or even the short-term) working of some number of existing systems of statewide sentencing guidelines. What is necessary is to be sensitive to the variety of conceptual and methodological problems raised by this technique of controlling

decision-makers' discretion; and to have a realistic understanding of what is involved in trying to deal with those problems through empirical research. We feel that our work over the past two years or so has provided us with at least some of that sensitivity, and a little of that understanding. The reader of this report who survives to the end of its concluding chapter may of course take a different view.

Overview of the Report

The plan of the remainder of this report is as follows. In the next chapter we present a historical review of legal and criminological concern, in both common-law and Continental jurisdictions, with the problem of controlling judicial discretion in sentencing. Over the past fifteen years or so, there has been a radical shift, in the United States, in prevailing views as to the proper objectives of sentencing of offenders; there has been a drastic falling-off of belief in what Allen (1964) called the "rehabilitative ideal", and a concomitant reassertion of importance of the notion of "just deserts" as a controlling principle in sentencing.[26] This shift -- which has had profound implications, not only for sentencing but for parole (von Hirsch and Hanrahan, 1979) and the organization of correctional systems -- has coincided, by and large, with an increase in empirical research on sentencing and other forms of decision-making in the criminal justice system;[27] and this increase in turn, can reasonably be said to have provided much of the impetus to the development of guidelines for parole and sentencing.[28]

It is a serious mistake, however, to identify the concept of sentencing guidelines -- in the sense with which this report is concerned -- with this shift from "rehabilitation" to "just deserts", or from "treatment" to "punishment". In reality, sentencing guidelines are only one technique -- as it happens, the most recently developed technique -- for attempting to control individual decision-makers' discretion in the choice of sentences; sentencing guidelines are not necessarily associated, intrinsically, with either a "treatment" or a "just deserts" rationale of sentencing. Guidelines are one method of achieving what has lately come to be called "determinacy" in sentencing; but the notion of "determinacy" can logically be applied with equal force to "treatment", incapacitation, "just deserts", deterrence, and many other commonly mentioned objectives of criminal sentencing.

As we show in Chapter 2 of this report, concern about the control of judicial discretion in sentencing, and the elimination of disparity, goes back at least a century; it antedates the shift which took place in most western

countries, in the late nineteenth century, away from a retributive or deterrent philosophy of punishment toward the "rehabilitative ideal". In order to understand what is distinctive about the modern notion of guidelines for sentencing or parole, it is necessary to contrast that notion with other techniques for attempting to control discretion (such as "presumptive" or mandatory sentencing, or legislative or judicial sentencing codes).

After presenting this historical background, in Chapter 3 of this report we analyze in some detail the concept of decision-making guidelines originally propounded in the early 1970's, by Professors Don M. Gottfredson and Leslie T. Wilkins. For a number of reasons, these two scholars' work (and that of their colleagues) has had a considerable impact on sentencing reform in the United States in the past decade. As we shall see, however, several of the attempts which have thus far been made to develop and implement sentencing guidelines have departed, in important respects, from the original Gottfredson-Wilkins concept. Moreover, that concept itself is, in some respects, vague. We need to understand these points of vagueness, in order to evaluate those efforts which have thus far been made to develop and implement guidelines which have claimed to follow the original Gottfredson-Wilkins model.

As we noted in an earlier section of this chapter, we were unable to carry out the research contemplated in our original grant proposal, owing to the New Jersey authorities' refusal to grant us access to their construction data set, or to facilitate interviews with judges and other actors in that state's criminal justice system. We did, however, manage to obtain a copy of the sentencing guidelines that were introduced in New Jersey late in 1978; and Chapter 4 of this report presents an analysis of those guidelines. A principal result of this analysis is that the New Jersey sentencing guidelines could be very much simplified in form, without sacrificing any important information which judges might wish to use in sentencing offenders. As they were originally introduced to the New Jersey judiciary, that state's guidelines took the form of a huge contingency table containing almost 2,000 cells, each one of which was intended to give information about prior sentencing practice for a particular combination of offense and offender type. We show, however, that substantially the same information could be presented by means of a table with no more than five rows and three columns, or fifteen cells in all.

Chapter 5 of the report describes the first of our two sub-studies, which was concerned with prisoners' perceptions of the fairness of their sentences, and with the relations

between perceived seriousness of crimes and perceived severity of punishments. The research on which this chapter is based was carried out at Rahway State Prison in New Jersey, in 1979 and 1980. Ideally we should have liked to conduct this research in the same state in which our intensive case study of guidelines development and implementation was done; for the timing reasons already explained, however, this proved to be impossible. Nonetheless, we believe that the findings of this chapter throw light on a number of problems, both methodological and substantive, relating to efforts to measure offenders' perceptions of the sentencing process and to the ability of sentencing guidelines -- the New Jersey sentencing guidelines in particular -- to alter those perceptions.

The next section of the report -- Chapters 6 through 8 -- describes the research which we conducted on the development and implementation of sentencing guidelines in Massachusetts. Chapter 6 describes the system of sentencing and parole in the Superior Court in Massachusetts, before the introduction of the guidelines. A historical overview of sentencing reform efforts in that state follows this description of the system, and particular attention is devoted to how the concept of sentencing guidelines came to be the most recent sentencing reform effort. Finally, the expected reception of sentencing guidelines by personnel involved in the sentencing process in Massachusetts is discussed. A large part of this chapter is based primarily on fieldwork (including not only observations of sentencing, but also interviews with judges, probation officers, prosecutors and defense counsel) in four counties in the state, in the summer of 1979.

Chapter 7 describes the process by which the Massachusetts guidelines were developed, and the steps that were taken (in late 1979, and early 1980) to introduce them into the sentencing system of the Massachusetts Superior Court. The actual structure of the Massachusetts guidelines is also analyzed in some detail, as are the reactions of judges, prosecutors, defense counsel, and others to that structure.

Chapter 8 -- which in some respects parallels our earlier analysis of the New Jersey guidelines -- is based in part on a reanalysis of the case-level data collected by the Massachusetts guidelines project. Also included in this chapter -- just before our secondary analysis of the data -- is an overview of the characteristics of the original Massachusetts case sample; this material is present in this report because it has not been previously released to the public by the Massachusetts sentencing guidelines project. The results of this reanalysis, we believe, have important implications for the evaluation of sentencing guidelines generally.

In Chapter 9, we give a brief overview of the statewide sentencing guidelines developed and/or implemented in three other states (Minnesota, Pennsylvania and Michigan). In a number of important respects, these guidelines contrast with those developed in both New Jersey and Michigan. They illustrate a number of important structural features of sentencing guidelines, which in some respects build on, and in other respects depart from, the concept originally propounded by Gottfredson and Wilkins. As is the case with both New Jersey and Minnesota, we have (for reasons explained above) no data on the impact of the guidelines on sentencing practice in any of these three states; at the time of this writing, only the Minnesota guidelines have been officially implemented, on other than an experimental basis.[29] We believe, however, that our analyses of the structures of these states' guidelines throws some light on very general problems raised by the concept of sentencing guidelines, which need to be taken into account in any full-scale evaluation of this technique for controlling discretion in sentencing.

Chapter 10 of the report presents the results of our other sub-study, which is based on some statistical analyses of the case-level data collected, for the purpose of developing sentencing guidelines, in Massachusetts and Minnesota. The guidelines developed in these two states differ, in a number of respects; that is, they prescribe somewhat different sentences for certain types of offense and offender combinations. What might be the results, if a population of offenders like those convicted in Massachusetts were sentenced in accordance with the Minnesota guidelines and vice versa? Part of the point of this chapter is to illustrate a distinction which we have come to believe is fundamental for an understanding of the working of techniques for controlling discretion in sentencing. This is the distinction between (a) the amount of <u>permissible</u> variation in sentencing or prison term-fixing, within a structure which prescribes some degree of constraint on individual decision-makers' behavior; and (b) <u>actual</u> variation in sentences or prison terms, given not only a particular structure but also the patterns of cases dealt with by the courts.

In conclusion, Chapter 11 of the report attempts to draw together the findings from the various lines of research described in earlier chapters, and to take some tentative steps toward outlining a general model or program for evaluating statewide sentencing guidelines.

Notes to Chapter 1

[1] This agency of the Department of Justice was of course known, at the time of our original proposal, as the National Institute of Law Enforcement and Criminal Justice (NILECJ); its title was officially changed in 1979. For the sake of simplicity we shall refer to it throughout this report by its present title and abbreviation.

[2] The topics to be investigated in these smaller studies, and their associated research designs, were proposed by us. However, offenders' perceptions of fairness and justice -- along with court administration and functioning, sentencing disparity, length of incarceration and implications for other states, which were to be considered in our main evaluation project -- were specifically mentioned in the original NIJ solicitation (our copy is dated 11 May 1978; see page 2).

[3] The states specifically mentioned (in a National Institute memorandum from W. Jay Merrill to Blair Ewing, dated 2 August 1978) were Alaska, Florida, Massachusetts, Michigan, Minnesota, Oregon and Washington. We present analyses of guidelines in two of these states (Minnesota and Michigan), together with a discussion of the statewide sentencing guidelines developed in Pennsylvania, in Chapter 9 of this report.

[4] The original solicitation actually referred to an unspecified number of "selected New Jersey counties;" our choice of four such counties was to some extent arbitrary, and was determined primarily by our estimate of budgetary constraints. The optimum number of units with which to study such inter-county variation is, of course an empirical question, and is not likely to be the same from one state to the next. In retrospect, however, we feel that we could reasonably have studied at least two more of Massachusetts' rather heterogeneous 14 counties, had our resources permitted. See below, Chapter 6, for a further discussion.

[5] These objectives were mentioned in the memorandum from Merrill to Ewing (cited in note 3 above). In addition, it was suggested that certain more limited evaluation (e.g. of certain elements, processes or stages in the development of guidelines) might be useful; it was also suggested that a longer-term evaluation effort analogous to the Institute's Research Agreements Program might prove valuable. These and other future evaluation strategies are considered at some length in our concluding chapter.

[6] These objectives were discussed in a meeting between Dean Don M. Gottfredson of the School of Criminal Justice, and NIJ personnel, on 7 August 1978; in the absence of the Principal Investigator during the summer months of 1978, Dr. Gottfredson assisted in the discussions on the final objectives and design of our project. We should like to acknowledge our indebtedness to Dr. Gottfredson for his assistance in this and other matters, not only in this period of re-negotiation of our original proposal but throughout our research.

[7] In fact, the solicitation referred to "evaluating the LEAA supported program, Statewide Sentencing Guidelines for the State of New Jersey" (emphasis added). According to information which we subsequently obtained from the New Jersey State Law Enforcement Planning Agency, the total grant made to the New Jersey Guidelines project over several funding cycles was $473,922, of which $426,530 -- or 90 percent -- came from LEAA. We were never able to ascertain whether this federal funding would have entitled us, as a matter of law (e.g. the Freedom of Information Act) to obtain the New Jersey data; the legal position is apparently complicated, and our grant budget had no heading for "litigation against recalcitrant state authorities". However, it may be thought -- and it appears that NIJ did think -- that the substantial federal involvement in funding the New Jersey guidelines development effort should have placed a strong moral obligation on the New Jersey Administrative Office of the Courts to make their data available not only to us, but to other researchers who might have used those data, not to criticize some bureaucrats in New Jersey, but to benefit criminal justice systems throughout the country. So much for moral obligation.

It is not clear just when the New Jersey authorities -- in particular, Judge Simpson -- decided that they would refuse to cooperate with our (and others') research. However, a summary by Gottfredson of a telephone call from Merrill on 3 August 1978, which was primarily concerned with the agenda for the meeting which took place at NIJ on 7 August, refers to a letter which was to be forthcoming from Ewing to Gottfredson; the memo also states that "Not mentioned in the letter is some question they have about New Jersey posture toward the evaluation." It may thus be that the New Jersey authorities had given some indication, at a fairly early date, of their unwillingness to be evaluated.

[8] This letter was addressed to Sparks and to Dr. William Rich of the National Center for State Courts, which had received a grant from NIJ to carry out a study of county-level guidelines (see Rich, et al., 1980, and Chapter 3 below); at the date of this writing, neither the National

Center, nor we, nor anybody else, has been given access to the New Jersey data. Judge Simpson's use of the word "cooperation" must surely be the most Pickwickian use on record. For the analyses of the New Jersey guidelines which we were able to carry out, see Chapter 4.

[9] At this meeting (which took place on 4 January 1979) Simpson complained that there had been a "lack of communication" between LEAA and their project; in addition to a concern about "premature" evaluation, Simpson also raised as objections to our request for cooperation the time demands which he thought our project would place upon him and McCarthy; a fear that our interviewing of judges would "scare" them and thus jeopardize the success of the guidelines; concern about racial discrimination in sentencing in New Jersey, which might lead to prison riots unless some plan were developed to deal with it; and a reluctance to disclose publicly information about variations in sentencing. At one point Simpson suggested that we try to amend the terms of our grant, and come to work for the Administrative Office of the Courts, before doing our "evaluation". We declined this offer. (Field notes of meeting, 4 January 1979.)

[10] This offer had been made on several previous occasions, and was subsequently repeated. We see no inconsistency between the offering of "technical assistance" (in the form of statistical consultancy, help with data analyses, dissemination of information about other projects, etc.) and the kind of impartiality required to conduct a scientific study of an innovative action program like that represented by the New Jersey guidelines. In fact, the New Jersey AOC subsequently drew on the technical expertise of other Rutgers University faculty, in connection with a study published in 1979 which purported to show that there was not, in the pre-guidelines data, any evidence of racial prejudice in sentencing in New Jersey. The methodology of this study, and the language of its conclusions, are discussed elsewhere (see Sparks, 1981a).

[11] Letter from Simpson to Sparks dated 2 March 1979.

[12] What is "reasonable" here cannot, of course, be stipulated a priori. For a general discussion of this problem in evaluation research in the social sciences, see Weiss, 1972; Weiss, 1970; and Rossi et al., 1979; and for a general discussion of systemic inter-relationships (e.g. between judges and probation officers) which may influence information in criminal justice flow, see, e.g., Emerson (1969).

[13] This is, of course, an optimistic estimate in view of the time usually required for coding, data cleaning,

analysis, etc., in a project such as ours. Moreover, even this period would probably not have allowed us to estimate the likely impact of the guidelines on such things as the flow of case processing, e.g. through changes in the proportions of charged offenders going to trial rather than pleading guilty (or the reverse). In the event, use of the Massachusetts guidelines was not formally mandated by the Superior Court until May 1981; thus, data collected on the use of guidelines up until that time would necessarily have been incomplete since only some of the judges were consulting the guidelines in all cases.

[14] However, see below, Chapters 6 and 7, in which we summarize the data which we were able to obtain (through interviews and observation) about the initial process (if it can be called that) of implementing the Massachusetts guidelines on a trial basis. It is important to distinguish the process of implementation of an institutional change (such as guidelines) from the short-run consequences of such a change. Again, however, there is only a vague and ill-defined borderline between these two things.

[15] It should be noted that the Minnesota Sentencing Commission has continued (in keeping with its legislative mandate) to collect data on sentencing practice since the Minnesota guidelines came into force on 1 June 1981. In time, it is expected that these data will furnish a basis for the "feedback" function envisaged by Gottfredson and Wilkins in their original formulation of the concept of decision-making guidelines (see, for a further discussion, Chapter 3, below); the data should of course, also provide some basis for an evaluation of judges' compliance with the Minnesota guidelines. See further, Chapter 9, and Chapter 11, below.

[16] For Massachusetts, see Chapters 6-8; for Minnesota, see below, Chapter 9.

[17] Our Advisory Committee consisted of Professors Don M. Gottfredson, Andrew von Hirsch and Donald M. Barry, of the Rutgers School of Criminal Justice.

[18] Our special consultants at this meeting were Professor Alfred Blumstein, Judge Morris Lasker, and Professor Herbert Solomon.

[19] See below, Chapter 10, for a discussion of the contrasts between these states in relation to our general model for evaluating statewide sentencing guidelines.

[20] At this time we had several discussions with Dr. Marvin Zalman about the possibility of conducting further

research on the guidelines which he and his colleagues were developing in Michigan; since that time Dr. Zalman has agreed to make available to us the data set from which the Michigan guidelines (see below, Chapter 9) were developed. We did not have the opportunity to obtain and analyze these data during the course of our (original) project grant; we hope, however, to be able to do this in the future.

[21] In such "games", judges, prosecutors and/or defenders might be given the facts of hypothetical cases and asked what sentences they would impose or recommend, and for what reasons. The "game" technique thus permits rigorous experimental manipulation of both the quantity and types of information on which such decisions are based. There are obvious problems of external validity raised by experimental research in the two states -- Massachusetts and Minnesota -- in which our intensive fieldwork would have been done. The combination of experimental with interview and observational data would have done much, we believe, to support the external validity of the experimental research.

[22] Nationwide data on the development of sentencing guidelines have in fact been collected, with the aid of an NIJ grant, by the Criminal Courts Technical Assistance Project, Institute for Advanced Studies in Justice, American University Law School. During the course of our grant, our project staff attended two meetings sponsored by the American University project; information obtained at these meetings supplemented our own national survey of guidelines development.

[23] See, in particular, Chapter 11.

[24] For a similar argument, see Rich, et al., (1980). We find ourselves in sharp disagreement with both the content and the form of many parts of this report; some of the issues on which we disagree with these authors are discussed in Chapter 3, below. However, on <u>this</u> point -- concerning the very short-term scale of many evaluation research projects supported by NIJ and other federal and private agencies -- we are in substantial agreement with the Rich et al. report.

[25] Out of what lawyers call "abundant caution", we ought perhaps to add that this statement is not intended to be critical: we have, after all, no information about the <u>reasons</u> why our continuation proposal did not meet with NIJ approval. In large part, however, this may be due to the fact that we were never notified by NIJ either officially or unofficially, that our phase two application would not be supported.

[26] See, for example, von Hirsch (1976); Singer (1979); Morris (1974); Twentieth Century Fund (1975). There is not by any means complete consensus among these or other authors as to what "just deserts" is supposed to mean.

[27] See, for example, Green (1969); Hogarth (1971); Hagan (1974); Lizotte (1978); and Sutton (1978). For a discussion of some earlier research see below, Chapter 2.

[28] Among modern studies, one of the earliest and most influential was done by Mannheim and Wilkins (1955); while this study was not explicitly concerned with the development of techniques for controlling discretionary decision-making (in this case, decisions to release from borstal institutions for young adult offenders in England), there is little doubt that it stimulated demands for consistency in decision-making policy. On the relations between prediction techniques (of the kind pioneered by Wilkins, in the field of criminal justice), and decision-making guidelines, see below, Chapter 3.

[29] The content of the Minnesota guidelines has also been modified by policy decisions several times since their introduction; thus, preliminary data -- had we been able to collect it -- would have been somewhat inconclusive.

Chapter 2: Concern About Variations in Criminal Sentences: A Cyclical History of Reform

Discretion and Disparity

The use of empirically-based guidelines as a means of minimizing unwarranted variation in sentencing has, of course, only a very short history. The more general question of proposed solutions to the problem of sentencing disparity, however, has been debated for as long as sentence lengths have depended upon judicial discretion.

The concept of discretion has itself been the focus of much dispute and discussion (for a useful summary and review see Davis (1969)). A century or so ago, it was argued by a number of legal writers that judicial discretion could not, by definition, be limited (or "structured" or controlled); the idea of "limited discretion", it was suggested, involved a contradiction in terms.

> When it is said by judges that a matter is in the discretion of a trial court, but that this is not an arbitrary decision, but one governed by rules, the word is used unadvisedly, and inference to be drawn from such language is erroneous. It cannot be said that a matter is left to the discretion of a judge, if that discretion (so called) be reviewable....
>
> To say that there are things in his discretion, but that he must use a "sound discretion" in reference to them; is to give him no discretion at all. (Judicial Discretion, 1880:506-7)

> How far may (discretionary) acts be reviewed by a higher court? If they are properly discretionary...they are in no sense subject to review. Any other answer contradicts the premises. It would be giving a power, but prohibiting its exercise. (Kaufman, 1883:568)

These views suggest that either there is discretion, which cannot be formally limited in its exercise in any way; or else there is no discretion at all, so that, for example, sentences prescribed by law would necessarily be imposed in each and every case. However, the views just quoted are confused in a number of respects. For one thing, as we will see, there is a sense in which a judge dealing with an individual case must exercise a kind of discretion in applying general rules to the facts of that case; he must decide whether or not the facts of that case fall under the concepts contained in one or more legal rules which may apply to the case, and often this involves an "exercise of judgement" which

is left up to him. But it does not follow that such "discretion" is not reviewable, e.g. by an appellate court; on the contrary, where legal rules in the strict sense are concerned, such appellate reviews are common in most jurisdictions.

Similarly, in most jurisdictions, judges typically have a range of sentencing alternatives open to them; they are thus permitted (indeed, they are required) to exercise their individual discretion in choosing a sentence from that range. But it does not follow that this discretion is not reviewable; nor, equally, does it follow that this discretion cannot be guided or controlled, e.g. by general principles or explicit legal rules.

The objective of controlling discretion in sentencing is usually said to be the prevention of "disparity". But what exactly is "disparity"? This term pervades the scholarly literature on sentencing, and is usually defined by a single phrase such as "when like individuals, committing like offenses, are treated differently" (Gaylin, 1974:3; von Hirsch, 1976:29). This definition is misleading in its simplicity. For example, do differences in the sentencing provisions of various state legal codes constitute "disparity"? If so, is it reasonable to separate legislatively-produced "disparity" from judicial "disparity"? What about variations between different regions within the same jurisdiction, e.g. between the northern and southern parts of New Jersey? To what extent should judges be expected to sentence offenders on the basis of community norms which may well differ in different regions? For example, given two towns in the same state, one having one serious crime per hour and the other having one serious crime per year, it might well be expected that two similar offenders, committing similarly serious crimes, would receive different sentences. Would that difference constitute "disparity"?

What this points to is that we need agreed criteria for defining what is a "like case", before we can decide whether two sentences are "disparate" if they involve "like cases" being treated differently. For example, in the case of the two towns with different crime rates, we need to know whether or not "frequency of serious crime" is agreed to be relevant to the judge's choice of an appropriate sentence; and we also need some agreed criteria by which it can be decided which crimes are "serious". Clearly there can be no agreed criteria for defining "like" cases, unless there is agreement on the purposes for which cases are being classified -- that is, on the purposes for which sentences are being imposed. In the hypothetical case of the two towns just mentioned, for example, the prevalence of serious crime might be regarded as

a legitimate factor in sentencing, if utilitarian aims such as general prevention were accepted; it would be otherwise if sentences were to be based on "just deserts" (von Hirsch, 1976).

Arguments about the proper role of judicial discretion, over the past couple of centuries, have in fact had a number of different undercurrents, which have shifted and mixed in complex ways as the general debate itself waxed and wained. On the one hand, there has been a continuing concern about the consistency with which judges exercised their choices among the sentences available to them; this by itself is a complex topic, since there are many ways in which decision-makers can be "inconsistent", and most of them have figured in accusations and pleas for reform at one time or another. On the other hand, there has been an equally persistent concern with the proper objectives of criminal punishment, which has carried with it arguments about the ways in which judicial discretion should or should not be used.

Perhaps the first objection to the range of discretion exercised by judges in sentencing was made by Cesare Beccaria in 1764 (Beccaria, 1963). Beccaria maintained that judges cannot determine punishment for another member of society because they do not have authority as legislators. Legislators have the responsibility to establish punishments for criminal acts because of the authority vested in them:

> Only the laws can decree punishments for crimes; authority for this can reside only with the legislator who represents the entire society united by a social contract (Beccaria, 1963:13-14).

As a corollary to this, Beccaria also felt that judges had no right to interpret the law. He felt if judges interpreted the "spirit" of the law for individual cases, "the law would be the product of a judge's good or bad logic, of his good or bad digestion", or of his passions, or the power or weakness of the accused (1963:16). Beccaria went on to state that judicial interpretation of the law leads to unwarranted disparities in sentence because of the inevitable judge, "who mistakes for a legitimate interpretation that vague product of the jumbled series of notions which his mind stirs up" (1963:16).

While this might look like a "technical" argument against discretion -- one that suggests that judges, being human, are bound to exercise discretionary choices in an irrational or inconsistent way -- it is also a political argument: only legislators, and not judges, have the right to attach penalties to crimes. Similar political arguments are found in

modern times; for example, it has been held that discretion should be eliminated entirely because "discretion in the criminal justice system aPparatus is tied to the present situation in our society, to discrimination, corruption and the abuse of power" (American Friends Service Committee, 1971:143).

Arguments like this, of course, ought to be backed up with empirical evidence (which has seldom been produced) that judicial discretion does in fact have the evil consequences claimed for it. Arguments based on disparity -- in the sense of unwarranted variation in sentencing, as a result of judicial discretion -- need to be backed with evidence that such disparity does in fact exist. In the next section we review briefly some of the efforts which have been made, since the beginning of this century, to show empirically that there is excessive and unwarranted variation in sentencing. These empirical efforts have tended to lead to proposals to reduce or eliminate disparity by controlling judicial discretion, or eliminating it entirely. At the same time, there have been continuing arguments about the nature and purposes of sentencing, which have also led to proposals for sentencing reform. We will see that remarkably similar arguments have been used in each of two cycles of "reform" over the past 100 years or so. The first of these, which reached its climax in the 1940's, led eventually to the "indeterminate" sentence; the second, which began in the early 1970's, involved the rejection of the "rehabilitative ideal", and the widespread adoption of "determinate" sentencing.

Early Empirical Studies of Sentencing

It was not until the early 1900's that statistical evidence of the range of discretion exercised by judges was compiled. In 1914, the Committee on Criminal Courts of the New York City Magistrate's Courts reviewed the more than 155,000 cases disposed of in that year to determine the extent judges' reputations for severe or lenient sentencing behavior were reflected in the dispositions of various classes of cases they had decided. The results, tabulated in their annual report, showed a dramatic variation in sentences given to apparently similar classes of cases.

Perhaps the earliest study containing a statistical examination of sentencing variation was published in 1919 (Everson, 1919). It analyzed data for the New York City Magistrate's Court cases for 1916. There were 42 judges in the Magistrate's Courts at that time who rotated at 15-day intervals among 28 different courts. This resulted in each magistrate presiding over a majority of the courts at one time or another during a given year. The study points out that

under this rotation system, "it can reasonably be assumed that each magistrate handles practically the same classes of cases as those handled by his colleagues" (Everson, 1919:91).

The results of this analysis showed that for peddling without a license, one judge fined all of his cases while another fined only ten percent and suspended sentence for the remainder. For disorderly conduct, one judge suspended sentence in just over two percent of his cases and another suspended sentence in 50 percent of his cases. One judge sentenced 80 percent of his vagrancy cases to the workhouse while another sent only about 17 percent there.

These wide variations in sentence were seen, by one commentator of that period, as unalarming.

> The magistrates are not to be condemned for this variation.... It is assumed that each magistrate is doing his duty as he sees it and is rendering his best service in disposing of cases brought before him (Everson, 1919:98).

It is important to note that at the time of this study, a city magistrate disposed of from 50 to 100 cases daily. As Everson noted, "He must get the facts quickly and decide quickly" (p.98). This circumstance can lead to arbitrariness in sentencing due simply to the lack of information necessary to make a reasoned decision.

It was further pointed out that, prior to 1914, a magistrate was never able to compare his work with that of his colleagues and therefore, "his personal pecularities were inclined to become accentuated" (p.98). Although an important step, these revelations were still only a distant precursor of contemporary sentencing guidelines.

An important study of individual differences in sentencing tendencies among judges was conducted by Gaudet, Harris, and St. John in 1933. They selected 7,442 cases from the court records of one county in New Jersey covering a period of nine years and a total of 21 different offenses.[1] Because all of these cases were sentenced by only six judges who rotated among jurisdictions, the authors felt that each judge could be expected, overall, to give sentences of approximately equal severity as they would likely receive a similar mix of cases.

The percentage of cases sentenced to imprisonment ranged from 34 percent to 58 percent among the judges. The imposition of suspended sentences ranged from a low of 16 percent to a high of 34 percent. It was also found that the

general sentencing tendencies of the judges did not greatly increase or decrease as they gained experience on the bench, and that pending reappointments had no marked effect on the overall severity or leniency of sentences. The authors conclude that these data indicate that the past environment and heredity of the judge must be the factors which influence the severity of sentence as experience and reappointment status apparently have no effect.

It was further concluded that previous studies of the seriousness of crime or of offenders' records that measured these things by length of sentence are of little value if the results of this study were typical of judicial sentencing tendencies.

> In other words the type of sentence received by a prisoner may be either an indication of the seriousness of his crime or of the severity of the judge (Gaudet et al., 133:815).

A secondary analysis of Gaudet's data was conducted for the purposes of this chapter and confirmed that the differences in incarceration rates among the judges were statistically significant.[2] However, a re-analysis of Gaudet's finding that experience on the bench made no difference in judicial sentencing tendencies did not support his result.[3] These findings negate Gaudet's explanation of the possible reasons for the sentencing variation, but they do confirm the existence of variation to a significant degree.

A 1940 study of sentence lengths was conducted by the Newark Evening News in the Essex County (New Jersey) Court of Common Pleas (Frankel, 1940:448-456). The analysis included 4,029 sentences of adult males imposed by four judges during three separate years (1932, 1935, and 1939).

The results indicated that for all offenses taken as a group there were "considerable differences" in the types of sentence imposed by the judges. For example, the use of probation among the four judges ranged from 29.6 to 40.5 percent of the cases, and commitments to the county jail were imposed from 2.2 to 11.9 percent of the time depending upon which judge sentenced the case. However, when a second analysis was done separating those defendants facing multiple charges and those having only one charge against them, the authors found "no striking differences" between the two judges they compared.[4] These two judges were also compared in an analysis of their sentences for seven separate offenses. It was again found that the relative frequency of the types of punishment selected by the two judges in each of the seven offense categories are "not too dissimilar." These latter

comparisons of the two judges were not tested for statistical significance, however.

A secondary analysis, done for this chapter, showed that for all seven offenses combined, there was in fact a significant difference between the sentence types chosen by the two judges.[5] The findings of this study, then, are in accord with those preceding it; there appeared to be significant variation among judges and jurisdictions even when sentencing similar types of cases.

These findings obtained additional support during this period at the federal level. The 1939 annual report of the U.S. Attorney General stated that

> My studies of the disposition of criminal cases in the Federal courts have led me to the conclusion that there frequently occur wide disparities and great inequalities in sentences imposed in different districts, and even by different judges in the same district, for identical offenses involving similar states of facts (Holtzoff, 1941:3).

The Committee on Sentencing, Probation, Prisons, and Parole of the American Bar Association also issued a report in 1939 discussing the findings of a U.S. Department of Justice nationwide study of the sentencing practices of federal judges and found:

> The sentencing records of many judges, as well as the judges' own statements concerning their sentencing practices, show the presence of arbitrary variance and numerous highly subjective factors and personal biases in the imposing of sentences (ABA, 1939:35-37).

Other Pressures for Sentencing Reform

The research reviewed above did not much question the purposes of criminal sanctions; it was aimed instead at showing that those sanctions had been inconsistently applied. As noted earlier, however, there was throughout the period in which that research was done a shift toward "treatment" or "rehabilitation" as the most important objective of sentencing; this led eventually to the adoption of "indeterminate" sentencing, under which variation (as measured by purely punitive criteria) was to be expected. This shift toward "indeterminacy" also led to a number of proposals for sentencing reform and the control of judicial discretion. We describe two of these proposals -- one European, one American -- in this section. Though neither of these proposals was in fact adopted anywhere, both were extremely influential; and

both illustrate techniques for controlling discretion in sentencing which provide an illuminating contrast with later efforts including sentencing guidelines.

Reform of the Italian Penal Code

An initial step in the evolution of sentencing reform occurred in Italy when the Commission for the Reform of the Italian Penal Code was nominated by the Italian Minister of Justice on September 14, 1919. Enrico Ferri was named president of the Commission and, although there were originally representatives from all the various "schools" of criminological thought, the final product reflected Ferri's strong positivistic orientation due to a number of resignations from the Commission. The orientation of the new code was clear as early as Ferri's speech at the Commission's inaugeration: "We shall try to bring the fulcrum of the law from the crime to the criminal" (Ferri, 1920:75).

With regard to sentencing, Ferri felt that individualization of sentences by courts was not possible because no judge could be expected to make detailed studies of every offender. Rather, Ferri felt judges should have enough information to place the offender in the proper class,[6] and then assign a sanction appropriate to the class. For those incarcerated for indeterminate sentences, Ferri thought future revisions of the sentence would be the task of "permanent committees in which judges, prosecutors, defenders...and with them psyciatrists and anthropologists would examine periodically those committed, with the guarantee of publicity, to determine if the term should be prolonged or not" (Cited in Sellin, 1957:481).

The sentencing provisions in the revised Italian Penal Code emphasized "the principle of the dangerousness of the offender" as its fundamental criterion (Ferri, 1921:355-77). This manifested itself in an elaborate guidelines structure consisting of a list of circumstances which indicate greater or lesser dangerousness of the offender.[7] These circumstances were to be used by judges in establishing the length and type of sentence. However, Ferri also describes some rather complex exceptions and modifications for cases where configurations of greater or less dangerousness occur (1921:645-7).

When these conditions and exceptions to the "conditions of dangerousness" are scrutinized, the arbitrariness of many of Ferri's criteria becomes apparent. For example, Article 77 states,

> If there concur together circumstances of greater or lesser dangerousness, the judge shall establish which are prevalent, in order to graduate the dangerousness of the accused and to apply, according to article 20, the sanction best adapted to his personality (1921:546).

This excerpt illustrates how Ferri's large number of prescriptive conditions combined to work against his goal of more specific sentencing rules. That is, there will always exist conditions where the rules are unclear, forcing judges to rely on unguided individual discretion. While it should also be noted that many of Ferri's "circumstances of dangerousness" are incorporated into American penal law as degrees of offense or elements of aggravation or mitigation, the revised Italian Penal Code was, nonetheless, an important step in the elucidation of a more specific sentencing rationale. [8]

Glueck's "Rational Penal Code"

In 1928, Sheldon Glueck published "Principles of a Rational Penal Code," which put forth his conception of the proper orientation and structure for a penal code (Glueck, 1928). Like Ferri, Glueck felt that the current emphasis of the criminal law on the offender, rather than the crime, called for individualization of treatment as well. He pointed out that the trend toward indeterminate sentences had, in practice, had little effect on the traditional "mass-treatment" method of dealing with offenders.

Glueck went on to criticize accepted sentencing practices based on legislative prescription of penalties for various offenses.

> Legislative prescription in advance of detailed degrees of offenses is individualization of acts and not of human beings, and is, therefore, bound to be inefficient (1928:467).

He further maintained that judicial individualization of offenders without "scientific facilities in aid of the court" was destined to "deteriorate into a mechanical process of application of certain rules of thumb or of implied or expressed prejudices" (p. 467).

Glueck's view represents an important move toward empirically-based sentencing guidelines inasmuch as he recognizes that pre-existing sentencing guides for particular offenses can never anticipate the diversity of human circumstances which may be present, and that there must be a more scientific method for determining sentence lengths than the mere application of "arbitrary" legislative rules.

Glueck also devoted much of his paper to a critique of Ferri's Italian Penal Code. In addition to its arbitrariness in selecting sentence lengths and types based on pre-conceived conditions, Glueck felt that Ferri's one-dimensional emphasis on the dangerousness of the offender was "both unjust and unscientific" and also "underemphasizes the rehabilitative possibilities of the offender" (1928:469). He pointed out that Ferri's criteria of dangerousness permanently labelled the offender prior to any treatment he might receive, forcing the judge to sentence speculatively, a task better carried out by parole boards, or as Glueck recommended, a "Socio-Penal Commission" (or "treatment board") made up of social scientists who would determine the type of treatment best suited to the individual offender as well as its duration.

These recommendations, in Glueck's view, provided for a more truly indeterminate sentence in which the correctional process would be expressly aimed at the offender rather than prescribed for a particular act. The specific means by which this "treatment board" would decide on the proper sentence, however, was not discussed by Glueck, and except for his references to the need for new and improved treatment methods, it remained unclear, in 1928, what the specific alternative to unchecked judicial discretion might be.[9]

Perspectives of Concern About Sentencing

Subsequent to the empirical studies of sentence variation during the 1930's and 1940's, it is clear that the growing interest in sentencing reform emanates from two perspectives. First, there is increasing concern about sentencing variation itself. Studies appeared which, for the first time, statistically documented the existence of wide sentence variation among judges and jurisdictions; and there was concern about the unfairness to offenders which this variation seemed to entail (Frankel, 1940:454, Holtzoff, 1941:3).

The second perspective, as expressed by Ferri and Glueck, saw the need for sentence reform out of a dissatisfaction with legislative rules which did not take into account the "treatment" potential of offenders.

Interestingly, both these perspectives are closely linked to the trend toward indeterminate sentencing throughout the country. By 1941, approximately three fourths of the states were operating indeterminate sentencing systems in one form or another, and the Conference of Senior U.S. Circuit Judges had just adopted a resolution favoring the adoption of an indeterminate sentencing law for the federal courts.

Among the arguments for a change to indeterminacy at the federal level was the view that the existence of an administrative agency (the parole board), to determine the actual period of imprisonment within the minimum and maximum set by the court, would serve to make penalties more uniform.

> The consequence necessarily emerges from the fact that a single board, or perhaps several boards, acting in close cooperation, determine the length of incarceration to which every prisoner should be subjected. In other words, all cases clear through the same channel (Holtzoff, 1941:6).

This board was also seen as taking the guesswork out of sentencing by reducing "arbitrary variances in the treatment of prisoners possessing similar case histories" (Morse, et. al, 1941:23).

Viewed in this way, the change to an indeterminate sentencing system would simultaneously correct both shortcomings of the existing system. First, much of the responsibility for sentencing would essentially be removed from the judges, as they would only set an indeterminate range of time to be served by the offender. A single parole board would set the actual time to be served at a later date. Second, the parole board would make their judgements based on the offender's treatment potential and his rehabilitation while incarcerated. Thus, indeterminate sentencing was seen both as an answer to wide sentencing variations and as a recognition of the individualized treatment needs of the offender.[10] That this involved a fair amount of self-contradiction seems to have been noticed by nobody.

The Cyclical Concern for Sentencing Reform

The arguments used in support of indeterminacy are interesting for a number of reasons, not the least of which is the fact that these same arguments have been widely used during the 1970's in support of determinate sentencing. In recent years, indeterminacy has been attacked on the same grounds for which it was established during the 1940's.

Criticisms of sentencing during the 1970's are strikingly similar to those of the 1930's and 40's. In both instances existing sentencing practice is being criticized due to its lack of uniformity and unfairness to offenders. Note the view of George Vold in 1941 as he argued in support of indeterminacy:

> The essential purpose of providing protection for society against the criminal is in fact facilitated by

> the generally greater consistency in operation of a single sentencing agency as compared with the variety of conflicting views represented in sentences imposed by a number of individual judges (Morse et. al, 1941:39).

Compare this with the view of Andrew von Hirsch in 1976 in arguing for determinate sentencing.

> The most obvious drawback of allowing wide-open discretion in the name of 'individualization' is the disparity it permits. Judges whose sentencing decisions are unchecked by general standards are free to decide similar cases differently (von Hirsch, 1976:29).

These arguments are essentially the same; however, they are being applied to exactly opposite philosophies of sentencing. Other examples can readily be found if one compares the rationale offered for sentencing reform during these two periods.[11]

The question now becomes, how could the same arguments be used in support of divergent sentencing philosophies? As it turns out, history appears to repeat itself here also.

Dissatisfaction with Indeterminacy: Determinacy Revisited

It did not take long for dissatisfaction with the "new" indeterminate sentencing system to arise during the 1950's and 1960's. In much the same way as with the system preceding it, questions began to emerge about whether the indeterminate sentencing system was achieving its goals. The first question related to the ability of indeterminate sentencing to reduce wide variations in sentences. This provoked the development of a number of methods to reduce this variation, such as sentencing institutes, appellate review, sentencing commissions, and sentencing guidelines. The ultimate result has been a nationwide trend back toward "determinacy", based on much the same rationale as was used in establishing indeterminacy some 40 years earlier.

One of the earliest challenges to indeterminacy was put forth by the Council of Judges of the National Probation and Parole Association (later renamed the National Council on Crime and Delinquency). Formed in 1953, the Council of Judges of the NPPA published, as one of their first projects, Guides for Sentencing, in 1957. This short book outlined the objectives of sentencing and made recommendations for achieving more equitable distribution of sentences and reducing disparity. The suggestions they offered, however, were very broad in scope and offered nothing specific for making sentences more equitable or fair. For example, "'there

can be no fixed formula' for sentencing because different individuals in trouble require different types of help" (1957:59).

It is interesting to note that the second edition of Guides for Sentencing (1974) included a separate chapter on sentence disparity. Citing the 1967 President's Crime Commission and the experience of Federal and state sentencing institutes, three more specific methods of reducing unwarranted sentencing disparity were offered: sentencing institutes, sentencing councils, and appellate review of sentences.

Sentencing Institutes

In 1958, the House Committee on the Judiciary issued a report and joint resolution "to improve the administration of justice by authorizing the establishment of institutes and joint councils on sentencing" (U.S. Congress, 1958:1). These institutes were established by statute "in the interest of uniformity in sentencing procedures."

This legislation reflected a growing concern among criminal justice officials and in Congress about wide variations in sentences that had been appearing for a number of years (see U.S. Congress, 1958:5-8, Sharp, 1959:9). Also addressed in the House Report were the releasing policies of the U.S. Parole Board, which were seen as the source of wide discrepancies in the amount of time served by offenders sentenced to incarceration.

At the time of this legislation federal judges fixed the maximum sentence (up to the statutory limit) leaving parole eligibility at one-third of this maximum. The joint resolution proposed that the judge also be permitted to specify the offender's parole eligibility date at any time up to one-third of the imposed maximum; alternatively, it was suggested that the judge should set only the maximum term, allowing the parole board to set an eligibility date. As the House Report claimed:

> This procedure in the case of a serious chronic offender would permit the judge to set both the maximum terms and the parole eligibility date at the statutory limits...In doubtful cases the judge could set a long maximum term and leave the matter of parole eligibility to the determination of the Parole Board (1958:9).

This legislation became into law in August 1958, and the first pilot institute on sentencing took place in 1959 under the auspices of the Judicial Conference of the United States.

Although no comprehensive evaluation of the effectiveness of sentencing institutes in reducing sentence variation has been conducted, subjective assessments have pointed to their educational function for judges. Following the pilot institute, "there was general agreement that the agenda for future meetings should combine discussions of specific cases illustrating sentencing problems with discussions directed toward the development of a sentencing philosophy" (Sharp, 1969:11; Levin, 1966:503). The fact that judges were becoming aware of the wide variations in philosophy and criteria used by them in sentencing provided an important first step in establishing the need for more specific guides in sentencing.

Subsequent to the pilot institute, 15 others were held, in the years through 1965. Three of these included federal judges from more than one circuit, the first of which was conducted in 1961 (Remington and Newman, 1962). While claims were made during the mid-1960's that decreasing federal prison populations and reduction in sentence disparities could be attributed, at least partially, to the sentencing institutes, this was never demonstrated to be the case. (Youngdahl, 1966:518-9). As one of the institutes' participating judges maintained, "We cannot expect to achieve uniformity of sentences but rather uniformity of procedures" (1966:519).

Sentencing Councils

In November 1960, the U.S. District Court for the Eastern District of Michigan initiated use of sentencing councils on a trial basis, after attending the pilot sentencing institute in 1959 and discovering the wide disparities in sentences within their own district. "Sentencing Councils" would involve weekly meetings of judges together, in which the sentencing decisions to be made in cases during the coming week would be discussed. While the trial judge retained ultimate responsibility for sentencing each individual case, it was felt that such a council would (1) provide an opportunity to assess issues of disparity using actual cases and presentence information in a group setting, and (2) serve to develop a consensus in sentencing philosophy among the judges (Doyle, 1961:28).

An examination of the first 203 cases considered by the sentencing council in Michigan found wide differences in sentencing recommendations among the judges for all types of offenses covering every type of disposition.

> In the case of an extortionist, dispositions recommended were for observation and study, and institutional sentences of 2 years, 4 years, and 10 years...In respect to a postal violator, the five judges

> indicated sentences of 2 years, 3 years, 4 years, 5 years, and 6 years (1961:29).

A unanimous disposition recommendation occurred in only 76 of the 203 cases, but the sentencing judge altered his original recommendation in 72 of the 175 cases recorded. It was also noted that there was "no substantial difference" in the number of increased sentences compared to the number of reduced sentences. Therefore, it appears that judicial sentencing conducted in a non-binding group environment affected the sentencing decision to individual judges to some extent.

The experience in Michigan continued and was adopted in other federal judicial districts with slight modifications. When the caseload and number of judges increased in Michigan, it became necessary to reduce the council meetings to smaller groups of three judges, rather than to assemble the entire bench (Levin, 1966:503). Also, a recommendation chart was developed to standardize the criteria used in sentencing decisions. The factors considered were prior record, family responsibility, work record, attitude, nature of offense, lack of adequate plan for probation, and an "other" category (1966:510). It can be seen that the general notions of fairness, which provoked the development of the sentencing councils, resulted in the adoption of more standarized sentencing criteria and, therefore, a more uniform philosophy.

An evaluation after the first five years of the Michigan experience found that between 40-50 percent of initial sentence recommendations were altered after council meetings in each of the five years. Overall, there were more reduced sentences than there were increased sentences (266 to 183). The evaluation concluded that,

> Prison terms are generally shorter...(and) the percentage of offenders placed on probation in one district has progressively increased from 45 percent five years ago to 60 percent today. Yet the percentage of probation violators has not increased (Levin, 1966:507).

A separate evaluation found similar results in declining commitments to prison and increased sentences to probaton (Hosner, 1970:20).

While the sentencing council has survived legal challenges, and has received the endorsement of the American Bar Association, it has certain limitations in its ability to reduce wide variations in sentence (Hosner, 1970:23-4; Diamond and Zeisel, 1975). First, councils only promote consistency in sentencing within a single jurisdiction -- in this case, a

federal district (Levin, 1966:508). Interjurisdictional variations are not addressed by sentencing councils. Also, judges sitting in rotating circuits or in geographically distant locations have obvious difficulties in discussing sentencing decisions with their colleagues. So while the sentencing council has proven useful in certain districts, especially in the initiation of newly appointed judges (1966:508), logistical and structural problems have precluded its widespread adoption.

Appellate Review

A reform, contemporary with sentencing institutes and councils, was appellate review of sentences. As was true for all of its predecessors, appellate review was developed primarily as a mechanism to reduce wide variations in sentence. Proponents of appellate review saw it as a way to correct an "unduly harsh sentence." Also, the opinions written by sentence review judges would "provide guides and standards" for trial judges to better utilize their discretion in choosing among the various theories of punishment in a particular case. "Thus, the tendency would be toward more uniform sentences within the jurisdiction." Further, this was seen as a possible method to reduce "discontent among prisoners" (Criminal Procedure, 1961:188-9).

By 1960, nine states had a system of appellate review whereby appeals against sentence were added to courts having jurisdiction over appeals against conviction. In these states, the appellate court was permitted to either affirm or reduce the sentence imposed by the trial court. Through this procedure it was hoped that sentences clearly departing from "normal" sentences imposed in the state, and lacking adequate justification, would be modified by the single, statewide appellate review procedure.

These systems of review were soon criticized, however, for not adequately achieving this objective. Some of the problems identified included: (1) The large amount of time and money needed to pursue a sentence review restricted the opportunity to exercise this option for many offenders, (2) Without the offender present his demeanor could not be assessed as in the trial court stage, (3) There was a temptation by the reviewing courts to "correct" non-reversible errors made by the trial court by reducing the sentence, (4) The opinions of appellate panels often "did not disclose any effort to articulate and establish criteria for sentencing," and (5) The inability to increase sentences "too short to achieve the relevant criminal law goals" prevented the establishment of consistent sentencing principles (Appellate Review, 1960:1461-2; Richey, 1978).

These criticisms were partially responsible for the development of somewhat modified forms of appellate review in Massachusetts, and later in Connecticut. In these states the appellate panel was able to both reduce and increase a sentence. Connecticut went further and required the appellate panel to provide written reasons for their decision on each appeal.

An analysis of the first 256 sentence appeals heard in Connecticut revealed that only 15 (6%) sentences were reduced while seven (2%) were increased. Such a small number of modifications provoked doubts about the viability of appellate review. This reluctance to alter sentences was explained as either being due to the type of criminal requesting a sentence review (only the most dangerous with nothing left to lose), or due to an "excessive deference" to the trial judge (Appellate Review, 1960:1464-5; Labbe, 1977:128).

The unfortunate effect is that prisoners have seen the infrequency of reductions as an indication that the process is a "sham". As a result the ability of appellate review of sentences to accomplish the goals it was designed for has been questioned (McAnany, Merritt, and Tromhausa, 1976:639; Appellate Review, 1960:1465-6).

In 1968, however, the American Bar Association recommended the adoption of appellate review of sentences on both the state and federal level (A.B.A., 1968). This recommendation, combined with increasing concern about criminal sentencing in general, led to the introduction of appellate review in an increasing number of jurisdictions. A 1977 survey reported that some form of sentence review was available in 23 states, although most reviewing courts tended to limit their review "to correcting only the most glaring abuses" (Labbe, 1977:123-8).

Nonetheless, the debate over appellate review has continued with opponents claiming that (1) appropriateness of sentence may not be a proper question for the judiciary, (2) there is a reluctance to alter sentences of lower courts due to resentment it may cause, (3) overcrowded appellate courts will be further burdened with sentence appeals, and (4) appellate panels will establish "acceptable sentences" for certain crimes and situations thereby discouraging trial judges from thinking for themselves. Alternately, supporters claim that (1) intervention by the executive occurs too seldom to substitute for an appellate review procedure, (2) the experience of states adopting sentence review indicates no sudden overcrowding of appellate courts had occured, (3) most sentences are imposed following guilty pleas giving trial judges little advantage in having observed the defendant's

demeanor in choosing the appropriate sentence, and (4) "acceptable sentences" would not result from sentence review but a "uniformly fair and equitable approach" to sentencing would be the result (Labbe, 1977:130-132).

The burden of proof has been upon the proponents of appellate review of sentences, and the evidence has been mixed. While most European countries have long considered the sentence as a matter of law, and therefore, reviewable (Mueller and LePoole, 1968), there is a continuing debate in the United States about the ability of appellate review to develop, in practice, more objective criteria for sentencing and guides for their application (Compare Appellate Review, 1960: 1466ff. with Thomas, 1967,1968). Further, even though more specific criteria for sentencing may be generated using sentence review, these criteria, no matter how specific must always be applied to an individual case. As a result, it is unclear whether the development of criteria alone can accomplish the goals of sentence review. The practicality of also developing rules for their application is even more suspect. Given the example of Ferri's Italian Penal Code, where extreme detail was used in specification of sentencing criteria leading to ambiguity about their application, the role of discretion versus specific rules for sentencing remains clouded in assessing the effect of appellate review of sentences.

Disagreement About Disparity

Because many jurisdictions allowing for appellate review of sentences have experienced only a small proportion of sentence modifications, the question of the true nature and extent of disparity has been questioned (DeCosta, 1968:59). Can the concern about wide variation in sentencing be greater than its actual occurrence in practice?

While many statistical analyses of criminal sentences have been conducted, most of which claim the existence of large and unwarranted variations, few have been methodologically sound enough to warrant repetition. Errors in sample selection, assumptions about randomness, disregard for paroling policies, and inappropriate statistical manipulation of data account for most of the errors found in the literature (See Baab and Ferguson, 1967; Seymour, 1973; Johnson et al., 1973; Tiffany, Avichai, and Peters, 1975; Kulig, 1975). Fortunately, efforts have now emerged which demonstrate the existence of disparity in certain jurisdictions and account for many of the methodological concerns noted above (See Hagan, 1974; Kilpatric and Brummel, 1976).

As a result, renewed consideration of measures to remedy these wide variations took place during the 1960's and 1970's now that any doubts had been removed as to its existence. This concern intensified when prison riots during the early 1970's were blamed partially on inmate unrest caused by disparity in sentences (Attica-The Official Report of the New York State Special Commission on Attica, 1972).

Discretion and Its Control

During the 1970's still another technique was developed for controlling discretionary decision-making in the criminal justice system: to wit, empirically-based guidelines. Most of the rest of this report will be concerned with this technique, and with its applications in particular jurisdictions. At this point, however, it may be useful to reflect briefly on the general problem of controlling discretion, in the light of the historical efforts to cope with this problem that have been sketched in this chapter. What exactly is the problem, and why exactly has its solution seemed so difficult?

A useful model with which to approach the kind of discretionary decision-making with which we are here concerned derives from the concept of "law" propounded by the German jurist Hans Kelsen (1961).[12] In brief, Kelsen argued that to say that there is a law, or a legal duty, concerning some kind of behavior is to say that there is a legal rule to the effect that a court (or some similar body) ought to impose a "sanction" of some kind if the behavior in question takes place. Thus, for example, to say that there is a law against theft means that there is a legal rule of the form "if somebody steals, he shall be punished." Kelsen conceded, of course, that the popular interpretation of the statement that theft is illegal is that there exists a legal rule of the form "Thou shalt not steal"; but he contended that such "secondary" rules were "contained in" the rules attaching sanctions, which he regarded as "primary" norms (see, e.g., Kelsen, 1961:61).

For present purposes, we need not concern ourselves with whether Kelsen's analysis of the concept of law is correct, or with whether rules of the form "If somebody steals, he shall be punished" are "primary" or "secondary" norms of law. (For a general discussion of this and related problems, see Hart (1961).) It is enough, for present purposes, to observe that we can indeed write down rules of the form "If X has occurred, then Y is permissible" -- for example, "If somebody steals, he may be punished" -- corresponding to any valid rule of the criminal law. Let us call such rules sanctioning rules; they prescribe some connections between antecedent conditions X (e.g. the commission of a theft), and the imposition of a "sanction" Y (for example, imprisonment or capital

punishment). These are of course the rules which are applied by a sentencing judge; and if there is discretion in sentencing, or disparity in the use of that discretion, then that discretion or disparity are concerned with the application of these sanctioning rules.

Before considering problems of discretion and disparity, however, let us look in more detail at sanctioning rules themselves. The first point to note is that, while we can indeed write down such rules, they will in many cases be quite complex: far more complex than statutory provisions (which they sometimes superficially resemble) may reveal. To begin with, the antecedent conditions X may be, and typically are, compound: that is, the single symbol X is really a shorthand for a whole set of antecedent conditions [x1, x2, ... xn] which licence a sanction. These conditions xi will usually (though not necessarily) include proof of the commission of an offense; but they may also include some facts about the person found to have committed that offense, e.g. that he or she is sane, or of a certain age. Typically the conditions usually have jointly to be fulfilled, before the sanction is licensed. (Thus, there must be a theft, and the offender must be legally sane, and legally an adult, etc., before imprisonment can be ordered.)[13] Similarly, the sanction term Y is typically compound: that is, it is a shorthand for a set of sanctions [y1, y2,...yn]. Typically, however, this set is what may be called a disjunctive set. That is, fulfillment of the antecedent X licenses either y1 or y2 or ....; but this too is not necessary, as combinations among the yi are possible (as in "split sentence" provisions for jail and probation, or the combination of a fine and imprisonment).

A second point about sanctioning rules is that they are neutral as to the purposes for which sanctions may be imposed, and the effects which it may be hoped that they will bring about. The Ferri (1921) code discussed earlier illustrates this point clearly (and Glueck's (1928) proposals do so less clearly). According to Ferri's draft code, if certain facts taken by him to indicate "dangerousness" were found, then certain sanctions would be permissible; this relationship could be described by a sanctioning rule of the form "If X, then Y". But such a rule could also be written, to describe the attachment of capital punishment to a finding of theft (by a sane, adult, etc., offender); this is so, whether the purpose of that sanction was deterrence, incapacitation, or "just deserts". As one of us has argued elsewhere[14], on a pure "treatment" model we need not even include among the antecedent conditions X the proof of the past commission of a crime; indeed, the logic of "positivistic" approaches (like Ferri's) to the problem of "social defense" suggests that we often would not do so. All that would be needed, on a purely

preventive model of the criminal justice system, would be conditions X (i.e., a set of antecedents xi) which reliably and accurately identified probable future criminals; and these might conceivably have nothing to do with crime in the past.

Finally, though we have so far confined the concept of sanctioning rules to the representation of valid legal rules, we can without difficulty extend it to cover other sorts of prescriptions governing the imposition of sanctions. For example, it may be that within a given legal system, with its set of rules of the form "If X, then Y", there may be other rules accepted by a particular group of decision-makers -- say, the judges in a particular city -- which so to speak lie within the general sanctioning rules of the system. By "lie within" we mean that the sub-group's rules must not prescribe something that would not be permitted by the general legal rules of the jurisdiction; such a sub-group, however, might well adopt much more detailed rules specifying connections between certain antecedent conditions and certain sanctions, provided that these did not conflict with the general legal rules. To take an even more extreme case: every individual judge (or similar decision-maker) might have his or her own set of rules of the form "If X, then Y". Provided that these prescriptions did not purport to authorize something not licensed by the more general rules of the jurisdiction, the two sets of rules might function simultaneously. In such a case, we might set out to study how far a particular judge consistently applied his own sanctioning rules, as well as asking about the relations between his rules and the (necessarily broader) rules of the jurisdiction.

Where do "discretion" and "disparity" come into all of this? Let us take discretion first. One element of discretion arises just because the prescribed set of sanctions Y typically contains a number of possible choices: the sentencing judge (or other rule-applier) may impose either probation, or a fine, or imprisonment, or (sometimes) nothing at all; and -- this is the important point -- the choice among these alternatives need not be governed by any other rules or formally stated policies. This is not by any means the only element of "discretion" involved in the application of sanctioning rules; but it is far and away the most important element.[15] Giving discretion to a sentencing judge is, in effect, handing him a set of sanctioning rules containing disjunctive sets of possible sanctions, and saying: "It's up to you: you choose."

"Disparity" is another matter. This term of abuse -- for that is what it is -- implies not merely variation in sanction-choosing, but unjustified variation: that in turn implies some objective or purpose, in terms of which the

decisions in question can be evaluated and found to be unjustified. Usually, it seems, the standard applied has been based on some notion of justice or fairness; that, in turn, rests on some notion of an equivalence between the seriousness of particular kinds of crimes, and the severity of particular sorts of penalties. (We return to this question, and to an empirical assessment of seriousness and severity, in Chapter 5.) However, it is important at this point to see that "disparity" in sentencing can come about in a number of quite different ways. First, it may be that all judges in a particular jurisdiction operate under the same set of sanctioning rules, but merely apply these rules to concrete cases in an erratic way; this is perhaps the "judicial digestion" case envisaged by Beccaria (1963). Second, however, it may that different decision-makers, operating within the same jurisdictional context, are consistently applying their own sets of sanctioning rules; but that these rules differ. Thus, at least some of the variation in sentencing over the past century, described earlier in this chapter, may have come about because some judges were in effect working with a set of rules shaped by the aim of "reform" or "rehabilitation", whereas others were working with rules shaped by the aim of deterrent punishment; these different sets of rules may have featured different antecedent criteria, and may have attached sanctions to those criteria in different ways. (The term "disparity" is sometimes used to refer to what is perhaps a special case of this: namely to sentences based in part on morally iniquitous factors such as race. "Discrimination" is perhaps a better term for this.)

Let us now re-examine the various efforts to control discretion sketched earlier in this chapter, using the concept of sanctioning rules just outlined. One way to control discretion in sentencing is to limit the possible sanctions Y which can apply to a particular set of antecedents X, and in the extreme case to limit them to just one measure or form of sentence. This is equivalent to making the sanctioning rule "If X occurs, then Y is permissible" read "If X occurs, then Y is mandatory" --where Y in this case contains just one sanction, e.g. capital punishment or imprisonment. This is precisely what was done, of course, in early nineteenth-century England, in which virtually all crimes were (by law) capital crimes; in more recent times, and in the United States, such mandatory sentences have been rare, and have been confined to particular offenses (some homicides, some firearms offenses, some narcotics offenses).[16]

At the other extreme, one might attempt to control discretion by providing very elaborate sets of antecedent conditions X, and attaching these to very specific sanctions Y: Ferri's (1921) code was perhaps an example of this, though

inspection of it will show that it probably would not in fact have accomplished much by way of fettering discretion among rule-appliers. In between these two extremes -- the complete abolition of choice, and thus of discretion, on the one hand, and intricately rule-governed choice on the other -- lie a number of techniques. Some of these techniques -- sentencing councils and sentencing institutes, discussed above, for example -- do not involve altering the rule-structure within which discretionary decision-making takes place; instead, they are essentially techniques for attempting to persuade or influence decision-makers to work consistently within a particular framework of rules. Other techniques -- such as so-called "presumptive" sentences, exemplified by California's Uniform Determinate Sentencing Act of 1976 -- place a rebuttable constraint on sanctioning rules; they say, in effect, that "If X, then Y shall be imposed -- unless further conditions W obtain, in which sanction Z may be imposed". The effect of rules of this kind is plainly like that of a particular judge, or group of judges, deciding to adopt some more detailed sanctioning rules, within a general structure that permits a wide range of discretionary choice; the difference is that the constraint on discretion in the case of presumptive sentences is embodied in the general legal framework prescribing the basic sanctioning rules, and is not an independent constraint on permissible choice.

Sentencing guidelines are also an intermediate technique for controlling discretion in decision-making; that is, they lie between the two extremes of mandatory sanctioning rules on the one hand, and elaborate ones on the other. The ways in which guidelines are supposed to work are described in the next chapter.

Notes to Chapter 2

[1] The crimes were chosen because of their relatively great frequency. The crimes which were used were: larceny, larceny and robbery; breaking, entering and receiving; breaking and entering; robbery; embezzlement; burglary; assault; battery and robbery; larceny from the person; assault and battery with intent to rob; violators of the Hobart Act (New Jersey's prohibition law); adultery; rape; assault; battery and rape; assault and battery with intent to rape; abuse; carnal abuse; and finally assault and battery with intent to abuse.

[2] $X2=355.1$, $df=15$, $p<.001$.

[3] The original study did not include tests for statistical significance, thereby permitting disparate results upon subsequent analysis. $X2=292.24$, $df=4$, $p<.0001$.

[4] "The figures are presented for only two of the judges having a large enough number of cases to make the findings significant" (Frankel, 1940:449-450).

[5] $X2=293$, $df=3$, $p<.001$.

[6] Ferri identified the classes of offender types: the born or instinctive criminal, the insane criminal, the passional criminal, the occasional criminal, and the habitual criminal.

[7] "Circumstances which indicate a greater dangerousness in the offender": "(1) Dissoluteness or dishonesty of prior personal, family or social life; (2) Prior judicial and penal record; (3) Abnormal organic and mental conditions before, during and after the offense, which do not constitute mental infirmity and which reveal criminal tendencies; (4) Precocity in committing a grave offense; (5) Having acted through ignoble or trivial motives; (6) Family and social relationship with the injured or damaged party; (7) Deliberate preparation of the offense; (8) Time, place, instruments, manner of execution of the offense, when these have rendered more difficult the defence by the injured or damaged party or indicate a greater moral insensibility in the offender; (9) The execution of the offense by means of ambush or strategem or through the commission of other offenses or by abusing the aids of minors, the deficient, the unsound of mind, the alcoholic, or by employing the assistance of other offenders; (10) The execution of the offense during a public or private calamity or a common danger; (11) Abuse of trust in public or private matters or malicious violation of special duties; (12)

Execution of the offense on things confided to the public good faith or kept in public offices or destined for public utility, defence or reverence; (13) Abuse of personal conditions of inferiority in the injured part or of circumstances unfavorable to him; (14) Having aggravated the consequences of the offense or having through the same act and not by mere accident damaged or injured more than one person, or having through one and the same crime violated various provisions of law, or the same provision of law at various times, by acts carrying out one and the same resolve; (15) Blameworthy conduct after the offense towards the injured or damaged party or his relations, the person present at the time of the offense; (16) In offenses by imprudence (negligency) having caused the damage in circumstances which made it very probable and easy to foresee."

The following circumstances indicate "less dangerousness in the offender": "(1) Honesty or prior personal, family and social life; (2) Having acted from excusable motives or motives of public interest; (3) Having acted in a state of excusable passion or of emotion through intense grief or fear or impulse of anger unjustly provoked by others; (4) Having yielded to a special and transitory opportunity or to exceptional and excusable personal or family conditions; (5) Having acted in a state of drunkenness or other form of intoxication not to be foreseen by the offender, through transitory conditions of health or through unknown material circumstances; (6) Having acted through suggestion coming from a turbulent crowd; (7) Having used, spontaneously and immediately after having committed the offense, all exertion to diminish the consequences or to make good the damage, even in part, if it be done with sacrifice of one's own economic condition; (8) Having in repentance confessed the offense not yet discovered or before being interrogated by the judge, immediately after the offense" (Ferri, 1921: Chapter II. Art. 21, 534-536).

[8] The new code was completed in 1921, but was never implemented by the Italian parliament due to post World War I unrest leading to the Facist revolution.

[9] In his late writings on this subject, however, Glueck did advocate the use of prediction tables, for which he and his wife had by then become well-known: see Glueck (1963),

[10] It should be noted that of the academics, federal judges, and criminal justice officials who commented favorably on the bill in 1941, most saw indeterminacy primarily as a means of reducing disparity in sentences, rather than as a method of individualized treatment (See Morse et. al, 1941:22-42).

[11] Compare the comments of Morse et. al (1941:28,41) with Frankel (1973: Chapter 4) and Gaylin (1974:165).

[12] This idea probably first appears in the writings of Hagerstrom (1931), and is also espoused by later Scandinavian "realist" jurists, e.g. Olivecrona (1971). For a general discussion see Hart (1961), esp. Chapter V.

[13] This set could be a disjunctive one, e.g. there could be one sanctioning rule attaching the same penalty or set of penalties to any number of crimes. It seems clearer not to treat the matter in that way, however.

[14] This matter is further discussed in a forthcoming paper by Sparks on "The Assessment of Seriousness and Severity."

[15] It should be noted there will still be an element of discretion, in one sense of that term, even if this happens; it will still be necessary to determine whether or not the facts of the instant case fit the description X. It can be argued that such determinations are capable of objective settlement in most cases, though in practice, of course, they may be largely influenced by judicial (or prosecutorial) attitudes, e.g. a dislike of mandatory sentences. Cf. Heumann and Loftin (1979).

[16] Mandatory minimum sentences merely limit the range of sanctions Y; they do not change the connection stipulated in the sanctioning rule.

Chapter 3: The Concept of Sentencing Guidelines

This chapter has two objectives. The first is to trace briefly the origins and history of the concept of sentencing guidelines, as this concept has developed over the past decade or so. The second is to examine in some detail certain features of the concept itself; and, in particular, to try to clarify certain features of this technique for controlling sentencing discretion about which there still seems to be a certain amount of confusion -- at least if one is to judge by some recent writings on the subject.

The origins and history of the concept are important for two reasons. The first is that the impact which this concept has had, on various states' efforts to promote change in sentencing policy and practice, seems to be due in part to the way in which the concept was originally developed, and -- more importantly -- to the way in which it was subsequently promoted. An extensive (federally supported) effort was made, in the years before our research began, to disseminate a "cookbook" approach to the construction of empirically-based sentencing guidelines which, in our opinion, is fundamentally misconceived. As we shall see, the sentencing guidelines actually developed and implemented in different states to date differ among themselves, and differ also from the models proposed by the concept's originators. In our opinion, some of these differences have been beneficial; some, however, have not.

The second reason for reviewing the historical origins of the notion of sentencing guidelines is that these origins help to explain some of the problems that are (in our opinion) inherent in this technique for controlling discretion in sentencing. An understanding of these problems is necessary, in turn, if we are to evaluate the technique itself (as distinct from its application in a few jurisdictions over the past few years). In order to clarify some of these inherent problems, we make use of the concept of sanctioning rules developed in the preceding chapter. We then show that in a number of respects sentencing guidelines cannot be "empirically derived" in any realistic sense of that term -- contrary to what the "cookbook" approach just referred to implies. Finally, we consider some problems of inducing or enforcing compliance with sentencing guidelines; these are central to the general model for evaluating guidelines which we discuss in the concluding chapter of this report.

Origins of the Concept

The origins of the concept of sentencing guidelines can be traced to a small study carried out (as part of a larger project) for the U.S. Board of Parole by Don M. Gottfredson and Leslie T. Wilkins, beginning in 1971.[1] This project was concerned with providing information to the Board's Youth Corrections Division, about parole practice concerning offenders sentenced under the Federal Youth Corrections Act (see Hoffman, 1975). Although technically eligible for parole at their first hearing (usually three to six months after reception into prison), most offenders sentenced under this Act were "continued" at that hearing, for an additional period in prison of anything up to about three years; most of the Division's decisions on cases, therefore, were essentially time-setting decisions, rather than yes-or-no decisions on immediate release. The object of the research[2] was to examine the apparent determinants of the lengths of time offenders would serve. For this purpose, board members were asked to score cases on each of several scales -- including four relating to perceived seriousness of the commitment offense, risk of parole violation, institutional misconduct, and participation in institutional programs -- before deciding on the recommended time of "continuation". It was found that perceived ofense seriousness and parole prognosis were the factors most highly associated with the outcome variables; and from regression analyses based on these factors, it was possible to calculate expected times to be served, for given combinations of perceived seriousness of offense and parole prognosis. Parole boards, it was suggested, could use such expected values -- which were presented to them in the form of a two-dimensional matrix, with cells containing expected numbers of months to be served -- both to describe or to assess policies and to reduce disparity in individual case decision-making.[3]

Shortly after this, the U.S. Parole Board launched a pilot project to test the feasibility of regionalizing its decision-making (see Gottfredson et al., 1975). Decisions to release on parole in individual cases were to be delegated to two-man panels of "hearing examiners", with provisions for appellate review of decisions by the whole board; there would thus have been substantial decentralization of decisions that had formerly been made under the auspices of one entity (namely the Board itself). The Board had been under heavy criticism for some years, on the ground that decisions to release on parole were often arbitrary and inconsistent; in part to meet these criticisms, the pilot project (based at five prisons in the northeastern United States) included provision for what were called "decision guidelines", the purpose of which was to convey information about paroling

policies in the past. These guidelines, developed late in 1971 by Gottfredson, Wilkins and their colleagues, used a six-level classification of offense seriousness based on median times served for different types of crimes, and a four-category parole prognosis scale known as the "salient factor" score.[4] In each of the 24 cells resulting from cross-classifying these two scales, median times served by prisoners paroled in the preceding two years were calculated; these medians were "smoothed" on the basis of visual inspection of the matrices [5], and the resulting values were bracketed with more or less arbitrary "plus and minus" ranges to yield a matrix like the one illustrated in Table 3.1.[6]

Insert Table 3.1 here

Guidelines of this form were subsequently adopted by the U.S. Parole Board (now known as the Parole Commission), and -- with slight modifications from time to time -- have been used by it since 1975 (see, for a further discussion, Gottfredson, Wilkins and Hoffman, 1978; Hoffman and Beck, 1974). Before tracing the further history of this concept, however, a number of important points should be noted. First, the techniques used to construct these first guidelines were extremely crude (and unashamedly so): they consisted merely of identifying the two factors -- offense seriousness and risk of parole violation -- that were most strongly associated with lengths of terms, then cross-classifying construction (and, later, validation) samples by these factors, which were grouped in a very ad hoc fashion; judgementally "smoothing" the resulting cell medians; and then arbitrarily bracketing those smoothed medians to provide the Board with ranges instead of "points", i.e. single terms. All that was wanted was to provide a rough estimate of the relative weights of the two variables used, so far as this could be revealed by past practice; there is not, in any of the reports on this early project,[7] any hankering after precise models of decision-making behavior. There was, it seems, some concern, even in this earliest report, to try to promote equity in the sense of similar terms for similar offenders;[8] but it was recognized from the outset that this could not be done in any precise fashion, by using a two-dimensional matrix.

Second, the "empirical basis" of the original parole guidelines is not in fact very clear. The original (Youth Correction Board) study was based, after all, on board members' subjective perceptions of offense seriousness, risk, etc.; no effort was made to establish the grounds for these perceptions, or to test their adequacy. In the guidelines later developed for (and eventually adopted by) the U.S.

Table 3.1: Average Total Time in Months Served Before Release*
(including jail time)

Offense categories:	Salient Factor Score (Probability of Favorable Parole Outcome)			
	9-11 Very High	6-8 High	4-5 Fair	0-3 Low
A (Low severity)	6-10	8-12	10-14	12-16
B	8-12	12-16	16-20	20-25
C	12-16	16-20	20-24	24-30
D	16-20	20-26	26-32	32-38
E (Very high severity)	26-36	36-45	45-55	55-65

*U.S. Board of Parole, Pilot Regionalization Project, Guidelines for Decision-Making, adult cases (adapted from Table 1 in Gottfredson et al., 1975). (Note: ranges were not provided for the highest severity offenses, e.g. willful homicide, because of insufficient numbers of cases.)

Parole Commission, more "objective" measures of offense seriousness and risk were substituted for the original subjective measures; but it is far from clear, in fact, just why this was done. So far as we are aware, no research was ever carried out by Gottfredson, Wilkins and their colleagues to try to show that the Parole Commission's subjective assessments were in any sense wrong; the shift from subjective scores to objective factors appears to have been motivated by little other than the fact that data on the latter could be routinely obtained from case files, and a feeling that such data were likely to be more reliable and valid that subjective assessments generally are. The shift may well have come about, however, because of apparent analogies to the problem of predicting parole violation -- a problem that had, of course, had considerable attention from both Gottfredson and Wilkins in the past.[9]

Third, the intended use of the guideline matrices originally developed by Gottfredson and Wilkins is not very clear to us either -- at least so far as their articles and research reports on the early projects reveal. The requirement to give reasons for a "departure" from the guidelines (i.e. a term falling outside the stipulated range) was incorporated into the appeal procedures for hearing examiners' decisions, in the rules that effected the Parole Commission's reorganization and adoption of the guidelines. But such departures -- whatever the reasons for them -- seem to have had no special status, as grounds of appeal under the Commission's rules;[10] in retrospect, the use of (or departure from) the guidelines in individual cases seems to have been much less important to Gottfredson and Wilkins, and to the Parole Commission itself, than the statement of revealed general policy which the guidelines would provide. Indeed, this notion of "making policy explicit" figured very largely in the early work which Gottfredson, Wilkins and their colleagues did with the U.S. Parole Commission; the issue of hearing examiners' subsequent compliance with the policies in question was seldom discussed.[11]

Fourth, the problem originally tackled by Gottfredson and Wilkins was not primarily one of reducing disparity or inconsistency between a number of different, autonomous decision-makers (though this was no doubt foreseen as possible when the Parole Commission's planned decentralization of case-level decision-making came into force). The early work by Gottfredson and Wilkins does not seem to us to have been much concerned with "averaging out" variations in term-fixing of the kind which might be found to exist between different judges, even in a smallish state -- though that is precisely the problem for which sentencing guidelines have subsequently been said to be a solution (see, for example, Kress, 1980).

The data from which the original Parole Commission guidelines were derived were, in effect, data on the decisions of a single body (namely the Commission itself); what the guidelines seem to have been intended to discover, or promote, was a kind of self-consistency on the part of that body, rather than the control of a large number of decision-makers, each of whom might reasonably have been expected to be operating in a self-consistent fashion, but with different objectives or principles which (perhaps quite legitimately) governed their decisions.

Fifth, and finally, the original Gottfredson-Wilkins parole guidelines were concerned with "term-fixing", i.e. with when imprisoned offenders should be released; since all of those to whom the guidelines would be applied were already in prison, there was no need to consider the so-called "in-out" decision, i.e. the decision to incarcerate in the first place. As we shall see, however, this necessarily dichotomous decision raises a number of fundamental problems for the use of guidelines to control discretion in sentencing.[12] In summary, the context in which Gottfredson and Wilkins and their colleagues originally developed decision-making guidelines for use in the criminal justice system, and the ways in which they developed those guidelines, differed in a number of respects from the context and the developmental methods that have characterized the attempted application of guidelines to the control of sentencing, especially on a statewide basis.

Still, there are evident similarities between the decision-making done by parole boards, and that done by sentencers; and these led Gottfredson and Wilkins to wonder how far their concept of empirically-derived guidelines might serve as "a judicial tool to aid in the sentencing of offenders" (Wilkins et al., 1976:20). A "feasibility study" of this question was begun in July 1974, and completed two years later. Data on sentencing were collected at the Denver (Colorado) District Court, and in the state of Vermont; these were analyzed in much the same ways as the parole data had been, and a total of five guidelines "models" (three for Denver, two for Vermont) were constructed -- also in much the same way as had been done in the earlier parole research.

Wilkins et al.(1976) were aware that sentencing decisions were not exactly analogous to paroling decisions, for a reason already alluded to: sentencing judges must decide whether or not to incarcerate a convicted offender, as well deciding how long he should be incarcerated.[13] They have tended to express this difference in a misleading way, by saying that the sentencing decision is a two-step or "bifurcated" one (see e.g. Wilkins et al., 1976:xxi). This is misleading, because

it suggests that the two decisions are somehow psychologically different -- yet there is no evidence for this claim (or at least none ever presented by Wilkins et al.). The truth is that the two decisions are logically related, in that the second ("How long?") question cannot arise at all unless the answer to the first ("In or out?") question has been "In". Thus the question of duration applies only to a sub-set of sentenced offenders, and may well be influenced by quite different factors from the question of incarceration or not. Given their recognition of the difference between the two decisions, however, it is surprising that Wilkins et al. seem to have ignored it entirely in their analyses of the Denver and Vermont data: their regression models used a single outcome variable, which took a value of zero for those not incarcerated, and time "inside" for those incarcerated. As Rich et al. (1980) have pointed out, this procedure can lead to serious mis-estimation. As we show in a later chapter,[14] the same procedure appears to have been followed in at least one other state in which sentencing guidelines have been developed, though the consequences of that mistake are not clear.

What is of most interest, however, is the similarity which Gottfredson and Wilkins seem to have perceived between the parole problems they had been researching, and the problems of sentencing, and the assumptions with which they began their sentencing research. For one thing, as they put it in their report (Wilkins et al., 1976:21-22) they

> ...intended to be constructive and not merely critical. We hoped to provide courts with a workable sentencing information system, upgrade the quality of probation reports and help judges in their most difficult task. In short, we consciously decided to work with the judiciary in a collaborative venture, and not on, around, or against judges. Our goal from the first was to assist judges rather than to study them.

After noting their opposition to legislatively mandated sentences as being "unrealistically rigid and mechanical", and acknowledging the need for some control on discretion, they stated that

> ...we were confident that there did exist an implicit policy formulation which acted as an underpinning for judicial decision-making in the sentencing area. Through careful analysis of present practice, we believed it possible to discover that implicit policy and make it explicit, thereby allowing in the future for clearer overall policy formulations, as well as more cogent review for individualized decisions.

> We began with these premises, believe that we have checked them and found them valid, and continue to hold them (Wilkins et al., 1976, emphasis added).

This is, of course, precisely the same notion that had animated the earlier research by Gottfredson and Wilkins on parole: it is the notion that there is a policy already there, albeit an "implicit" one; and that all that is needed is to analyze past practice, whereupon that policy will be revealed. But once made "explicit", this policy can and should be used by judges -- whose policy, after all, it really is -- to guide future decisions in individual cases, as well as providing the basis for clearer formulations (or reformulations) of the policy itself.

The validity of the underlying assumption here can certainly be questioned. If all that is meant by saing "a policy exists" is that, in general, judges regard seriousness of current offense and prior record as important determinants of severity of sentences, then the statement may be accepted as true; but it does not make much of a claim nowadays (though it might have been regarded as stronger stuff in the heyday of "indeterminacy"). If it is intended in any stronger sense, however, the statement that "a policy exists" becomes highly problematic for many jurisdictions. We will return to this point below. One further aspect of the original Gottfredson-Wilkins feasibility research should be noted at this point, however. Their notion of working "with the judiciary in a collaborative venture" -- of doing "action research" like that which they had earlier done with the U.S. Board of Parole -- seems to us have obscured, perhaps inadvertently, the extent to which anyone involved in developing sentencing guidelines -- whether or not these claim to be "empirically based" -- is in fact deeply involved in the formulation of what is likely to become future sentencing policy, and is not merely carrying out some statistical analyses which may or may not be used, eventually, by someone else who is the "real" policy-maker. For example: Gottfredson and Wilkins, and their colleagues, have repeatedly characterized their guidelines as "descriptive, not prescriptive" -- the general implication being that they were merely carrying out some analyses which would reveal a policy that had already been there, all along (as was the case, in the original Youth Corrections Board study). Yet after this initial study, it seems clear to us that their intentions were perfectly clear: they wanted their findings to be translated into action, and were not merely trying to "make explicit" a policy that had been there along, and would continue to be applied (or not), in the same way in future. As their reports very often make clear, they wanted their findings to be translated into a set of sanctioning rules, in the sense

explained in the preceding chapter of this report: that is, it was apparently intended by them that judges should in the future follow a rule of the form "If X, then Y is permissible" -- where the X and Y terms were defined by their research. In the earlier publications by Gottfredson, Wilkins and their colleagues, this concern with such things as equity, justice, rationality, and fairness, is explicit: they were certainly not trying merely to explain sentencing behavior, or merely to predict it. They wanted to change it, to improve it -- to make it more just.[15]

The ways in which they tried to accomplish this end seem, in retrospect, to have carried with them some dangers. As Rich et al. (1980) have noted, the "descriptive, not prescriptive" notion has had a large place in the rhetoric used, over the past decade, to describe sentencing guidelines -- though, as we shall see in the next section, that rhetoric mostly did not come from either Gottfredson or Wilkins. To put the matter more strongly: the catch-phrase "descriptive, not prescriptive" has been used to try to sell the concept of sentencing guidelines, especially to the judiciary. Yet, as we shall show, the concept of sentencing guidelines is inherently prescriptive, in the sense that such guidelines are intended to serve as sanctioning rules governing decision-making in the future; if this were not the case, the analysis of past sentencing practice could have been no more than an academic exercise, of a kind in which Gottfredson and Wilkins definitely were not engaged.[16]

Vulgarization of the Concept of Guidelines

As we have seen, the study done by Gottfredson and Wilkins in Denver and Vermont concluded that it was both "feasible" and "desirable" to structure sentencing discretion by means of empirically-based guidelines (Wilkins et al.,1976:xx). These "findings" were swiftly translated into social action. A further grant (number 76NI-99-0102) was made by NILECJ to the Criminal Justice Research Center at the State University of New York at Albany (CJRC) to support the implementation of sentencing guidelines in a number of jurisdictions. In the event, four county-level court systems were involved in this phase of the research. One was the Denver District Court, which had taken part in the original feasibility study; another was the Essex County Court in Newark, New Jersey, which had had an "observer" role in the feasibility study. The other two were the Cook County Circuit Court in Chicago, and the Maricopa County Superior Court in Phoenix. In addition, research staff of the guidelines project gave technical assistance to the Philadelphia Court of Common Pleas, which was independently involved in developing sentencing guidelines. In all five of these jurisdictions,

sentencing guidelines were used to some extent for varying periods in the years 1977-79, though the Chicago, Denver and Phoenix guidelines were soon superseded by new state penal codes, and the Newark guidelines were made obsolete by the statewide guidelines introduced in New Jersey in 1978.[17]

At the time of the second (implementation) grant to CJRC, however, a change in the dramatis personae occurred: Gottfredson and Wilkins ceased to be involved with the sentencing guidelines project.[18] The principal investigator on the implementation grant was Jack M. Kress, who had been co-director (with Gottfredson and Wilkins) of the feasibility study, and who at that time (1976) was a member of the faculty of the SUNY-Albany School of Criminal Justice.[19] Also involved in the implementation study were Arthur M. Gelman and Joseph C. Calpin, who had been members of the project staff during the feasibility study. Subsequent reports on sentencing guidelines show Kress, Gelman and Calpin as authors, in varying orders.

Concurrently, NILECJ had made a contract (number J-LEAA-022-76) with the University Research Corporation (URC) in Washington, D.C., to "design, coordinate and deliver training workshops" on selected criminal justice topics designated by NILECJ, as part of the Institute's Executive Training Program in Advanced Criminal Justice Practices.[20] After receiving the final report on the Gottfredson-Wilkins feasibility study, the Director of NILECJ[21] had decided that

> From the judicial and public awareness of the importance of even handedness in sentencing, the feasibility demonstrated by recent research, and the need of State Planning Agencies....it is concluded that Developing Sentencing Guidelines should be in the 1977 Executive Training Program. (Kress et al., n.d.:7).

After an initial meeting with NILECJ staff, a planning conference (attended by "national experts") was held in Silver Spring, Maryland, on the subject of developing sentencing guidelines workshops; several such workshops were held during 1977 and 1978. Kress, Gelman and Calpin were substantially involved in this exercise in dissemination; they are listed as co-authors of some materials prepared under the URC contract and published by NILECJ, they participated in workshops, and they also acted as consultants to a number of researchers involved, in various states, in research aimed at developing sentencing guidelines. In particular, Kress made at least one presentation to a group that included representatives of the Massachusetts judiciary, at a time when sentencing reform was under consideration in that state; while this may not have been the decisive factor which led judges to support the

development of sentencing guidelines in that state, it almost certainly had some impact.[22] Kress and Gelman also made visits to Minnesota and Michigan, respectively, at the time when the development of guidelines was under consideration in those states (see Parent, 1979; Zalman et al., 1979:xii).[23]

It is not our intention to try to evaluate, in exact terms, the impact of the dissemination program carried out under the second (implementation) grant to CJRC and under the URC contract. In addition to the workshops and consultation just mentioned, the tangible products of this dissemination effort included a do-it-yourself "methods manual" (Gelman, Kress and Calpin, 1977) and a made-for-television movie on "Developing Sentencing Guidelines" starring -- if that is the right word -- Gelman. We do not know what proportions of their intended audiences were actually exposed to these products; but what is important for present purposes is the concept of sentencing guidelines that was promoted by Kress and his colleagues during this period, and the kind of sentencing reform that was envisaged; that is a scientific issue, and not merely an aesthetic one.

Consider first the "trainer's handbook" developed for the URC workshops (Kress et al., n.d.). In addition to the planning conference mentioned earlier, these workshops were apparently stimulated by responses to a series of questions about sentencing and guidelines, from "40 judges from across the nation" (Kress et al., n.d.:8). It was found that "less than one percent"[24] of these judges were familiar with sentencing guidelines, but that 98 percent wrote comments interpreted as being in favor of considering guidelines, including "If I don't (help develop sentencing guidelines), then I'll be stuck with something I didn't have a voice in" -- a sentiment which we found, during our research, to be expressed by judges from a number of different states.[25] From this kind of evidence it was concluded that the CJRC feasibility study "had not only proved sentencing guidelines to be a feasible instrument, but should (sic) serve as the central core of a training workshop on sentencing" (Kress et al., n.d.:9, emphasis added).

The workshop manual provided for a total of 14 sessions. In the "presentation outline" for the second of these, it is stated that guidelines "provide a narrow sentence range based upon the actual sentences imposed by fellow judges in similar cases involving similar offenders"; it is also stated that "these sentences are not made up by researchers, but are statistically established after analysis of thousands of sentencing decisions within that particular jurisdiction" (Kress et al., n.d.:21). At the end of this session attention is given (under the heading "Value of sentencing guidelines")

to "Need for the judiciary to act on its own to help ensure equitable sentencing", with the further elliptical caution that "Alternative is for legislature to do it through mandatory type sentencing proposals" (Kress et al., n.d.:22). New legislation in California, Maine and Indiana was referred to in the next session.[26]

At a simple factual level, these materials are grossly misleading. The "narrow sentence range" to be provided by sentencing guidelines was not, in any of the research to which the statement could have referred, "based upon the actual sentences imposed by fellow judges in thousands of similar cases". Indeed, as we show below, the width of guideline ranges is necessarily judgemental if not arbitrary: there is no purely statistical technique by which optimum ranges can in general be derived. It is also absurd even to imply -- much less to say, in plain English -- that the CJRC guidelines were based on "analysis of thousands of sentencing decisions". The construction samples in the Denver and Vermont feasibility research contained total N's of 200 each; the validation samples contained 221 and 113 cases, respectively. In their report, Wilkins et al. (1976:13, 58, 85) state that, because of missing data, the Denver guidelines were actually constructed from data on only 120 cases. In their reanalysis of the Denver data, however, Rich et al. (1980:67) state that "because of frequent missing observations, the Denver guidelines models could not have been estimated by more than 50 cases". (Nearly the same conclusion was reached by Hewitt and Little (1981) in their reanalysis of the Denver data.) It is true that in their do-it-yourself manual, Gelman, Kress and Calpin (1979:11) suggest that "a sample of 1,000 to 4,000 sentencing decisions would appear to be adequate" in most jurisdictions; and samples in the range of 1,000 to 2,000 were used in some of their later research. But to imply -- as the written training manual, at any rate, clearly does -- that the original CJRC guidelines had been based on samples of that size is disingenuous at best. The later references to the "need for judiciary to act on its own", with the threatened alternative of "mandatory type sentencing bills" may seem merely like shabby hucksterism, given that the workshops were aimed at judges; but at least they are approximately honest.

In subsequent workshop sessions, participant judges were to have chosen sentences appropriate for three hypothetical cases, using "decision-game" techniques of the kind developed by Wilkins (Wilkins and Chandler, 1965). The materials pertaining to these sessions contain some assertions with which one might argue;[27] but they do not seem on their face to be obviously misleading. The later sessions on developing and using guidelines, however, appear to be based entirely on the original Denver and Vermont sentencing guidelines, as

developed by Wilkins et al. (1976). It thus seems to be assumed that the factors used in those guidelines, and the weights attached to those factors, will be equally applicable to the jurisdictions of workshop participants; this in turn implies that those factors, and those weights, are likely to emerge empirically, from analyses of data from workshop participants' jurisdictions. This is not, of course, necessarily true.

In the twelfth session of the workshops, the process of developing sentencing guidelines (collecting data, etc.) was to be described; the presentation outline for this session is brief, though there are references to the manual developed by Gelman, Kress and Calpin (1979), which is discussed in more detail below. Presumably these moderately technical issues were not thought likely to be of much interest to judges. Participants were, however, told in session 13 to consider the feasibility of implementing guidelines in their own jurisdictions; they were also given a list of "some steps which could be taken toward implementing sentencing guidelines" -- the last of which is "Keep record of principal opponents of guidelines and reasons for opposition". Immediately after this recommendation for an "enemies list", the trainer's handbook provides an outline for a grant proposal to be titled "A Proposal for the Development of Sentencing Guidelines", which was presumably to be used to obtain federal or other funds.

It may be argued that it is unfair to evaluate these workshops on the basis of the trainer's handbook, which is generally only in outline form; the workshop presentations themselves may well have been more complete and detailed. Nonetheless, reading the manual often gives the impression of an account of the concept of guidelines that is at best simplistic (and at worst -- as with the reference to "thousands of cases" discussed above -- misleading). For instance, in session 10 on "types of guidelines", participants were apparently told that "the computation of guidelines sentences can be accomplished by almost anyone in the court system with relatively little training", in an estimated four to six minutes per case; this may be true for some jurisdictions or for some types of guidelines, but (in our experience) it is not true for others.

This impression of conceptual simple-mindedness is confirmed by a reading of the "training manual" (Gelman, Kress and Calpin, 1979) which "describes for researchers the specific procedures to be followed in constructing a sentencing guidelines system."[28] Much of this manual consists of rather brief instructions, at approximately the level of an undergraduate textbook in research methods,

concerning such things as designing a coding form, defining missing values, "cleaning" data and the like; sample coding instructions and a computer program are included.[29] What is of interest, for present purposes, is the concept of sentencing guidelines which emerges from this set of cookbook-like injunctions (e.g. "Having collected the sample, the data must be keypunched and verified. Verification will help minimize, but not eliminate, keypunch errors": p. 12, emphasis in original). In brief, the reader is told to carry out the following steps, in order: First, review univariate frequency distributions of all variables on which data have been collected; then cross-tabulate independent or predictor variables with the dependent variable, "best classified" as 'in' or 'out'; correlation coefficients such as Pearson's r are also suggested at this stage, with data reduction being based on a rough cutoff point of $r=\pm.20$ at a significance level of .005, and missing data being handled by pairwise deletion in the first instance.[30] Multivariate analyses -- including multiple regression and discriminant function analysis ("which can be expected to produce results similar to those provided by multiple regression analysis when the dependent variable is dichotomized"[31]) are the next step. The same statistical analysis techniques are then to be repeated, with "length of incarceration" as the dependent variable. The results of these last analyses will not be "used directly" in the development of sentencing grids, but it is said that they "will provide some additional insight into sentencing practices"; it is noted that "certain variables may be identified whose effects on the sentencing decision are limited to the 'how long' as opposed to the 'in/out' question."[32]

The next step is described as "model choice". After having completed correlational and multivariate analyses, the researcher is supposed to have a good idea of the "10 to 20" variables which seem to have the potential for most accurately predicting sentencing decisions; he is then to generate "models", determining what combinations of those items most accurately predict sentences. "General", "generic" and "crime-specific" models are briefly described, as is a "bifurcated" model using one set of items to predict the decision to incarcerate, and another set to predict length of incarceration. The next step, called "model development", is said to involve

> ...the process of incorporating the ten to twenty items into various combinations of sentencing guidelines models. Assuming that a two-dimensional model has been decided upon, two separate scales must be developed, one focusing on the crime or offense and the other on the offender. Different combinations of variables can be

> used to develop each scale. The range of scores on each scale determines the number of cells within a grid, although identical sentences in contiguous cells may eventually lead to a merging of columns or rows. (p. 17, emphasis added.)

It is noted by the authors that one of the findings of earlier guidelines research was that judges consider the "real" offense, rather than the offense of conviction, which may reflect plea-bargaining; it is noted that there may be a problem of assessing the seriousness of offenses, and it is suggested that "intra-class ranking" by judges is one solution to this problem, though it is also noted that a "seriousness modifier" reflecting such things as use of weapon, degree of injury, etc., may be needed, to help in developing a model with "increased predictive power". Then comes this paragraph:

> There are no rules as to which variables are to be used in the development of guideline models or what specific weights are to be assigned to them. It is an iterative process of testing, modification and retesting. However, the predictive ability of each individual variable can be checked through a Mean Cost Rating or the Index of Predictive Efficiency (p. 18).

It is then recommended that the "models" thus developed be applied to a separate validation sample, using a sample "at least one-third and hopefully one-half to two-thirds the size of the construction sample". Cases are then to be cross-classified in the cells of the matrix or matrices resulting from this "development" exercise; an "in/out" line is to be drawn through the matrix, by inspection of the modal category of sentences in each cell. The median is recommended as a measure of length of incarceration, with a "small" range, e.g. ±12.5 percent, marking off departure cases; it is noted, however, that final decisions on these last two points "must be made by judges in their role as a policy body" (p. 22). There is a further brief discussion of presentation of the resulting guidelines to judges, implementation, and feedback and review sessions.

This manual seems to us to be open to criticism on a number of grounds. We do not refer here to its often simplistic notions of data analysis;[33] nor to its occasionally baffling statements about weighting,[34] the process of validation,[35] or methods of conceptualizing and measuring the seriousness of offenses.[36] These technical lapses -- or perhaps they are just stylistic infelicities -- could and probably would be patched up, by any competent social researcher. A much more serious criticism of the do-it-yourself recipes in this manual is that they debase

completely the concept of sentencing guidelines as a species of sanctioning rule --that is, as an instrument which is ultimately intended to guide or even control judicial decisions on sentencing in the future. As a result, most if not all of the fundamental issues of social policy and morality -- that is, issues of what is just or expedient for sentencing in the future, as distinct from what may be revealed by some necessarily cursory analysis of data about sentencing in the past -- are almost completely ignored. True, there are occasional references to questions which "judges in their role as a policy body" must decide -- such as the appropriate measure of the central tendency of length of incarceration, and the use (presumably this should be the width) of durational ranges.[37] But these questions are, by and large, trivial; on several much more important issues, considerations of policy are either distorted or ignored, in favor of the Index of Predictive Efficiency.[38] Some examples of this insensitivity are as follows:

(1) It is several times stated, and other times implied, that the sole criterion for inclusion of variables in sentencing "models" or guidelines is their contribution to predictive accuracy: at one point it is suggested that there may be "10 to 20" variables that have sufficient potential for this purpose.[39] Nowhere in this or any other part of this manual is it suggested that it matters what these variables are; all that is said to matter is their predictive accuracy. But what if it should happen, for example, that some morally iniquitous variable such as race or ethnicity were shown to be the strongest predictor of sentences in jurisdiction X in the past? What if a morally dubious variable -- sex is perhaps one example at the present day, early childhood environment another -- were shown to predict? What if the strongest predictor of sentences in the past were something totally nonsensical, such as length of big toe or the possession of red hair?[40] Nothing in Gelman, Kress and Calpin (1979) even allows for the scrutiny of such variables, let alone for their exclusion.

It may be said in the authors' defense that the prior research in which they were involved had not shown that variables such as those we have just mentioned were much predictive of sentences. That fact -- if it is indeed a fact -- may redeem one's faith in the judiciary of Denver, Philadelphia, Vermont, Chicago, Phoenix and Newark, New Jersey. (Yet while morally iniquitous variables may be superficially missing from the guidelines developed by those bodies, the effects of such variables on less controversial included items were never estimated.) But there are other variables which as well may be associated with, and even predictive of, sentences in the past, which are not so easily

ruled out. For instance, at one point Gelman, Kress and Calpin (1979:17) blithely inform their readers that "judges consider the 'real offense' in deciding what sentence to impose." There are, indeed, well-known legal and ethical arguments in favor of judges' doing this.[41] But there are also strong arguments against their doing it; and, as we shall see in a later chapter of this report,[42] the use of the "real offense" as opposed to the offense of conviction was explicitly ruled out by the Minnesota Sentencing Commission in constructing that state's guidelines. Yet on the purely predictive approach advocated in the do-it-yourself manual, debate over the propriety of this course of action could not even arise, if "real offense" were strongly predictive.

(2) A related point is that Gelman, Kress and Calpin (1979:17) simply assume that a "two-dimensional model" has been "decided upon" by those involved in developing guidelines. Exactly why they should assume this -- apart from the fact that this was the kind of model with which they were most familiar -- is not clear. (Since they had, just previously, referred to "generic", "crime-specific" and "bifurcated" models, none of which is two-dimensional, there may be some doubt about their views on this point; their assumption of a two-dimensional model for the rest of their exposition, however, is tolerably clear.) Yet why should a two-dimensional set of guidelines be anywhere near optimal -- even if, as is said elsewhere in the manual, as many as "10 to 20" variables may have important power for predicting past sentences? In other words: if reproduction of the (predicted) past is the name of the game, why should this exercise be handicapped by the choice of a two-factor decision rule -- unless, of course, it is assumed that all of those factors that can appropriately be included in guidelines must be subsumable under the two broad headings of "factors relating to the current offense" and "factors relating to the offender" and can be most predictive when combined in a additive model? Here is one of several instances in which what are patently questions of social policy seem to be smuggled into what is presented as a purely predictive exercise.

(3) When it comes to drawing the "in/out" line, the crude empirical approach of Gelman, Kress and Calpin (1977) becomes even more obviously absurd. The designation of a particular cell (in the two-dimensional matrix already assumed to have been decided upon) as "in" rather than "out" is to be "determined by the modal category of the sentences within that cell, and by analysis of contiguous cells" (p. 19, emphasis added). What if the mode includes only 55 percent of the cases in the cell? What if it includes only 51 percent? In these instances the guidelines prescribed decision would only accurately "describe" a slight majority of cases, yet this problem is not considered at all.

It is no answer to these questions to say (as the authors do say) that "the predictive line (sic) is drawn so as to minimize errors in predicting the 'in/out' decision". For it is perfectly possible, depending on the distribution of cases within cells, that an "in/out" line that maximizes predictive accuracy in that sense over the whole table will entail "out" sentences for cells in which the modal numbers of sentences (in construction and or validation samples) were "in" sentences. The authors of the manual state that "the logic of the guideline concept demands that, as the offense becomes more serious and/or the offender's unfavorable characteristics become more pronounced, the probability of incarceration and the length of that incarceration should increase" (1979:19). Where exactly this "logic" is supposed to come from is not clear; we will return to this point below. For the moment, the point to note is that this "logic" may well not emerge from a statistical analysis of past sentencing practice in this or that particular jurisdiction. What is the neophyte guideline constructor to do, if it does not? Is he to follow past practice, designating cells as (presumptively) "in" or "out" according to what happened in the past? Is he to engage in a "smoothing" exercise of the kind that Gottfredson et al. (1975) engaged in with the original parole guidelines? Or is he to follow the "logic" of the guidelines concept -- assuming that it can be figured out exactly what this is?

(4) After "development", it is expected that the researcher will have identified "five or six 'best' models in terms of their predictive power" -- though no clear criterion of predictive power is mentioned. These five or six models are then to be tested -- in a none-too-rigorous fashion[43] -- on a validation sample; and the "two or three best predictive models" are then to be presented to the judges "sitting en banc as a policy decision-making body". The researchers are instructed to "make it clear that their role so far has been empirical, that is, to describe the court's current sentencing practices" (Gelman, Kress and Calpin, 1977:21-22). But will they necessarily have done this, in fact? Suppose that the "two or three best predictive models" are very different, but about equal in predictive efficiency (however measured)? Which one best describes past practice? Suppose further that, for example, race of offender is a strong predictor of sentence, perhaps because it is accounting for variance in dispositions that is not captured by imperfectly-measured prior-record variables? In such a case, race might be a good predictor; but it would not follow that previous sentencing practice in the jurisdiction in question had been racially biased. Indeed, none of the research procedures described in the do-it-yourself manual seems aimed at discovering ("making explicit") previous sentencing policy, that is, at uncovering the grounds on which previous sentences were based and

supposedly justified. (A possible exception is the use of judicial rankings as a measure of seriousness of offense; but this is said on p. 17 of the manual to be only one method of handling this problem.)

(5) Finally, while it is admitted that the issue of range width (around median terms in matrix cells) is an issue of policy to be resolved by the judges, and it is (virtually) admitted that "plus or minus 12.5 percent" is an arbitrarily-chosen figure, no consideration is given, anywhere in the manual, to the nature of the cases in the construction and validation samples which fall outside the chosen range. Here, surely, is something which an empirical analysis of past sentencing practice can easily and usefully provide: namely, some indication of the kinds of cases that had, in the past, resulted in extreme deviations from general sentencing practice, e.g. in receiving very long or very short prison terms. Technically, this is a relatively simple matter; it involves an examination of regression residuals, and identification of any extreme outliers.[44] Yet no mention of this is made in the manual; indeed, there is virtually no consideration given to skewness or other aspects of distributional shape.[45] Finally, while the two or three final models are to be chosen on the basis of their efficiency at predicting the "in/out" decision, no consideration is given to the cases in which the "best" predictions are wrong, e.g. to the cases in which the models predict incarceration but an "out" sentence was in fact given. Was there something about these cases -- some legitimate but rarely-present factor, for instance -- that made the predicted disposition genuinely inappropriate? The researcher who follows this manual will not be advised even to ask this question.

One possible explanation for such failures in prediction, of course, is that they reflect the idiosyncratic practice of one or more judges, who sentence in very different ways from the rest of their colleagues. The manual says nothing whatever about this possibility -- no doubt because the resulting "models" were intended to be peddled to the judiciary, rather than to, say, a legislature or sentencing commission.

In summary, the concept of sentencing guidelines which emerges from the materials shows only minimal sensitivity to the issues of morality and social policy inherent in the development and implementation of this technique for controlling judicial discretion and bringing about a change in sentencing practice. The materials contain, for example, no discussion of the aims of sentencing, and the possible ways in which these may affect, and be affected by, sentencing guidelines. Nothing is said about the relevance to those aims

of the factors contained in the guidelines "models" or the guidelines themselves; nor is any consideration given to the justice or morality of factors from which guidelines might be constructed. (For example, there is no consideration in the manual of the question whether it is appropriate to include "social stability" measures (to say nothing of prior record measures) which may be race-linked or class-linked.) The concern with promoting equity and justice which characterized the report on the feasibility study (Wilkins et al., 1976) has disappeared completely; developing sentencing guidelines has become a mere exercise in research technique, judged solely by predictive accuracy. Finally, where empirical research could contribute to the development of adequate guidelines -- for example, in identifying sources of variation in prior sentencing practice, and trying to identify reasons for extreme outliers -- the manual says nothing at all.

In fairness, it must be pointed out that in his latest writings on this subject, Kress (1980) takes a very different approach to the subject. His aim in this book is avowedly at achieving a rational and just sentencing system; he notes that this must inevitably involve explicit consideration of normative elements, and argues that one function of an empirical analysis of sentencing practice "serves to open it up for clear discussion and cogent review" (Kress, 1980:227). He also advocates that guidelines be developed by a sentencing commission, rather than on behalf of the judiciary themselves. Given the emphasis which Kress places on the normative or prescriptive elements inherent in guidelines, it may be wondered why he continues to place so much emphasis on statistical analyses of past practice -- which, in our reading of his latest work, seem almost irrelevant to the "prescription for justice" which he advocates there. However that may be, the contrast between the dissemination materials and Kress's latest work is very striking. (For another very different approach to the construction of decision-making guidelines, see Gottfredson and Wilkins, 1978, esp. 285.)

As we noted earlier, the evaluation of the dissemination grant was not an objective of our research. As we show in later chapters, none of the statewide guidelines developed to date seems directly to have followed the procedures specified in Gelman, Kress and Calpin (1979) or in Kress et al. (n.d.). In at least three of the five states (Massachusetts, Michigan, and Minnesota), the researchers involved in developing guidelines were certainly aware of the work of Kress and his colleagues, however; at least, all three groups possessed copies of Gelman, Kress and Calpin (1979) which we saw in their offices.[46]

Nor are we concerned to engage in criticism of the "methods manual" just for the sake of criticism -- if only because we are unsure whether such criticism should be aimed primarily at those who wrote the manual, or those who funded and disseminated it. Our point is that the concept of decision-making guidelines has been distorted, in some subtle and not-so-subtle ways, by many of those who have written about that concept; the work of Gelman, Kress and Calpin illustrates many of those distortions, and may be the source of most of them. It is vital to clear these distortions out of the way, if we are to understand the strengths and limitations of this technique for controlling discretion, and to assess the ways in which the technique has been applied in statewide jurisdictions to date.

Guidelines as a Species of Sanctioning Rule

At this point, let us skip ahead somewhat, ignoring the problems involved in constructing sentencing guidelines in order to consider the ways in which guidelines may be used to control judicial discretion in sentencing practice. For this purpose we make use of the concept of sanctioning rules developed at the end of the preceding chapter: that is, rules addressed to judges (or other decision-makers), which take the general form "If X then Y", where X is some set of factual circumstances and Y is a sanction or disjunctive set of sanctions which may be applied if X is found. Viewing guidelines as a species of sanctioning rule can be a clarifying idea in two ways. First, it can facilitate comparison both with other techniques for controlling discretionary decision-making and with the virtually unfettered discretion that is characteristic of "indeterminate" sentencing systems. Second, it can help to show exactly how guidelines "work": that is, it can illuminate the features of guidelines -- in the form originally proposed by Gottfredson and Wilkins, but also as exemplified in the statewide systems studied by us, and described in later chapters -- which operate to constrain decision-makers' behavior when dealing with particular cases.

For this purpose we begin by considering the use of guidelines to prescribe the length of incarceration of offenders sent to jail or prison. As already noted, this "term-fixing" decision may be made either by a sentencing judge or by a parole board or other administrative agency;[47] and it is the kind of decision for which guidelines were originally developed. In addition, as we shall show in a later section, the dichotomous "in-out" decision poses some special problems where guidelines are concerned.

At first sight, guidelines prescribing lengths of term may seem typically to take a very simple form: "If x1 and x2, then y1→y2 is permissible", where x1 characterizes the offense in question, x2 characterizes the offender's prior record, and y1 and y2 indicate the end-points of the permissible range of length of incarceration; the symbol "→" is used to indicate that the permissible sanction lies within a range, rather than being a set of points.[48] The descriptors x1 and x2 can be in words; alternatively, as in most of the guidelines developed by Gottfredson, Wilkins and their colleagues, they can be numerical scores, calculated according to formulae contained in separate rules (which are not addressed to the decision-maker in the same way that sanctioning rules are).[49] The situation is actually a little more complicated than this, however, since sentencing guidelines typically provide that it is permissible for the judge to "depart from" the prescribed range y1→y2 in certain circumstances. It seems, therefore, that the sanctioning rule embodied in guidelines must be written in the following way: "If x1 and x2 and (nothing special), then y1→y2 is permissible", or alternatively, "If x1 and x2 and NOT (x3, x4,...xk) then y1→y2 is permissible", where the bracketed set of terms (x3, x4,...xk) is conjunctive, so that none of those things can be present.[50] The first of these -- with the (nothing special)" term -- represents the Gottfredson-Wilkins formulation of guidelines, since they did not include an explicit list of factors that would justify departure from the prescribed range. But it is perfectly possible to include a list of such factors; the guidelines developed in Minnesota and Pennsylvania, for example, do something of this sort.[51]

This may seem like a pretty unsatisfactory situation. What sort of a rule can it be, that says in effect "Do such-and-such, if (nothing special) is present in the case"? Dissatisfaction may be heightened, when it is realized that even very explicit lists of the form (x3, x4,...xk) tend, almost invariably, to be "open-ended"; that is, they tend to end with an "et cetera" clause, or to have an xk that says in effect "anything else like that". In fact, however, this kind of "open-endedness" is not something inherent in sanctioning rules in general, or guidelines in particular; instead, it is a feature of language in general. The X terms in sanctioning rules are descriptive; they set out possible states of affairs which, if present in a particular case, make permissible the sanction Y. Now, any time words are applied to the world, there is some "open-endedness" potentially present; inevitably there can be borderline cases in which it is open to argument whether or not a particular descriptive word or phrase applies. This is so, even for very simple cases. We all know, for example, what the word "cow" applies to. But there is no set of logically necessary and sufficient conditions by

which the word "cow" can be defined; a cow does not need to be alive or to have four legs, for example, since neither a dead cow nor a three-legged cow constitutes a contradiction in terms. Of course, if the beast before us is outrageously different from the standard or paradigm case of a cow, then we may refuse to apply the word "cow" to it. (If, for example, it is made out of metal and plastic, has slots for coins between its horns, and emits rock music when coins are inserted and its tail is pulled, then it probably is not a cow, but is rather a juke box designed to look a little like a cow.) So it is with the X terms in sanctioning rules; they too have an inherent penumbra of vagueness. Of course we may make up further rules designed to settle borderline cases; but since rules logically cannot determine their own application, they too will contain an inherent element of "open-endedness". Lawyers have developed a variety of techniques for rationally settling questions of this kind.[52]

Nonetheless, it is important to note that there is a substantial difference between a sanctioning rule containing merely a "(nothing special)" clause, and one containing a explicit list of the form "(x3, x4,...xk)" -- even if this list contains an "et cetera" clause, and is explicitly said to be "nonexhaustive" -- as is the case, for example, with the Minnesota guidelines.[53] Rules of the first form -- which represent the Gottfredson-Wilkins conception of guidelines -- say nothing about the kinds of factors which may serve as justifiable grounds for departure from the prescribed range. Of course, it may be felt that there is a fair moral consensus among those to whom the rule is addressed, so that no decision-maker would be likely to try to justify a departure by citing a morally iniquitous or nonsensical factor of a case as "something special". Still, the sanctioning rule representing Gottfredson-Wilkins guidelines says nothing on its face about this; a judge dealing with a particular case thus could cite an irrelevant or nonsensical fact, without actually violating the sanctioning rule. This is not so clearly the case with guidelines which provide an explicit list of permitted grounds for departure. Though such lists cannot be absolutely exhaustive, even if they do not contain an "et cetera" clause -- because of the "open-ended" feature of language discussed above -- they still place some constraint on permissible grounds for departure. The guidelines -- and thus the sanctioning rules which they entail -- may also specify factors which may not be used to justify departures from the specified sanctioning range; this is the case, for example, with the guidelines developed in both Minnesota and Pennsylvania.[54]

Under the original Gottfredson-Wilkins concept of guidelines, and in most of the guidelines developed in

different states to date, there is no formal limitation on the sentence which the judge may impose *if* he believes that a departure from the prescribed range is justified; he may do nothing at all, or impose any available measure including imprisonment up to the statutory maximum. But there is, of course, no reason why some such formal limitation should not be provided, so that even if the sentencing judge decides that a term outside the normally prescribed range is justified, he is still limited in (for example) the *extent* to which he can go outside that range. This is the case, in fact, in the guidelines developed (but not yet implemented) in Pennsylvania: "departure" sentences are in general limited to a term provided in one of the adjacent cells of the guidelines matrix.[55] The easiest way to represent this situation analytically is to assume that such a limitation involved the existence of a *further* sanctioning rule, in addition to the one stipulating the basic guidelines range.[56] Finally, there may also be further rules whose purpose is to limit discretion even within the normal range; in Minnesota, for example, the guidelines appear to provide that the mid-point of the stipulated range of incarceration is presumptively to be imposed in the typical case.[57]

Sentencing, then, is an activity that may take place in the context of a *set* of what we have called sanctioning rules; these rules are directed to the sentencing judge, and it is intended that he should follow them (or, at a minimum, that his behavior should not contravene them).[58] The rules contain descriptions of antecedent conditions X, which may be of various kinds; for each such set of conditions, sanctions of various kinds are then made permissible or mandatory. The outer limits prescribed by this set of rules are those contained in statutes, which in most jurisdictions provide only maximum terms for various offenses, leaving judges free to choose any sentence not exceeding that maximum. If there are sentencing guidelines, these provide narrower and more definite prescriptions, within the statutory maxima; such guidelines would constitute a second set of sanctioning rules directed to the sentencing judge, narrowing the discretion which the legislative rules conferred on him. We might also suppose that an individual bench or judge would devise still more detailed rules for further controlling discretion, within the "space" provided by the guidelines. Evidently there is a hierarchy of authority associated with these different sets of rules, so that, for example, the guidelines could not permit or require sentences which exceeded what was allowed by statute. It should be noted, however, that any of these sets of rules may provide that more than one sanctioning rule may apply to certain kinds of cases, leaving the choice between alternative *rules* (not merely alternative sanctions specified by the same rule) to the discretion of the sentencer. For

example, in all Anglo-American jurisdictions statutes provide for maximum terms of imprisonment which can be imposed for different offenses; in many of those jurisdictions there are also rules providing for terms of imprisonment for different types of persistent offenders, e.g. "three-time losers". In some cases these latter rules (which usually have the avowed aim of incapacitating potential recidivists) may permit prison terms that are much longer than the maximum punishment for the offender's most recent offense, leaving it up to the judge to decide which of the two sanctioning rules he should apply.[59]

We can now more or less formally compare, in terms of their sanctioning rules, sentencing guidelines and some of the other techniques for controlling discretion in sentencing which were discussed in the preceding chapter. At one extreme, consider the situation in which there are no formal constraints of any kind on the judge's discretion, except for the statutory maximum: in this situation, the sanctioning rules take the form "If (x1, x2,...xk), then $y_0 \rightarrow y_\mu$ is permissible", where (x1, x2,...xk) is some set of antecedent conditions, and $y_0 \rightarrow y_\mu$ designates the whole range of legally permissible sentences, from the minimum y_0 -- which may be doing nothing at all -- up to the statutory maximum term of imprisonment y_μ. In this situation, the judge's discretion is indeed unlimited apart from the statutory maximum, since there are no further sanctioning rules which he need consider and apply. He may of course have his own individual rules or policies, and may conscientiously try to follow these in particular cases; he may even succeed, more or less, in doing this correctly and consistently in most cases. But just because different judges may have different "personal rules" of this kind, it may well be that cases fulfilling the same initial conditions X would be dealt with in very different ways.

Next, consider a situation in which the law formally leaves a lot of discretion to the individual sentencer, but at the same time provides an elaborate and complex set of rules intended to govern the use of that discretion: examples would include the penal codes drafted by Ferri (1921) and Glueck (1928), which were discussed in the preceding chapter. These codes can be conceived of an sanctioning rules containing highly complicated disjunctive sets of antecedent conditions X -- in Ferri's code, these were the indicia of temibilita or "dangerousness"; to these there are attached various sanctions, depending on whether the offender is thought to need "treatment", incapacitation, or a mere reprimand. The X's and Y's in such rules may be more or less precisely defined, and the connections between them either permissive or mandatory; the result is that while the rules seem to constrain discretion to a very high degree, the individual

decision-maker still in fact has quite a lot of discretion in rule-application, both in fitting the X's to the facts of the case before him, and in choosing among the Y's which may apply to what he takes the combination of X's in that case to be. It would indeed be difficult (though by no means impossible) even to write down the sanctioning rules created by such a code, in simple sentences of the "If X, then Y" form.

At the other extreme, discretion may be curtailed entirely, by making sentences mandatory: in this case, the sanctioning rule makes a penalty Y required, and not merely permissible, if certain conditions X are found to be present in a particular case. Here there may still be some room for one kind of discretion in rule-application, namely the discretion involved in applying the description of offense and/or offender contained in the antecedent conditions X. For example, a statute provides for a mandatory jail or prison term for those found guilty of possession of an unlicensed firearm. But is one guilty of such "possession" if he has found a gun, is driving to the police station to turn it in, and is stopped for speeding on the way? [60] The sanctioning rule, however, allows the sentencer no further choice among penalties once this question is settled: <u>if</u> X, then Y is <u>required</u>.

In between these two extremes -- intricately rule-governed discretion on the one hand, and the abolition of discretion on the other -- lie a number of other techniques for controlling discretion, including sentencing guidelines. As we have noted, where length of incarceration is concerned, the sanctioning rule for guidelines takes the general form "If x1 and x2, then y1→y2 is permissible"; this narrows discretion, at least in the normal (or "nothing special") case, since the range y1→y2 is narrower than the range $y_0 \rightarrow y_\mu$ defined by the maximum penalty. It is obvious that, all other things being equal, guidelines will introduce more constraint on discretion, the narrower the range y1→y2 is. When it is narrowed to the vanishing point, so to speak -- that is, when the sanction Y is a single term of incarceration rather than a range -- then the rule specifies what is usually called <u>presumptive</u> sentencing; this is exemplified by California's Uniform Determinate Sentencing Act of 1976, which prescribes a single term -- e.g. two years -- for the normal case, instead of allowing choice within a range. [61]

There are, of course, other techniques avowedly aimed at controlling judges' discretion in sentencing. Sentencing councils and sentencing institutes, discussed in the last chapter, are two such techniques; the set of sentencing principles enunciated by the English Court of Criminal Appeal (Thomas, 1970) represents another. These techniques have

taken different forms, in different times and places; and it is sometimes unclear just how much control over individual decision-makers' behavior they are supposed to exert. The answer to this question, we suggest, is to be found by asking how far such techniques actually have the effect of creating further sanctioning rules (as the English Court of Criminal Appeal's principles clearly do); if this does not happen, as we suspect it does not with many "sentencing institutes", then the supposed technique is mere window-dressing.

Of what does the element of control consist, in the case of sentencing guidelines? First, and most obviously, guidelines (for length of incarceration) prescribe a narrower range of permissible terms for the "nothing special" case than is allowed by the statutory maximum. On the original Gottfredson-Wilkins model, the sanctioning rule imposes no further constraint: the only limitation on judges' discretion is an obligation (usually moral, rather that statutory, in force) to give reasons of some kind for any sentence outside the stipulated range. As we have already noted, however, guidelines may also prescribe the grounds on which departures from the "normal" range may be based, as well as those on which they may not be based; they may also limit the extent to which departing sentences may differ from the normal range. These further rules are intended introduce a greater degree of constraint or control over the behavior of the sentencing judge.

Guidelines may aim to control discretion in another way, however: to wit, they may make the descriptions in the antecedent condition X much more precise than the descriptions contained in criminal statutes. Offense categories such as burglary or rape are typically defined by legislatures in fairly broad terms; as a result, such categories will contain a wide range of cases of different degrees of seriousness. Part of the discretion accorded to the sentencing judge then consists of his being allowed to decide whether the case before him is a very bad rape or a trivial burglary; and in these decisions -- which may be "value judgements", but which ultimately turn on issues of fact -- judges may obviously differ among themselves. On the Gottfredson-Wilkins concept of guidelines, broad statutory categories are refined, and antecedent conditions are more precisely defined in terms of an offense seriousness score (x1), and an offender score (x2). These scores may be calculated in a number of ways; their effect, however, will generally be to partition broad statutory categories, thus introducing a number of sanctioning rules with relatively precise antecedent conditions in place of a single rule with broadly defined ones.

At the same time, guidelines as proposed by Gottfredson and Wilkins have a certain amount of flexibility. They define the normal or "nothing special" case in terms of two factors only, leaving other things which might legitimately influence terms within the prescribed range to be assessed by the sentencing judge. There is no logical necessity for this two-factor approach; sentencing guidelines could easily issue in sanctioning rules of the form "If x1 and x2 and x3 and x4, then y1→y2 is permissible", where x3 might refer to some facts specific to particular types of offense, such as extreme vulnerability of the victim, and x4 might refer to the offender's likelihood of recidivism.[62] (The guidelines recently developed by Zalman et al. (1979, 1980) in Michigan reflect something of this approach; we discuss these in Chapter 9.) Alternatively, such factors can perhaps be taken into account in calculating offense and offender scores in a two-factor model. There is an obvious tradeoff, in practice, between detail and precision in defining antecedent conditions, and flexibility in the sense discussed here; where the line is drawn between these things will no doubt vary from one jurisdiction to another. But some degree of flexibility is obviously desirable, if guidelines are to be workable in practice; to see this, we need only reflect on the probable unworkability of intricate rule-systems like the Ferri (1921) and Glueck (1928) codes discussed in the last chapter.

We deal with the question of enforcing compliance with sanctioning rules in general, and sentencing guidelines in particular, in the concluding section of this chapter. Before proceeding further, however, we wish to emphasize three things about the concept of sanctioning rules which we have employed in this section to analyze guidelines and other techniques for controlling discretion. The first is that this concept draws attention to a central fact about law (in particular, the criminal law) as a means of social control: this is that the use of force involved in criminal sanctions is itself rule-governed, and is permissible only under conditions spelled out in what we have called sanctioning rules. To be sure, those sanctioning rules may be few in number; they may be very vague and general. But there must be some such rules, if we are to speak of a legal system at all. Suppose that in a certain society there was widespread agreement that a certain set of rules of the form "Thou shalt not steal" should govern the behavior of individuals; these rules may have emerged from a moral consensus owing to habit, tradition, etc.; or they may have been imposed by some powerful elite. Suppose further that there are no rules of the form "If X then Y", which spell out what may and may not be done to those who violate the first set of rules; police, victims or vigilantes may shoot anyone whom they define as a violator, but there are no official constraints on the use of such force. Such a

system might indeed produce a great deal of conformity; but we would surely be reluctant to speak of it as imposing legal controls. That rules constraining official behavior are a central element of the concept of law may seem obvious, even to those ignorant of jurisprudence; yet this concept is curiously absent in some recent sociological studies of law.[63]

In fact, of course, the everyday behavior of judges, prosecutors and other actors in the criminal justice system is very highly and consciously rule-governed, not only by rules of law in the strict sense but by policies and principles that can be construed in exactly the same way as sanctioning rules. The authority (and, in one sense, the legitimacy) of these rules may vary; the rules themselves may be changed from time to time; and they may be inconsistently applied to particular cases. But that is not to say that no rules exist.

In our interviews with prosecutors in four Massachusetts counties, for example, we consistently found that most assistant district attorneys claimed to have fairly settled policies which they tried to follow in making recommendations as to sentences.[64] The extent to which these policies were articulated varied, in the four counties, as did the extent to which they were laid down by the district attorney himself as distinct from his individual assistants. Moreover, prosecutors were often unwilling to be very specific about their rules, for the same reason often given by judges: namely, that cases can vary a great deal, in many ways, so that "rigid" rules could not be applied. But this simply says that their rules were of the form, "If x1 and x2 and...xk, and (nothing special), then Y is the right recommendation". As we have seen, such rules have a certain open-endedness to them; but that does not mean that they are not rules at all.

Our second point is that, once it is recognized that discretion in sentencing (and analogous decision-making) is typically governed by at least some rules -- prescriptions that are meant to be followed by judges and others to whom they are addressed -- the concept of sanctioning rules can be used to compare different (actual and possible) rule-structures, as we have done in this section, to bring out the different amounts of individual discretion in rule-application which they allow. This factor can in turn be analyzed empirically, in a variety of ways. For example, what are the relations between the specificity and complexity of sanctioning rules, and the extent of compliance with those rules in practice? What differences are likely to be found between guidelines or other sanctioning rules promulgated by a legislature, and those adopted more or less formally by judges themselves? Finally, how do the rules work in practice? In

most American jurisdictions, of course, the antecedent conditions X in sanctioning rules are to some extent a matter of negotiation, through plea-bargaining. What proportion of variation in sentencing is due to this aspect of the rules, as distinct from the choice of a sanction Y given that some X has been agreed to obtain?

This brings us to our third point about the concept of sanctioning rules, which is that it brings into focus the vital distinction between the behavior prescribed by the rules, and the behavior of individual actors who are supposed to apply those rules to particular cases. We saw in the last chapter that there was general concern about unwarranted variation in sentencing, in the United States, even a century ago; there was concern about the use of judicial discretion even before the rehabilitative ideal and the ideology of "treatment" became entrenched in criminal justice systems, and this concern persists even though that ideal and ideology are being abandoned. But "disparity", or unwarranted variation, can come about in at least two very different ways. On the one hand, it may be that two or more judges share the same set of sanctioning rules (whether these be guidelines, or e.g. personal rules on which they happen to agree); but that they may apply these rules to particular cases in inconsistent ways, because of e.g. indigestion, infantile experiences, or class-membership. On the other hand, it may be that two or more judges are applying sanctioning rules in a consistent fashion, but that they are working with different rules. For example, one judge may have a belief in general deterrence; this may lead him, in effect, to apply a sanctioning rule prescribing heavy prison sentnces in certain cases. Another judge may believe instead in some notion of "rehabilitation"; this may lead him to apply a very different sanction to the same case. Alternatively, the two judges' sanctioning rules may prescribe the same sanction Y in certain kinds of cases; yet the judges may still sentence differently because one judge's antecedent conditions X for applying that sanction includes some elements which the other judge ignores. (This would be the case, for example, if the treatment-oriented judge thought that coming from a "broken home" were an indicator of a condition that psychoanalysis would cure. We have encountered stranger beliefs among judges -- not, we hasten to add, in Massachusetts.)

The distinction between these two sources of variation in sentencing is of obvious importance. Many critics of the judiciary -- including some of those discussed in the last chapter -- seem to have jumped prematurely to the conclusion that variation in sentencing is due entirely to the erratic application of an agreed-on set of what we have here called sanctioning rules. The truth of the matter, however, may be

that judges are perfectly consistently applying very different sets of sanctioning rules. This point has important implications for sentencing reforms of various kinds, including the introduction of sentencing guidelines. We will discuss these implications in more detail below.

The Supposed Empirical Basis of Sentencing Guidelines

As we have noted, the first decision-making guidelines by Gottfredson and Wilkins were based on a statistical analysis of past parole decision-making practice; and as we shall see in later chapters, it is equally the case that all of those responsible for the sentencing guidelines developed to date have taken as their starting-point some kind of statistical analysis of past sentencing practice, the purpose of which was ostensibly to identify those factors most commonly associated with variation in sentencing in the days before discretion in sentencing was to be "structured" by their work. In New Jersey, for example, the analysis was based on information on about 15,000 persons sentenced in 1976-77, on whom about 1,000 items of information per case had been collected.[65] The report containing these guidelines states that "it should be emphasized that the purpose of sentencing guidelines is not to persuade judges regarding what is the 'right' sentence or the 'best' sentence" (McCarthy, 1978:6), and elsewhere repeats the "descriptive, not prescriptive" claim made from time to time by Gottfredson, Wilkins and their colleagues (see, for example, Gottfredson, Wilkins and Hoffman 1978:chap.7, passim; Wilkins et al. 1976: 31-32; Kress, 1980: 1-12; also Zalman (1979); contrast Gottfredson, Wilkins and Hoffman 1978:141, 159). As we also saw in an earlier section, the concept of sentencing guidelines popularized by Gelman, Kress and Calpin (1977, 1979) laid great emphasis on statistical analyses of past sentencing practice as a first step toward developing guidelines; indeed, these researchers seem to have regarded "predictive efficiency" as the primary, if not the only, criterion for inclusion of an offense or offender variable in their sentencing "models".

Given the rhetoric of "description, not prescription", and its associated history of empirical research it is natural to assume that sentencing guidelines must be empirically based, and must primarily reflect sentencing practice in the same jurisdiction in the past. It takes only a moment's reflection, however, to see that this is not the case; and that the much-touted empirical basis of guidelines is by no means intrinsic to the construction of an instrument for controlling decision-makers' discretion. Such an instrument -- whether it is presented in matrix form like that illustrated in Table 3.1, or some other form -- consists essentially of a set of sanctioning rules; and such a set of

rules obviously could be made up, by a legislature, a sentencing commission, or a parole board, without any reference whatever to past practice. This is in fact precisely what happened with the Oregon parole board's guidelines, which were first developed in 1975 and given statutory authority in 1977. No analyses of past decision-making practices were carried out before these guidelines were formulated; instead, the board, under the chairmanship of Mr. Ira Blalock, simple made up the ranges of time which they thought appropriate to be served by different types of offenders. It is in fact unclear just how far Blalock and his colleagues were trying, in creating their guidelines and the associated definitional rules, to reflect past paroling practice in the state.[66] What is clear is that they did not carry out any detailed statistical analyses of those past practice, and of course it is clear that they did not need to do so. They simply prescribed.

It may still be argued that it is useful to begin with an empirical analysis of past practice, and that "obtaining an empirical description of current sentencing behavior is a reasonable first step in the process of sentencing guidelines development" (Zimmerman and Blumstein, 1979:2). For one thing, there may be a genuine feeling that what was done in the past was by and large right. We suspect that this comfortably conservative view has fairly widespread support, especially among the judiciary, though it is difficult to get anyone to admit this in public. Even if there is not a general feeling of this kind, however, it may well be felt (perhaps even correctly) that it is politically expedient to carry out an analysis of past practice as a preliminary to devising a prescriptive instrument. For example, it may be thought necessary to demonstrate the existence of a substantial amount of variation in sentencing, in order to show that there is a real need for guidelines; or it may be felt that it will be comforting to the judiciary to engage in a little preliminary number-crunching, calculated to show that sentences in the future will not be too different from sentences in the past.[67] It may have been for this reason that the Minnesota legislation directed that state's Sentencing Commission to "...take into consideration current sentencing and release practices..." devising guidelines (Minn. Laws 1978, ch. 723; Minn. Stat., ch.244 et seq.; see Minnesota Sentencing Guidelines Commission 1980:1). At any event, one must start somewhere; and empirical evidence concerning past practice is better than mere guesswork. However this may be, it should not be forgotten that it is perfectly possible to construct sentencing guidelines in the back-of-an-old-envelope fashion followed by the Oregon parole board; these might be called "guidelines by fiat".

It is in fact doubtful that most of the statistical analyses carried out by guidelines researchers to date have shown anything at all about past sentencing policies, as distinct from revealing some uniformities in past sentencing practice. (In some cases, as we shall see in later chapters, it is clear that the analyses could not have shown anything about policies.) To some extent this may have been obscured by the fact that the two primary factors used to construct most sentencing guidelines -- seriousness of current offense and prior criminal record, defined in various ways -- are the two things that have been found, in virtually every study of sentencing practice, in every jurisdiction, to be most strongly associated with variations in the severity of punishment. (Any study which does not come up with this result has almost certainly measured something incorrectly.) Thus those who (like Gottfredson and Wilkins) have described their findings as revealing the main elements of an "implicit policy" were almost certainly not wrong: it is difficult indeed to imagine a set of sanctioning rules, for a particular judge or a particular jurisdiction, in which seriousness of offense and prior record did not play an important role.

The trouble is that, as we noted earlier, this does not make much of a claim. To describe a past sentencing policy is to describe the set(s) of sanctioning rules actually used by judges in the past. At a minimum, this involves showing how such things as "seriousness of offense" and "prior record" were defined by the judges themselves; and what other factors (if any) were included in the antecedent or X-conditions in the sanctioning rules which they used. And the fact is that in none of the research done to date, as a preliminary to constructing guidelines, has anything been asked about those things. Instead, the research has been based on information retrospectively collected (usually from pre-sentence reports) about offenses and offenders; there is no guarantee that judges were even aware of that information, and still less reason to think that they necessarily based their sentences on it.[68] Analyses of such data might perhaps succeed in showing something about the objective correlates of sentencing; they can show nothing about sentencing policies in the sense of sanctioning rules.

In other words: a multivariate analysis of sentences imposed on a sample of past cases -- like those carried out by Wilkins et al. (1976) and other researchers -- can at best provide an "external" description of what judges actually did, "on the average", in sentencing those cases. This is not necessarily a description of what any judge actually considered, in sentencing the cases in that sample; it does not necessarily show anything about the grounds or reasons for those sentences, or about their causes.[69] None of the

research done to date, with a view to formulating "empirically-based" sentencing guidelines, has been based on anything that could reasonably be called a theory of judicial decision-making.[70] In the absence of such a theory, there is no way to interpret the results of the research (e.g. the coefficients of a regression equation) as describing sentencing policy. As Rich et al. (1980) have pointed out, an analysis of cases sentenced by a number of different judges will necessarily produce a description of sentencing practice on the average; and this average may not well describe the behavior of any individual judge in the group.[71] Even an analysis of data on the cases sentenced by a single judge, however, will by itself show nothing about the sentencing policies followed by that judge, i.e. the sanctioning rules which he or she actually applied in sentencing those cases.

This is not to say that an accurate and detailed "external" description of past sentencing practice is of no use, as a preliminary to constructing sentencing guidelines. For one thing, an analysis of past sentencing practice may convincingly show that a morally iniquitous factor such as race was strongly associated in the past with variations in sentence severity; in addition to showing that there is a need for guidelines in order to eliminate this iniquity, such an analysis may suggest ways of "purging" future sentencing policy of such racial influences.[72]

The conceptual, methodological and statistical problems surrounding multivariate analyses of past sentencing behavior have been discussed at some length by one of us elsewhere,[73] and will be considered at various places later in this report; it is thus unnecessary to present a detailed discussion of them at this point. The following points may however be briefly noted. First, it may be necessary to construct not just one, but several, models of sentencing practice, since it is likely to be the case that factors associated with the decision to incarcerate offenders may be different from those associated with lengths of terms for those incarcerated. As we pointed out earlier in this chapter, this fact was recognized by Wilkins et al. (1976), though in fact they constructed models using a single outcome variable by in effect scoring non-incarcerative sentences as being of zero months; as we show in Chapter 8, the same thing was done by the Massachusetts researchers.[74] Second, the correct specification of models of sentencing behavior may be very difficult indeed, not only because of the limitations of the data typically used in such analyses, but because there may be factors which, though relatively rare, are nonetheless important determinants of sentences when present. Third, at least in theory, some care is needed in choosing the statistical methods used to estimate such models; for example,

ordinary least-squares regression techniques may give misleading results when used to predict a dichotomous outcome variable such as the decision to incarcerate.[75]

Fourth, the overall fit of such models to data on past sentencing -- as measured by R2 or some analogous statistic -- is in general of much less importance than the significance and robustness of the coefficients of variables included in the model; the importance of things such as prior record may be clearly demonstrated statistically, even though (because there is a lot of variability in sentencing) the total proportion of variance accounted for may be very low.[76] Finally, and for the reason just mentioned, it is important to examine carefully the ways in which such multivariate models fail to fit data on past sentencing, e.g. by examining regression residuals; we illustrate this point in our reanalysis of the Massachusetts guideline data in Chapter 8.[77]

When all is said and done, then, adequate multivariate analysis of data on past sentencing can be a very difficult matter; it is certainly a far more complex and difficult enterprise than is suggested by the Gelman, Kress and Calpin (1979) manual discussed earlier in this chapter. As we shall show in later chapters, the analyses done by the several statewide guideline projects on which we obtained information in the course of this project often failed, for methodological reasons, to provide as detailed and accurate a description of past sentencing behavior as they might have done. If all that is wanted is a very rough description of that practice, which will identify a couple of factors (called "seriousness of offense" and "prior criminal history") from which a two-dimensional matrix can be constructed, these methodological issues may not matter very much; that was certainly the case, as we have seen, with the original Youth Authority guidelines created by Gottfredson and Wilkins. However, if the guidelines are to take a form other than the familiar two-dimensional matrix, then accurate description of past sentencing practice may become much more important. For example, as we shall see, the current Massachusetts guidelines take the form of a fairly direct transformation of a set of (unstandardized) regression coefficients yielding an "expected" sentence given certain attributes of the offense and offender; these coefficients were estimated from a sample of cases sentenced in the past, though they were to some extent modified on the basis of explicit considerations of policy. In such a case, mis-estimation of the relevant regression weights may have a direct effect on the "expected" sentences prescribed by the guidelines.

The concept of empirically-based sentencing guidelines has not infrequently been attacked, on the ground that it will lead to the institutionalization of injustices (like racism) which may have characterized sentencing policy in the past. But this criticism losed its force, if the distinction between description of (past) sentencing behavior, and the prescription of (future) sentencing policy, is recognized and clearly maintained. In fact, there is no need whatever to base guidelines (or any kind of sanctioning rules) on the results of research on past sentencing practice -- however politically expedient it may be to do this. It follows, a fortiori, that highly sophisticated multivariate modelling of past sentencing behavior is not needed in order to construct guidelines -- however aesthetically pleasing or scientifically useful such modelling might be.

This is not to say that empirical research and statistical analysis are of no use at all, to those involved in constructing sentencing guidelines, or in assisting in other kinds of sentencing reform. For example, as we describe in Chapter 9 of this report, those responsible for developing the Minnesota guidelines also developed a statistical model for estimating the impact of the guidelines on prison populations; they were thus able to show the consequences of different combinations of incarceration ratios and lengths of terms which might have been adopted. In later chapters of this report, we illustrate a variety of statistical and numerical techniques for assessing the structure of a guideline matrix, and for estimating its possible impacts on future sentencing practice; and in Chapter 10 we show how comparative cross-jurisdictional research can be carried out, so as to show the impact of guidelines on different offender populations. Finally, of course, empirical research is needed at some point, to determine whether or not guidelines introduced in a particular jurisdiction have made any difference to sentencing practice in that jurisdiction; though we were not able to conduct research of this kind ourselves, during the life of this project, we are firmly of the opinion that it ought to be done.

All of these examples, however, are of research that is done after the creation of guidelines or other sanctioning rules: make up the guidelines first, by fiat if necessary, then do the research. This is precisely the opposite of what is implied by the concept of "empirically based" sentencing guidelines that has been most widely promoted to date.

Description vs. Prescription, Reemphasized

In this section we show that even if guidelines purport to be "empirically based" in a very strict sense, there are

two fundamental elements which must necessarily be shaped almost entirely by judgemental or policy considerations, and cannot simply be a translation into policy of a description of past sentencing practice. These concern (1) the so-called "in-out" decision, and (2) the width of the prescribed "normal" range of jail prison sentences.

The Decision to Incarcerate

Suppose that, following the original concept of Gottfredson and Wilkins, we construct a guideline matrix the rows of which are defined by offense types and the columns of which are defined by prior criminal record; and suppose (though this is not strictly necessary) that the rows and columns are ordered from least to most serious in each dimension (as they are, for example, in Table 3.1 above). It will almost certainly be found that the probability of incarceration increases directly, and in a fairly orderly fashion, as one passes across the rows and down the columns of this matrix, from its upper left to its lower right.[78] There is no difficulty in principle in estimating such conditional probabilities of incarceration -- though in practice it may be as difficult to do this efficiently as it has historically been to predict elsewhere in criminology.[79]

Unfortunately, a probability of imprisonment is of very little use, where sentencing guidelines or other sanctioning rules are concerned: there is very little that a decision-maker can do with such a probability, when applying the rules to a particular case. Suppose that a statistical analysis of past sentencing practice showed that, say, 70 percent of all cases falling within a given cell in a guideline matrix in the past had been given "out" sentences such as probation or a fine. How can judges be instructed to comply with this finding, when sentencing in the future? They cannot send three-tenths of the offender to prison -- at least until more elegant forms of "split sentence" can be invented, than now exist in most jurisdictions; in any case, that is obviously not the meaning of the finding of 70 percent "out" in the past. Nor, however, can they easily comply with a prescription to the effect that only 30 percent of the group of offenders falling into that cell in the future should be incarcerated. Which 30 percent is that to be? It may indeed be that some further criteria (beyond those used to construct the matrix) can be found, that will discriminate reliably between the 70 percent of cases given "out" sentences, and the 30 percent sent to jail or prison. But this is by no means guaranteed, since the 70-30 split in that cell's cases in the past may reflect a complex set of factors which it is impossible to identify, or it may have reflected random variation between judges. The only purely "statistical" way

of complying with a prescription based on the empirical findings would be to toss a biased coin -- designed to come up "heads" three times out of ten, on average -- when dealing with cases in that cell; and this procedure is unlikely to commend itself to anyone.[80] The only alternative, however, is to declare that cases falling in that cell in the future shall presumptively be treated as "out" cases.

Such a presumption need not be entirely unguided; for example, it is possible to supplement it with a description (addressed to judges) of the kinds of factors which should or should not be considered in deciding whether or not the presumption should be over-ridden. Moreover, it is possible to provide for a presumptive "out" sentence, and still provide a "normal" range of months or years to be served if the presumption is over-ridden; both Minnesota's and Pennsylvania's guidelines, for example, do this. The need to rely on a presumptive "in" or "out" decision, however, does away with the flexibility inherent in the provision of a normal range, which was noted earlier to be a distinctive and useful feaure of the Gottfredson-Wilkins concept of guidelines (and which of course remains intact in the case of parole guidelines, where the concept was originally developed).

Moreover, the choice of which cells to treat as presumptively "in" or "out" is clearly a matter of judgement, and not something which can be completely settled by empirical evidence relating to past practice. Where ought one to draw the line? It was noted earlier that Gelman, Kress and Calpin (1979) recommended use of the modal sentence in each matrix cell, to draw the "in-out" line; similarly, Zimmerman and Blumstein (1979), in their reanalysis of the Denver construction data, treated matrix cells in approximately the same way, in order to test the predictive accuracy of their model of past practice. But it is surely unlikely that this cutting-point would ever be accepted in practice, even if it were shown to be a statistically reliable one; it would clearly entail a much greater change in sentencing policy than has ever been suggested by anyone advocating empirically-based guidelines. Even if the split between "in" and "out" cases is a fairly sharp one -- 70-30, say, or 80-20 -- it may be difficult if not downright arbitrary to declare that cases receiving the less common outcome after implementation of the guidelines must be regarded as departures from the guidelines -- unless the grounds for departure are quite strictly specified (as they are, for example, in Minnesota and Pennsylvania guidelines).

It may be thought that the presumptive, and thus relatively inflexible, character of the "in-out" decision can be avoided by designating "out" sentences as being of zero

months, and including them in guideline ranges in some matrix cells; this is done, for example, in the guidelines currently being tested in Michigan.[81] Similarly, under the Massachusetts guidelines it is possible for an offender to have an "expected" sentence, calculated by using the guidline formula, of zero months; indeed, zero (i.e. non-incapacitation) is the lower range limit for cases with a guideline score or expected sentence of between one and six months. But the difficulty with this approach is that it gives judges virtually no guidance on a crucial question, to wit: should this offender be incarcerated, or not? A guideline cell containing a range of "0 - 18" (months) constrains only the upper end of that range; an "out" sentence is by definition not a departure from the normal range, for which a special reason must be given, but then nor is any sentence of incarceration of up to 18 months. Such an approach thus does very little by way of "structuring" judicial discretion..

In summary, the problem is that empirical analysis of past sentencing practice can yield only probabilities of imprisonment, conditional on various offense and offender attributes; and it is very difficult to turn such probabilities into effective prescriptions for future sentencing, since it is not easy to follow a rule that says something like "Do such-and-such 35 percent of the time". It may well be that, as Zimmerman and Blumstein (1979) have suggested, one can by research on past practice identify three groups of cases: a group with very high rates of incarceration in the past (presumptively treated as "in" under the guidelines); a group with very low rates (presumptively treated as "out"); and a middle group (for which there would be no presumption, and thus no guidance given to decision-makers). The difficulty remains, however, that designating all of the cases in a cell as presumptively "in" or "out" seems likely to lead to changes in sentencing practice (and not merely a continuation of the empirically-ascertained practice of the past), unless the distribution of cases in that cell was in the past extremely skewed (e.g. with only five percent or so not receiving the modal disposition). Consider a cell in which 80 percent of the pre-guideline cases were imprisoned, for example. If this cell is designated as presumptively "in" under the guidelines, there seems a clear danger that the proportion of cases falling in that cell in the future will rise, ceteris paribus, unless it should happen that judges will find grounds to rebut the presumption in just 20 percent of the post-guideline cases. It is not easy to see how they can be given guidance of a kind that is calculated or likely to bring this about.

Widths of Prescribed "Normal" Ranges

There also seems no way in which the question "How wide should 'normal' ranges be?" can be answered, purely by an analysis of past sentencing practice. Guidelines developed to date display wide variations in this respect. Those in Minnesota and Pennsylvania, at one extreme, average plus or minus five percent around mid-ranges; by contrast, in Massachusetts the guidelines have a range of permitted variation of plus or minus fifty percent around the "expected" or guideline sentnce.

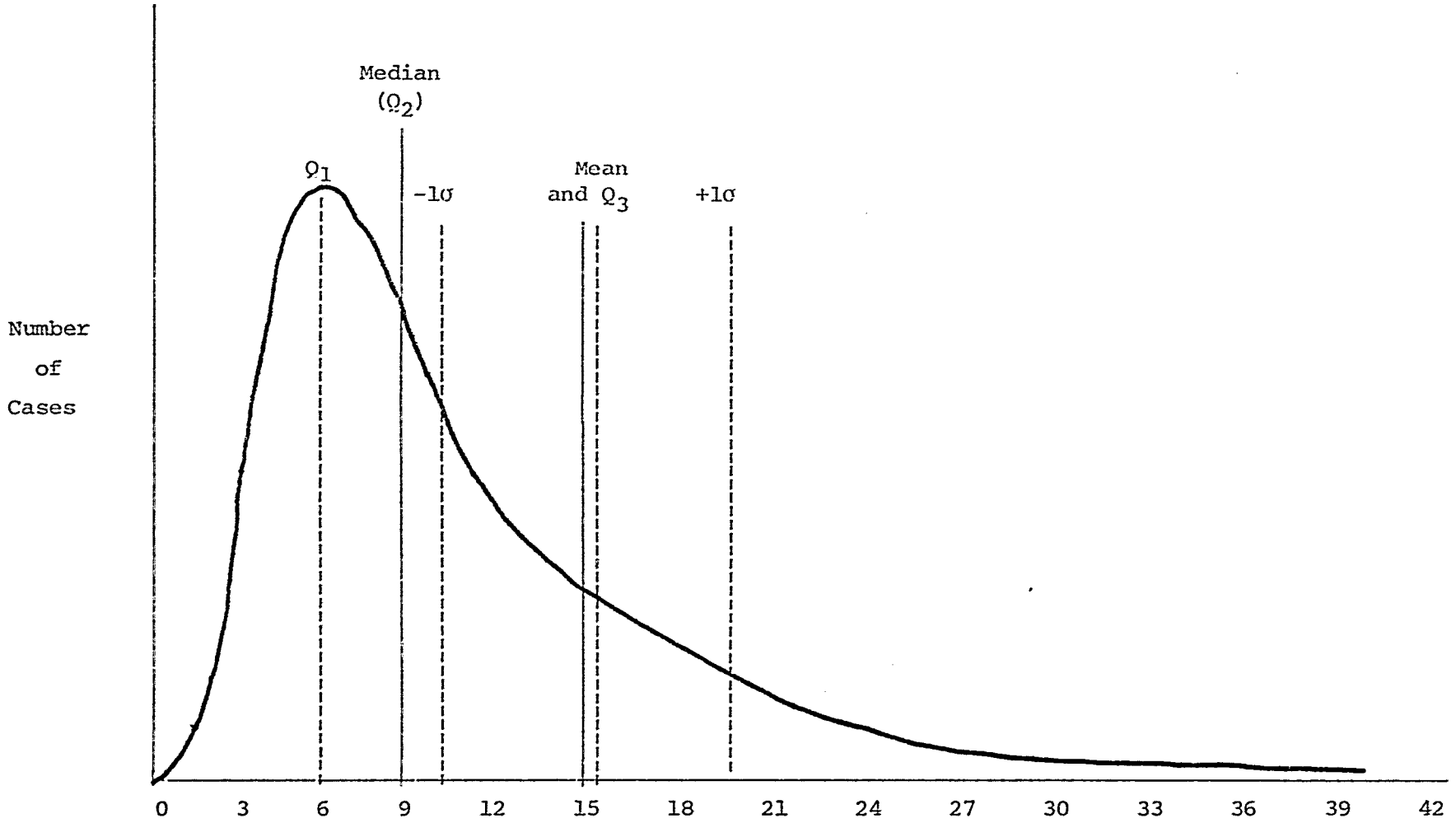
To be sure, one may be lucky. Inspection of the frequency distributions of lengths of terms in particular cells may show that these cluster within a reasonably narrow range, in most cells; an examination of the few cases falling outside that "natural" range may show that they had features which justified them in being treated as "departures". But it may also turn out that neither of these things is the case: in which case the question "How wide a range?" -- which is really the question, "What terms should be required to be justified, if imposed in the future?" -- has to be faced squarely as a question of policy, without any pretence of an appeal to past sentencing practice.

A statistical problem may arise, however, depending on the shapes of the frequency distributions of jail or prison terms and the ways in which the permitted range is defined. There is good reason to believe that, in general, the lengths of terms of incarceration imposed by judges (where they have effective control over this) are not symmetrical, but in fact are positively skewed, with a long right-hand tail (see, e.g., Banks (1964)). If this is so, then the choice of measures of location or central tendency, and of spread around that mid-point, may make a considerable difference to the resulting guidelines, and may in effect build in an increase in average terms. This situation is illustrated in Figure 3.1, which illustrates such a skewed distribution (ignoring its characteristic discontinuity).[82]

Insert Figure 3.1 about here

We have assumed that the mean of this hypothetical but not unreasonable distribution of past prison sentences is 15 months, and that its standard deviation is 4.5 months (since the distribution is roughly gamma or negative binomial in shape, its variance is bound to be greater numerically than its mean). However, as Figure 3.1 shows, the median term -- the term above and below which 50 percent of the cases lie --

Figure 3.1: Hypothetical Frequency Distribution of Lengths of Prison Terms



is much lower than the mean; in this case we have assumed it to be about nine months. The lower quartile is tolerably close to the median (here we have assumed it to lie at six months); but because of the shape of the distribution, the upper quartile will lie at a value which is more distant from the median, since we have to move further to the right to encompass 25 percent of the cases in that direction.

Now, how are we to construct a guideline range from these data? We might begin by characterizing the distribution's location by its mean, and setting the end-points of the range (arbitrarily) at one standard deviation above and below the mean; this would give a range of 10.5 months to 19.5 months, i.e. a spread of nine months centered around the former mean. But the effect of this would obviously to declare that in future cases given terms of less than 10.5 months would have to be treated as "departures" from the guidelines; yet as Figure 3.1 makes clear, over half of the cases in the pre-guidelines data had terms of less than 10.5 months. (At the other extreme, only terms of over 19.5 months would have to be treated as "departures"; yet there might well be some highly unusual cases not quite so far out in the tail, which should <u>not</u> be included in providing for a "normal" guideline range.)

If we begin by taking the median and then using the first and third quartiles, however, we will not be much better off. It is true that the median -- here, nine months -- more adequately picks out the center of the distribution than the mean. The interquartile range, however, is from six to 15 months; that is a spread of 10 months, but one that is centered on 10.5 months -- which, if the absolutely typical case is assumed to be given a term in the middle of the range, may lead to an increase of some 12 percent in the terms given such cases, compared with past practice! Moreover, if we assume that the distribution of prison terms in the past was reflective of the true "badness" or desert of the caes involved, it will be more difficult to find "departures" from the old middle case in an upward direction, than in a downward one.

Of course, such a situation need not obtain; it may be that the distribution of terms of incarceration, for any group of cases sentenced in the past, will be more or less symmetrical (or it may even be skewed to the left, though this seems to us unlikely). If things are as depicted in Figure 3.1, however, using either the mean and standard deviation, or the median and quartiles, will likely lead sentences imposed under the guidelines to be different from what they were in the past, in ways that might not be intended. All things considered, <u>if</u> it is wanted to reflect past practice, it seems

better to start with the median of the old distribution, and bracket that with a range defined in terms of "plus or minus x percent" (as Gottfredson and Wilkins and indeed most others who have so far developed guidelines have done). In Figure 3.1, for example, we might set a guideline range defined by the median plus or minus one-third, i.e. 9 + 3 = from 6 to 12 months. This would mean that -- if the distribution of "desert" continued as it had been previously -- more cases would in future have to be treated as "departures" on the upper or more severe end, than on the lower one. But that is surely no bad thing.

Controlling Compliance with Guidelines

We noted earlier that the only real element of control over sentencers' behavior, according to the original Gottfredson-Wilkins concept, lies in the (moral) obligation to give reasons of some sort if a sentence outside the prescribed "normal" range is to be imposed. Evidently the amount of control which that obligation imposes will be a function of the widths of those ranges; if they are wide enough, virtually any legal sentences will fall within them, so that the prescriptions contained in the guidelines would impose virtually no structure on sentencers' behavior. This element of control can be enhanced if, in addition to having reasonably narrow ranges, the guidelines are accompanied by rules (like those in Minnesota and Pennsylvania) which specify permissible grounds for departure; and it is obviously still further strengthened if (as in Pennsylvania) the extent of departure is also limited by rules. Even so, rules cannot determine their own application; so the question remains, how can we help to insure compliance with sentencing guidelines of whatever kind -- or with other kinds of sanctioning rules?

This is a general problem inherent in the concept of law as a technique of social control, and it is not something special to guidelines; at least, it is true for any legal rule which can be analyzed as implying a rule of the form "If X occurs, then Y is permissible." For example, in most developed legal systems there are rules that provide that if it is proved that one person, through his negligent conduct, caused some kind of harm or injury to another, he may be made to pay compensation for this; there are also rules to the effect that if there exists a contract between two parties (defined so as to include e.g. an "offer", "acceptance", "consideration", etc.), and if one party to the contract fails to do what has been agreed, then the other party may obtain damages (or, in some cases, may compel the defaulter to carry out his end of the bargain). But what is to stop the judge from finding that there was indeed negligence causing harm, or a breach of contract, but simply refusing to award damages as

the rule provides? Similarly, what is to stop the judge from imposing a punishment which is not permitted by a sanctioning rule, or failing to impose one that is required by it?

There is, it seems, no legal or logical absurdity in providing that the legal system contain further sets of sanctioning rules, the antecedent or X-conditions of which refer to behavior of judges which contravenes the general sanctioning rule; so that, for example, a judge could be hanged or thrown off the bench if he were to impose a sanction which was not permitted, or failed to impose a mandatory one.[83] But this is, of course, not usual (except in cases of extreme judicial malfeasance). Instead, the usual remedy for what are seen to be incorrect applications of the rules is by allowing the presumptively wronged party to appeal to a higher court -- this is, one having the legal authority to over-rule incorrect or improper actions of courts of first instance. There may of course be several layers of appellate courts of this kind, thought in most common-law jurisdictions the higher ones are restricted to settling disputed points of law as distinct from rehearing issues of fact.

In our opinion, appellate review of this kind is the only way, ultimately, in which compliace with sentencing guidelines or other sanctioning rules is likely to be insured. The guidelines developed in Minnesota and Pennsylvania contain provisions for some appellate review of decisions; a similar provision would be possible in Massachusetts, where appellate review of sentences to state prison has existed since 1943.[84] Such appellate review might do much more than just correcting inappropriate applications of the guidelines in particular cases; in addition, it might help to spell out, in greater detail than legislatively-written rules are likely to be able to, the kinds of fact-situations in which departures in either direction are appropriate -- much as, in England, the Court of Criminal Appeal (now the Court of Appeal, Criminal Division) has done since 1907 with the "principles" of sentencing which it has enunciated in appeals against sentence (Thomas, 1970).

Other techniques for trying to bring about compliance with guidelines (or other sanctioning rules) exist. The recruitment, selection and training of judges is one that is commonly suggested; the use of sentencing councils or sentencing institutes to impart information to judges and/or change their attitudes is another. But the first of these things is, in practice, likely to be influenced by many considerations other than a desire to promote compliance with a set of sanctioning rules for dealing with offenders; and we have already indicated our skepticism about the second. It might of course be the case that in a particular jurisdiction

there was clear consensus among the judges as to the appropriate sentences for most types of cases; and it might also be, in such a jurisdiction, that if guidelines were implemented they would be scrupulously followed by all judges in all cases. But it is difficult to see why there would be a need for sentencing guidelines -- especially empirically-based ones that purported merely to "make explicit" the policies underlying previous judicial decision-making -- in such a jurisdiction. The legislature might of course want to promulgate guidelines in such a situation, in order to change sentencing policy from what it had been; but those guidelines would clearly not be "empirically-based" either, at least in the sense in which that term has come to be used over the the past few years.

Conclusions about the Concept of Guidelines

It should by now be evident that we are more than skeptical about the concept of "empirically based" sentencing guidelines. Whatever may have been the practical merits of the original (parole) research by Gottfredson, Wilkins and their colleagues, and however appropriate it may initially have seemed to extend the ideas underlying that research to sentencing, the concept of sentencing guidelines subsequently promoted by Kress and others seems to us to be both incoherent and misleading. To the extent that they are seriously intended to be considered by judges, sentencing guidelines are never merely descriptive; and there is no necessity whatever to carry out an analysis of past sentencing practice before constructing guidelines. If such an analysis is done (and if it is methodologically adequate) the translation of the results into a prescriptive instrument (a set of sanctioning rules) is by no means an automatic process; the analysis itself, if done in the way that most have been done to date, can show little if anything about past sentencing policy, and entails nothing at all about policy for the future. Finally, in two important respects -- concerning the "in-out" decision and the widths of "normal" ranges of incarceration -- analysis of past practice is next to useless.

As we will show in the next few chapters of this report, those statewide guidelines which have been developed to date have only approximately adhered to the "empirically based" approach. The five states in which there have been efforts to develop guidelines can in fact be ranged on an approximate continuum, ranging from the mindlessly empirical to the intelligently policy-oriented; in none of those states, however, has the naive "predictionism" espoused by Gelman, Kress and Calpin (1979) been followed.

In this chapter we have dealt with the concept of guidelines itself, and have deliberately ignored the operational context in which sentencing guidelines must be used in actual practice. We have done this because that context is likely to differ from jurisdiction to jurisdiction to such an extent that little can usefully be said about it in general terms. One point should however be emphasized. By and large, the proponents of sentencing guidelines have concentrated exclusively on judicial decision-making, and on the construction of prescriptive instruments for "structuring" that decision-making. But sentencing does not take place in a judicially-controlled vacuum; and it is unrealistic to consider techniques for controlling judges' decisions in isolation from the activities of other actors -- in particular, prosecutors and defense counsel -- in the criminal justice system. Those other actors may have an important role in shaping the definition of cases that are subsequently sentenced by judges; through the use of their discretion, and the negotiation involved in plea-bargaining, they may influence the ways in which offenses and offenders are classified, before the guidelines are applied to them. In our field studies in four counties in Massachusetts, we found considerable variation (which we describe in Chapter 7) in prosecutors and defense counsel's attitudes to the sentencing process and the possible use of guidelines; similar variation almost certainly exists in most other jurisdictiions, which is likely to affect the impact of any guidelines which are developed and implemented.

We have also neglected, in this chapter, the broader political issues raised by sentencing guidelines. As we have seen, the earliest efforts to promote this kind of sentencing reform were aimed at judges themselves -- in part, perhaps, because there seemed little likelihood at that time (i.e. the early-middle 1970's) that guidelines could be legislatively mandated. Moreover, if guidelines are developed for a county-level jurisdiction (as was done by Gottfredson, Wilkins and their colleagues in their original feasibility study), the use of legislation to create and implement guidelines is obviously not an issue, since counties do not legislate. In can be argued, however, that the responsibility for developing and implementing statewide guidelines properly lies with the legislature and not the judiciary; and that any attempt to limit the statutorily-mandated discretion of the judiciary as a whole (as distinct from that of individual judges) is really a kind of ought to be left to elected representatives rather than to the judiciary itself, legislating, whatever the judges themselves may say about it. We express no views as to the merits of this argument; we note merely that it is an argument that we heard seriously presented (in Massachusetts), and one which needs to be seriously considered by advocates of guidelines as a kind of sentencing reform.

The very idea that there should be statewide guidelines, after all, itself involves a decision of social policy. Why should sentences for a given type of offense and offender be the same in urban areas as in rural ones? A probably apocryphal story says that it is regarded as more serious to shoot a cow in eastern Oregon than it is to shoot one's wife in western Oregon. Is such a view necessarily wrong -- given the economic and social importance of cattle in the eastern part of that state? Again, we express no opinion on the merits of this view; we note merely that the desirability and importance of uniformity in sentencing across the whole of a state has so far received far too little attention from those involved in developing sentencing guidelines. (As a technical matter, it is easy to incorporate regional variation in sentencing into guidelines, e.g. by constructing separate matrices or by putting a term for "region" in a regression-like formula like that used in Massachusetts.)

Despite our reservations, however, we believe that as an element in a set of sanctioning rules, the concept of guidelines is a useful one. The idea of stipulating a normally permissible range (rather than a presumptive point) allows for discriminations which we believe it may often be reasonable to make; the notion of a "departure" from the normal range, with its obligation to state reasons, is also useful -- provided that the permissible grounds for those departures are explicitly listed, provided that the amount by which sentences may depart from the normal range is also rule-controlled, and provided that the use of the guidelines can be enforced by a meaningful system of appellate review.

Whether statewide sentencing guidelines -- with or without the conditions just listed -- will "make a difference" to sentencing practice is a question which we would, at one time, have liked to try to answer; and in the concluding chapter of this report we show what kinds of research must be done in order to answer it. But that research will have to be done by somebody else, if it is done at all; we are no longer interested in doing it, and do not regret that we aren't.

Notes to Chapter 3

[1] In this account of the origins of the notion of guidelines we have relied primarily on published accounts of the early research conducted for the U.S. Parole Commission (see, in particular, the papers reprinted in U.S. Department of Justice, 1978a, 1978b). We have, however, also had the benefit of the retrospective accounts of both Gottfredson and Wilkins on certain points. Neither of them is responsible for the views expressed in this chapter, with which they may well disagree.

[2] The research in question was carried out at the Research Center of the National Council on Crime and Delinquency, in Davis, California, where earlier work on parole prediction had been done by Gottfredson.

[3] The way in which this would be done was described by Hoffman (1975:352-53) as follows: "After reviewing a case, the parole board member would complete the rating scales and make his recommendation. He would then check his recommendation against the matrix provided. If he found that his recommendation varied from the expected decision by more than a given amount (e.g., two months), he would be alerted to specify the considerations resulting in this decision or to reconsider his recommendation." It was apparently not envisaged that the board member would first consult the matrix, and then decide whether or not special features justified a different term. Yet this is percisely what is proposed under many subsequently-developed guidelines schemes -- whether or not it is what has happened in practice. This point is discussed in more detail below.

[4] This was in fact a collapsed version of an eleven-point Burgess-type scale, said to predict recidivism; it had a point-biserial correlation with parole violation of .318 (see Hoffman and Beck 1976:71). Somewhat different scores were later developed.

[5] Separate matrices were constructed for adults and young offenders.

[6] Gottfredson et al. (1975:6) state that the "discretion ranges" were determined after informal discussions with board members and hearing examiners, and "while arbitrary", were to some extent proportional to the size of cell medians. Inspection of Table 3.1 will show, however, that this is only approximately true; the ranges are proportionately broader in the left-hand half of the table, and do not vary much until the fifth row. This point is discused further in later sections.

[7] See, for example, Gottfredson et al., 1975; the interest in modelling decisions appears to have been Wilkins's (see, for instance, Wilkins and Chandler, 1965; Wilkins, 1975). But as we will note further below, no research on individual judges' or parole examiners' decision-making behavior was ever carried out by Gottfredson or Wilkins in their research on guidelines.

[8] See, for example, Gottfredson et al., 1975:1-2.

[9] E.g. Mannheim and Wilkins, 1955; Gottfredson and Ballard, 1965; Gottfredson and Beverly, 1962; Gottfredson, 1967.

[10] However, terms outside the guidelines range were required to be reviewed by the regional Administrative Hearing Examiner: see the rules reported in Hoffman and DeGostin (1974).

[11] In fact, up to that time the Board of Parole had publicly denied that it had any policies at all; it contended that it provided "individualized" treatment with decisions made on a case-by-case basis. Public criticism of the inequity which seemed to result from this seems to have been what made the Board eager to have Gottfredson and Wilkins discover what their "implicit" policy had really been all along. Compare Moliere's character who was relieved, as well as surprised, to discover that he had been speaking prose all his life. (Gottfredson has informed us, however, that in private the Board was concerned about disparity, quite apart from the criticism which their previous stance had engendered.)

[12] This point is discussed further in a later section, and see Chapter 11.

[13] In fact, their report actually says that "a parole board's decision to release or retain an inmate was not at all the equivalent of a sentencing judge's bifurcated decision" (Wilkins et al., 1976:20, emphasis added). This is presumably a slip. The quotation incidentally shows, however, that parole boards -- like judges -- sometimes do make dichotomous-outcome decisions.

[14] This issue will be discussed further in Chapter 8 below.

[15] See, for example, Gottfredson et al., 1975; Hoffman and Stone-Meierhoefer, 1977; Hoffman and DeGostin, 1975; Wilkins et al., 1976, esp. 1-19, 83-108; Gottfredson, Wilkins and Hoffman, 1978, esp. Chapter 2.

[16] It should be repeated at this point that we do not rely, for this interpretation of the intentions of Gottfredson, Wilkins and their colleagues, on their retrospective accounts, but rather on what they said at the time. They should not be assumed to agree with our interpretation of their past.

[17] See Kress (1980); Rich et al. (1980). On the New Jersey statewide guidelines, see below, Chapter 4.

[18] Wilkins continued to have a peripheral association with the implementation grant; in particular, he attended the planning conference early in 1977, discussed in the text below. Gottfredson did not get out of the guidelines business, but did further work on state parole systems: see Gottfredson et al. (1978).

[19] The title page of Kress (1980) identifies the author as being associated with the SUNY-Albany School of Criminal Justice. In fact, Kress left the faculty of that school in 1979.

[20] There is some uncertainty as to the sponsorship of the products of this contract and CJRC implementation grant. For example, Kress (1980) cites the manual on developing sentencing guidelines (Gelman, Kress and Calpin, 1979), as having been published by University Research Corporation in 1978. Our copies of this manual, however, are published by the Office of Development, Testing and Dissemination of NILECJ; one is dated 1979, another November 1977. Both are marked "Copyright 1977 by Criminal Justice Research Center." In this volume it is stated that the "methods manual" is the fourth report in a series of which the third is titled The Analytical Basis for the Formulation of Sentencing Policy. But this title is cited by Kress (1980) as having been published by the U.S. Government Printing Office in 1980. At the time of this writing we have not been able to obtain a copy of this report.

[21] At this time the Acting Director of NILECJ was Mr. Blair Ewing.

[22] See below, Chapter 7. This statement is based in part on interviews which we conducted in 1980 with members of the Massachusetts Superior and Supreme Courts.

[23] According to Kress (1980), he and his colleagues had only minor involvement with those responsible for the New Jersey sentencing guidelines (see also McCarthy, 1978). It is open to question whether or not this was a good or a bad thing: we discuss this further below. See also Martin (1981).

[24] "Less than one percent" of 40 judges is something under half a judge, of course; we have no idea what the quoted statement is supposed to mean. The figure of 98 percent presumably refers to 39 of the 40 judges interrogated.

[25] It was mentioned by several of the judges who attended the meetings on guidelines development sponsored by the American University, Criminal Courts Technical Assistance Project, in Cincinnati (March 1979) and Washington (April 1980), which members of our project staff attended.

[26] See Kress et al., n.d. In fact, the Indiana and Maine statutes did not introduce "mandatory type" sentencing; in particular, the Maine law merely abolished parole, leaving judges to fix "flat time" sentences within a very broad range.

[27] It is not clear from the brief outline description, for example, whether the objective of the "decision game" exercises was to demonstrate disagreement in sentencing to participating judges, or -- as was intended originally by Wilkins and Chandler -- to study judges' use of information, and the reasons underlying their choices and decisions. These uses of the "decision game" technique require slightly different procedures; both are important.

[28] In the introduction to this manual it is stated that it is "not a primer in social science research", and that "it is assumed that the reader has a basic understanding and experience with (sic) social science methodology and statistical analysis" (p.1). It is difficult indeed to reconcile these statements with some of the contents of the manual, e.g. the remarks on coding and "data cleaning" on pp. 12-13.

[29] The apparent purpose of this program is to "test" specified models by showing the proportions of cases they classify as "in" and "out", and to display cell median terms and the widths of ranges of ±12.5 percent of the cell medians. We have not attempted to run this Fortran program, thought it appears to be at least syntactically correct.

[30] Gelman, Kress and Calpin, 1979:14; the suggested significance level really is .005. That is, this figure is not a typographical error on our part; but it may be one on theirs. It is also suggested that researchers should re-run correlations using listwise deletion of cases with missing values; presumably this is to see whether this makes any difference. For multivariate as distinct from bivariate procedures, listwise deletion is recommended by the authors. Cf. Rich et al. (1980), where the impact of the latter procedure is discussed in relation to the Denver guidelines.

[31] The results from these two versions of the general linear model will of course be equivalent within limits of rounding error, unless a computational error has been made. Cf. Rich et al. (1980); Gottfredson, Wilkins and Hoffman (1978).

[32] This is indeed a real possibility, since the "how long" analyses should be based on a sub-set of the sample. However, so far as we are aware, in none of the CJRC analyses were separate models fitted. See also our discussion of the Massachusetts guidelines in Chapter 8.

[33] For example, there is no mention of the possible existence of suppressor or distorter effects on the bivariate analyses (cf., e.g., Rosenberg 1968); and Pearson's r is recommended despite the admission that many sentencing data do not meet its assumptions for inference, so that the use of rank-correlation coefficients (Kendall's tau, Spearman's rho) "will serve as an additional check". As a sample descriptive statistic, of course, r needs no assumptions at all. In any event, the suggested "rough cutoff" of $r=\pm.20$ is baffling, even if the significance level intended is .05 and not .005; an r of $\pm.90$ would presumably be suspect unless it attained some level of significance.

[34] See pp. 18-19 of the manual, where Burgess-type unit weights are recommended, with "at least a ten percent increase or decrease in incarceration rate across categories" as compared with the base rate; the reason for this recommendation is unclear to us.

[35] It is recommended that the validation sample should be drawn from cases sentenced at a later time than the construction sample, so as to see whether there have been any significant changes in court policy since selection of the construction sample! This is plain nonsense, of course; validation of a statistical model requires the assumption that there have been no such changes, hence the second sample should be drawn from the same population (in this case, the same time period) as the first. Continued monitoring of decision-makers' behavior, to see whether it has changed so that the original model no longer describes it, is also important; but this is a separate matter. It is also stated (on p. 21) that there is apt to be a fair amount of "shrinkage" of predictive accuracy for length of incarceration, so that "it is possible that the validation sample will have to be combined with the construction sample to form an experience table rather than a prediction tool." We do not know what this means.

[36] The suggested ranking procedure (Appendix D, pp. 85-85 of the manual) is unclearly described. However, it appears to be aimed at nothing more ambitious than three- or four-category ranks (rather than, say, Thurstone scales with a claim to interval-level measurement). For a more adequate treatment of this approach in the context of parole decision-making, see Hoffman, Beck and DeGostin (1975).

[37] Some statistical implications of this policy choice are discussed below. It is not clear why judges should be thought competent to choose between the mean and the median as measures of location.

[38] No citation is provided for this statistic; it is in fact a proportionate-reduction-in-error measure, invented by Ohlin and Duncan (1949), and similar to the Mean Cost Rating later developed by Duncan et al. (1953). On the latter statistic see the appendix by Lancucki and Tarling in Gottfredson, Wilkins and Hoffman (1978), where it is shown that the sampling distribution of MCR is related to that of Kendall's tau.

[39] However, it is noted on the same page that "a limited number of information items, perhaps even as few as four to eight, will account for nearly all of the variance in the dependent variable." This is of course more realistic.

[40] In the Denver pilot study (Wilkins et al., 1976:122), data were apparently collected on offender's height and weight; in the construction of the New Jersey sentencing guidelines an effort was made to obtain information on offender's father's education (though this was missing in 94 percent of the cases: McCarthy, 1978).

[41] See, for example, the discussion in Coffee (1975); the U.S. Parole Commission has routinely done this for some years.

[42] This is discussed in more detail in Chapters 9 and 11 below.

[43] See note 35 above; there is no discussion of how much "shrinkage" is allowable, thought it is arguable that this is necessarily a judgemental matter. For the view that double cross-validation and jacknifing (Mosteller and Tukey, 1977) may be the best procedures in contexts of this kind, see Larntz (1980).

[44] See Chapter 8, below, where we illustrate this procedure using the Massachusetts guidelines construction data.

[45] Except that it is suggested that some predictor variables may have highly skewed marginals; the implications of this are not discussed. On the characteristic skewness of lengths of term see below.

[46] See also Martin (1981). We have no first-hand knowledge concerning the New Jersey or Pennsylvania guidelines researchers.

[47] On a strict "treatment" orientation, under which the prisoner is not supposed to be released until he is "reformed" or "rehabilitated", the board's decision would presumably be dichotomous: release or not. (For an early statement of this view, in support of the idea that an administrative agency rather than judges should decide terms of incarceration, see Saleilles (1916).) The parole guidelines introduced in Oregon in 1977, which are explicitly based on "just deserts", provide by contrast that the prisoner's release date shall be fixed within six months of the day he enters the institution.

[48] Though it may be that in practice judges restrict themselves to just a few points, and avoid terms of e.g. seventeen months. Evidence of this discontinuity in the use of numbers by sentencers was first found by Galton in 1895: see Banks (1964).

[49] The legal status of such rules probably differs between jurisdictions. In Minnesota, the "scoring rules" have the force of statute law; in Massachusetts, by contrast, judges assign scores for injury and weapon use. See below, Chapters 7 and 9. The rules are not "addressed" to the judge, however, in the sense that they themselves contain no sanction which he is directed to apply.

[50] Given the usual conventions on the use of brackets, the set would be written as a disjunction, since NOT (A or B or C) is equivalent to NOT A and NOT B and NOT C, which is what is intended.

[51] See below, Chapter 9.

[52] These techniques are based on reasoning by analogy: see, for example, Cross (1961) and Levi (1953); and for the argument that the same techniques are used in the learning of scientific concepts, see Kuhn (1977).

[53] The Minnesota Guidelines Commission's report (1980:30) actually says "nonexclusive"; but presumably "nonexhaustive" is meant.

[54] See below, Chapter 9.

[55] The rule is that mitigated departure terms may be from the next right-hand cell of the matrix, unless the normally prescribed cell is in the right-hand column, in which case they can be from the adjacent cell above; aggravated terms work in just the opposite way.

[56] The Minnesota guidelines matrix provides both a presumptive term of incarceration, and (in some cells) a range equal to about five percent around that term. The Minnesota Commission's report (1980: 29-30) refers to deviations from both the range, and the mid-range, as "departures" requiring reasoned justification; the intention, however, is apparently that this applies only to terms outside the range. This is discussed further below.

[57] It is in principle possible to incorporate such provisions into a single rule; but there is no point in doing so, since there is no contradiction in assuming that the judge must simultaneously apply more than one non-mandatory sanctioning rule to the instant case -- unless, of course, the rules themselves contradict each other.

[58] The qualification is necessary since some sanctioning rules may say only what must not be done, and may give no positive direction as to what must be done. Rules prescribing maximum permissible sentences are an example.

[59] The same issue of choice between competing rules may come about if the rules in question represent different policies subscribed to by the individual judge, e.g. deterrence for some offenses but treatment for some types of offenders. This is one possible source of disparity: others are discussed further below.

[60] An example often cited to us by Massachusetts prosecutors during our fieldwork: it may be apocryphal. Besides redefining "possession", there are other ways of avoiding such unwanted results. For example, Heumann and Loftin (1977) found evidence that Detroit judges, required to give a mandatory two-year term for use of a gun, adjusted downward the sentences they imposed (consecutively) on offenders, so that the total sentence was about what had been given earlier in similar cases.

[61] The California law actually provides for three "base terms" of imprisonment; the middle one of these is presumptive. Departure from the middle term can only be to one of the two end-point terms; however, no other term in between can be used.

[62] One might expect that if more factors are used in this way, the range prescribed would be narrower; part of the purpose of providing a range, according to Gottfredson and Wilkins, was to allow leeway for variations in such other factors without having to specify them in the guidelines. In practice, however, it has not always worked that way: see below, Chapter 7, on the Massachusetts guidelines.

[63] See, for example, Black (1980). There are of course other criteria for distinguishing between legal and non-legal social control; and it is arguable that even within the legal sphere the most important form of social control is provided by the law of contract rather than the criminal law. See Sparks (1980c) for a discussion.

[64] See Chapter 6. These rules or policies were of course largely influenced by the need to obtain a conviction, by a guilty plea if possible.

[65] See McCarthy (1978); and Chapter 4.

[66] In interviews with one of us (Sparks) in 1979 (in the course of research on the Oregon parole guidelines) Blalock asserted that there had not been an attempt to mirror past practice, on the ground that their had been no consistent practice prior to the introduction of the guidelines. He then explained that the matrix had been constructed in part by reference to the maximum time that an offender would have to serve, given full "good time", and the board's desire to make the longest such terms (i.e. those in the lower right-hand corner of the matrix) sufficiently shorter in order to induce offenders to leave prison on parole rather than "maxing out" and being discharged not under supervision.

[67] This may have seemed especially important to Gottfredson and Wilkins, each of whom has informed us (in personal communications) that they saw little prospect of legislative mandates for sentencing guidelines (like the one subsequently to emerge in Minnesota), so that "self-regulation" by the judiciary seemed the best bet. See also Kress (1980), on the importance of involving the judiciary. This approach does not entail an empirically-based methodology; but it is easy to see that it might be found more congenial by the judges who in the end would have to use (or refuse to use) the guidelines.

[68] In our fieldwork in Massachusetts, for example, we were told by several judges that if there had been a trial (rather than a plea) they paid little attention to the pre-sentence report, since they felt that by the end of the trial they "knew" the offender personally. See further below, Chapter 6.

[69] The distinction between the cause of an action (in this case, a choice of sentence) and a reason for an action, though sometimes blurred, is clear enough here; reasons, which might include some general policies or aims held by the judge, must be distinguished from the grounds for a sentence, i.e. the factors on which the sentence is consciously and explicitly based. Note that there may also be causes of a sentence, e.g. racial prejudice or sexual attitudes, of which the judge is not consciously aware. Thus even a significant association between race and sentences does not by itself show that judges were deliberately racist.

[70] Though some of the writings of Wilkins have dealt with aspects of decision-making theory, these were not rigorously applied in the sentencing feasibility study (1976); nor, indeed, have they been integrated into anything like a systematic theory of decision-making in sentencing, in our opinion.

[71] In their reanalysis of the Denver data, Rich et al. (1980) show that there were in fact several differences between the judges in that study, in the weights which they attached to different factors. It is perhaps not surprising that none of those who have so far attempted to develop empirically-based guidelines have published analyses of sentences by different judges -- given that the idea was in most cases to sell the idea of guidelines to the judges, it was presumably thought more diplomatic not to stress this cause of variation.

[72] Fisher and Kadane (1981) have argued, in our view correctly, that the way to do this is not to leave race (or other iniquitous variables) out of modelling equations, but to include such variables and then remove them from the guidelines; if this is not done coefficients for such things as current offense and prior record may be mis-estimated. Our Massachusetts data show a somewhat different pattern below, as we discuss in Chapter 8. We discuss this still-unresolved issue in a forthcoming reply by Stecher and Sparks to the Fisher-Kadane paper. See also Chapter 9 below.

[73] See Sparks, (1980b).

[74] As we shall see in Chapter 8, it may also be necessary to develop separate models for cases sentenced after trial and for those in which there is a plea, since different factors may figure in those cases.

[75] Though in practice this use of ordinary least-squares -- in what is sometimes called a linear probability model -- may give substantially the same results

as more theoretically correct techniques such as probit analysis or logistic regression: cf. Berk (1980), Zimmerman and Blumstein (1979); and for a recent review see Gottfredson and Gottfredson (1981).

[76] We owe this point to Professor Frank Fisher. In fact, the R square values obtained by most guidelines researchers have been in range .40 to 55, which is not at all bad by social science standards. See, for instance, Zalman (1979); and compare Sutton (1978a, 1978b).

[77] See also Berk (1981) for a discussion of regression errors and residual problems which he notes may arise even if the estimated model is well-specified and contains genuinely causal variables. This is of course not the case with most (if not all) models of sentencing behavior.

[78] Though the progression may not be all that orderly, even if the concepts of offense seriousness and prior record are appropriately defined. For some techniques which can be used to examine this and other aspects of decision-making guideline matrices, see Chapter 9, below.

[79] See, for recent reviews of this aspect, Simon (1974), Gottfredson and Gottfredson (1981).

[80] However, some utilitarians might argue that such a procedure would be a desirable one since it would obtain the same amount of deterrence without imposing the threatened sentence in all cases, and ex hypothesi offenders would not know what they would receive if convicted. We have heard a version of this argument presented by Professor Norval Morris.

[81] See below, Chapter 9, for a description of the Michigan guidelines. It should be noted that even if this practice is thought appropriate for the guidelines, it does not follow that preliminary descriptive analyses should use a single outcome variable: to do so will lead to mis-estimation of regression weights for offenders actually incarcerated. (See below, Chapter 8, for an illustration with the Massachusetts construction data.) Contrast Rhodes (1981).

[82] On which see Banks (1964); the finding that "penal deserts" are in practice not a continuous variable, and that some terms are not in practice used by judges at all, was first noted by Galton.

[83] It should be emphasized again that in general the notion of judicial conformity to sanctioning rules required only that the judge's behavior should not contravene the rules; it is only where the rule provides a mandatory sanction that positive non-compliance can easily be demonstrated.

[84] See Zeisel and Diamond (1977); and below, Chapter 6, for a discussion of the Massachusetts sentence appeal provisions. As we make clear in that chapter, we do not regard that appellate procedure as an adequate one for enforcing compliance with guidelines.

Chapter 4: The New Jersey Sentencing Guidelines

The New Jersey sentencing guidelines were introduced to the judiciary of New Jersey on October 23, 1978. The culmination of a two-year research project conducted by the New Jersey Administrative Office of the Courts, they were the first statewide sentencing guidelines to be implemented in the United States.

Use of the guidelines by judges in New Jersey has not been legally mandated; but it was strongly urged by the Administrative Office of the Courts. The guidelines were said to be intended to serve as an "information tool" for sentencing judges -- to provide them with information on the typical sentences imposed (in the year ending 31 September 1977) on various types of offenders (McCarthy, 1978:2). Thus, the New Jersey guidelines purported merely to provide a description of prior sentencing practices throughout the state, for combinations of specific offense types and offender characteristics. It seems clear that in reality, however, the New Jersey guidelines were intended to do much more than merely describe. Even though they were said (in the report just cited) to be "merely advisory", it was plainly intended that they should to some extent influence the sentences subsequently imposed by New Jersey judges; that is, it was intended that they should prescribe sentencing behavior as well as describe it.

The main focus of this chapter is on the structure of the New Jersey guidelines. We show that the guidelines in their present form are unnecessarily complicated, and that a much simpler format can provide almost exactly as good a description of sentencing practice in New Jersey. In addition, however, we argue that the New Jersey guidelines in their present form provide little if any guidance to those judges who may wish to use them in order to reduce disparity in sentencing. We also argue that there are serious issues of principle raised by the construction and use of the New Jersey guidelines; and that for several reasons those guidelines represent an inappropriate method of controlling sentencing discretion.

Of necessity, however, we must begin this chapter with a fairly detailed description of the New Jersey guidelines; for they do not resemble the usual kinds of guidelines described in earlier chapters. Our main thesis is that the information contained in the New Jersey guidelines can be much more economically presented; but it will be impossible for the reader to understand that thesis, unless he has a full appreciation of the lack of economy in the guidelines as they now stand.

Description of the New Jersey Guidelines

The New Jersey sentencing guidelines consist of two reports, each containing information on sentencing which was derived from an analysis of all sentences imposed in New Jersey in 1977 (about 16,000 cases in all). The first volume, containing 158 pages, gives information on eleven types of offenses -- Breaking and Entering, Robbery, Assault, Sale of Controlled Dangerous Substances, Possession of Controlled Substances, Larceny, Forgery, Fraud, Weapons, Rape, and Lewdness. The second volume, containing 224 pages, deals with four other categories and 99 types of offenses for which there were too few cases in the data base to permit statistical analysis; for these offenses, the information in the guidelines consists of a brief description of the facts in each case (e.g., "set fire to paramour's bed"), and the sentence imposed in that case. We shall have nothing more to say about this second volume.

The first volume, after some brief introductory material, contains a 33-page discussion of the concept of guidelines, a description of the research methodology by which these were constructed, and a brief discussion of their use. It then gives information on each of the eleven categories of offense mentioned in the preceding paragraph. Each of these categories of offense is dealt with separately; it is for this reason that the New Jersey guidelines have been described as "crime-specific", though this description is not in fact quite accurate since most of the eleven categories contain cases involving several different statutes. (For example, the category of Assault contains offenses from five different sections of the New Jersey statutes, and includes atrocious assault and battery and assault with intent to kill as well as "threatening to take a life".)

Before the sentencing judge can even begin to select the appropriate guideline sentence for a person convicted of an offense in one of the eleven categories, he or she must first determine the offender's numerical scores on each of the five offender characteristics These are called Prior Criminal History, Amenability to Non-Custodial Supervision, Exacerbating Factors, Community Background, and Actions Since Arrest.[1] The possible values of the scores on each factor are the same, no matter what the type of current offense may be. Scores on the Criminal History variable can be 0, +1, or +2; Amenability to Non-Custodial Supervision and Exacerbating Factors may be scored either 0 or +1; Community Background and Actions Since Arrest may be scored either -1 or 0. However, the definitions of the five attributes, and the way in which scores are derived on them, differ from offense to offense. As an example, the scoring procedure for those convicted of Breaking and Entering is shown in Figure 4.1.

Insert Figure 4.1 here

It will be seen that a burglar may receive a score of +2 on the Criminal History variable if he has six or more adult convictions or juvenile petitions (or, presumably, both) for any offense; or if he has four or more convictions, etc., for any crime; or three or more convictions, etc. for breaking and entering, or two or more incarcerations (of any kind whatsoever). A person convicted of an offense in the category of Robbery, however, would receive a Criminal History score of +2 if he had four or more convictions for any offense, or two or more convictions for any crime, or one or more prior convictions for robbery, or one or more incarcerations. Thus, in no case would the two offenders' Criminal History scores be calculated in exactly the same way.

For seven of the eleven offense categories, information relating to all five of the offender attributes is used to locate the appropriate guideline sentences. For two offense categories -- Forgery and Lewdness -- only the attributes of Criminal History, Exacerbating Factors and Community Background are considered. The sentencing guidelines for Rape offenses use those three attributes plus the factor of Actions Since Arrest, while the Weapons category uses those three things plus Amenability to Non-Custodial Supervision. An overview of the attributes used for the different offense categories is given in Table 4.1. As this table shows, only three attributes are included in the guidelines for all eleven offense categories: namely Criminal History, Exacerbating Factors and Community Background. But as we have noted, even these three things are defined in different ways, for different offense categories. (As Table 4.1 also shows, the number of things which may be considered exacerbating factors varies greatly. In the case of Forgery offenses, only one thing -- whether the total cash value of the offenses was greater than $1,000 -- is listed; in the case of Larceny offenses there are ten possible exacerbating factors of which any two will lead to an Exacerbating Factor score of +1.)

Insert Table 4.1 here

The scoring procedures just described lead to two different numerical assessments of offenders. We refer to the first of these as a configuration, using this term to refer to a particular pattern of scores on the (usually) five attributes. Thus, suppose that an offender has a score of +2 on Criminal History, +1 on Amenability to Non-Custodial

Figure 4.1: New Jersey Sentencing Guidelines

Breaking and Entering Score Derivation*

A) CRIMINAL HISTORY

1. Score 2 if total adult convictions or juvenile petitions sustained for any offense is greater than five, or if total adult convictions or juvenile petitions sustained for any crime is greater than three, or if total adult convictions or juvenile petitions sustained for any similar offense is greater than one.

2. Score 1 if total adult convictions or juvenile petitions sustained for any offense is between two and five inclusive, or if total adult convictions or juvenile petitions sustained for any crime is between one and three inclusive, or if total adult convictions or juvenile petitions sustained for any similar offense is equal to one or two, or if the total incarcerations is equal to one.

3. Score 0 is none of the above is true.

B) AMENABILITY TO NON-CUSTODIAL SUPERVISION

1. Score 1 if the offender was under criminal justice supervision at the time of the offense, or prior probation was negatively evaluated, or presentence report indicates that offender is drug dependent.

   Score 0 if none of the above are true.

C) EXACERBATING FACTORS

1. Score 1 if the crime included one or more of the following exacerbating variables:

   i) Offender convicted also on a weapon charge
   ii) Goods stolen included those of only sentimental value
   iii) No strong need for money, money was "extra" or for "fun" only
   iv) Person was apparently present in the structure entered
   v) Offender committed multiple breakings and entries, consider convictions only
   vi) Offense included property damage over $100

2. Score 0 if the crime did not include one or more of the exacerbating variables.

D) COMMUNITY BACKGROUND

1. Score -1 if the offender was employed, in military, or in school at the time of offense and has a job, military, or school to go to now, or score -1 if the offender contributes to the support of other persons.

2. Score 0 if the above conditions are not met,

   (NOTE: THE ONLY SCORES POSSIBLE ARE -1 OR 0)

E) ACTIONS SINCE ARREST

1. Score -1 if the offender has voluntarily entered a drug or alcohol treatment program, secured employment, made restitution, sought psychiatric help, entered school, sought skills or trade training or otherwise attempted to rectify past mistake, AND has entered a guilty plea.

2. Score 0 if the above conditions are not met.

*In each case the five factor scores must be calculated on the sentencing sheet. Then the appropriate cell and score can be identified on the guideline matrix, which contains the summarized information on the sentences meted out to offenders possessing the same factors (similarly situated). Judges will receive specific guideline information with each presentence report.

Table 4.1: Overview of New Jersey Sentencing Guidelines

Attribute Inclusion by Offenses

Offense Type	Criminal History	Amenability to Non-Custodial Supervision	No. Exacerbating Factors		Community Background	Actions Since Arrest	No. Matrix Cells (Total = 396)
				No. Needed			
Break & Enter	X	X	6	1	X	X	48
Robbery	X	X	6	3	X	X	48
Assault	X	X	4	1	X	X	48
Drug Sale	X	X	6	2	X	X	48
Drug Possession	X	X	5	1	X	X	48
Larceny	X	X	10	2	X	X	48
Weapons	X	X	4	2	X		24
Fraud	X	X	3	1	X	X	48
Forgery	X		1	1	X		12
Rape	X		4	1	X	X	16
Lewdness	X		4	2	X		8

Number of factors needed for a rate score of 1

Supervision, 0 on Exacerbating Factors and Community Background, and -1 on Actions Since Arrest; his configuration could be written (+2 +1 0 0 -1). An offender with scores of +1 on Criminal History, Amenability to Non-Custodial Supervision and Exacerbating Factors, 0 on Community Background and -1 on Actions Since Arrest would have a configuration of (+1 +1 +1 0 -1). Once an offender's configuration has been determined, the judge may use the "48 cell locator sheet" (like the one shown in Figure 4.2) to determine the "cell number" corresponding to that configuration; using this "cell number", or the configuration itself, he or she can then locate the row of the "guideline matrix" (see Figure 4.3) which gives information on sentences for cases of that type. The term "cell number" is in fact something of a misnomer; it refers, in fact, to a particular row of the guideline matrix, the columns of which correspond to particular dispositions.

Insert Figure 4.2 here

The second numerical measure derived from the offender attributes is called a cell score; this is simply the algebraic sum of the elements making up the configurations. Each of the two configurations mentioned in the preceding paragraph has the same cell score, namely +2; and as Figure 4.3 shows, there are nine other configurations which also have this cell score. (For ease of reference, we have put heavy lines around cell number 7, with its configuration of (+2 +1 0 0 -1); this is because we will use this row of the Breaking and Entering matrix to illustrate further discussion later in this chapter.) For seven offense categories which make use of all five offender attributes, cell scores may range from -2 to +4. The two extreme scores can obviously be obtained in only one way each; but all of the other cell scores can be obtained in several different ways. Thus, eleven different configurations will yield cell scores of +2, and other eleven will yield cell scores of zero; scores of +3 and -1 can each be obtained in five different ways; and a cell score of +1 can be obtained from no less than fourteen different configurations or possible combinations of attributes scores.

Insert Figure 4.3 here

So much for the rows of the guideline matrices. What about the columns? As Figure 4.3 shows, these give the total number of cases in the construction sample with that configuration; the percentage of those cases incarcerated; the

Figure 4.2: 48 Cell Locator Sheet

CRIMINAL HISTORY	AMENABILITY TO NON-CUSTODIAL SUPERVISION	EXACERBATING FACTORS	COMMUNITY BACKGROUND	ACTIONS SINCE ARREST	CELL NUMBERS
2	1	1	-1	-1	1
2	1	1	-1	0	2
2	1	1	0	-1	3
2	1	1	0	0	4
2	1	0	-1	-1	5
2	1	0	-1	0	6
2	1	0	0	-1	7
2	1	0	0	0	8
2	0	1	-1	-1	9
2	0	1	-1	0	10
2	0	1	0	-1	11
2	0	1	0	0	12
2	0	0	-1	-1	13
2	0	0	-1	0	14
2	0	0	0	-1	15
2	0	0	0	0	16
1	1	1	-1	-1	17
1	1	1	-1	0	18
1	1	1	0	-1	19
1	1	1	0	0	20
1	1	0	-1	-1	21
1	1	0	-1	0	22
1	1	0	0	-1	23
1	1	0	0	0	24
1	0	1	-1	-1	25
1	0	1	-1	0	26
1	0	1	0	-1	27
1	0	1	0	0	28
1	0	0	-1	-1	29
1	0	0	-1	0	30
1	0	0	0	-1	31
1	0	0	0	0	32
0	1	1	-1	-1	33
0	1	1	-1	0	34
0	1	1	0	-1	35
0	1	1	0	0	36
0	1	0	-1	-1	37
0	1	0	-1	0	38
0	1	0	0	-1	39
0	1	0	0	0	40
0	0	1	-1	-1	41
0	0	1	-1	0	42
0	0	1	0	-1	43
0	0	1	0	0	44
0	0	0	-1	-1	45
0	0	0	-1	0	46
0	0	0	0	-1	47
0	0	0	0	0	48

Figure 4.3: 48 Cell Guidelines Matrix: Breaking and Entering or Entering

CRIMINAL HISTORY	AMENABILITY TO NON-CUSTODIAL CARE	EXACERBATING FACTORS	COMMUNITY BACKGROUND	ACTIONS SINCE ARREST	CELL NUMBER	SCORE	NUMBER OF CASES	PERCENT INCARCERATED	STATE PRISON		INDETERMINATE		COUNTY TERMS 12 MOS. OR LESS		COUNTY TERMS MORE THAN 12 MOS	
									NUMBER OF CASES	MEDIAN TIME	NUMBER OF CASES	MEDIAN TIME	NUMBER OF CASES	MEDIAN TIME	NUMBER OF CASES	MEDIAN TIME
2	1	0	-1	0	6	2	66	63%	15	5 yr	12	5 yr	13	9 mo.	1	18 mo.
2	0	0	0	0	16	2	56	58%	6	4 yr	6	5 yr	21	6 mo.	None	-
2	1	1	-1	-1	1	2	32	51%	5	5 yr	5	5 yr	5	9 mo.	1	18 mo.
2	1	0	0	-1	7	2	133	50%	23	5 yr	14	5 yr	25	12 mo.	4	18 mo.
1	1	1	-1	0	18	2	26	50%	3	4 yr	5	5 yr	5	11 mo.	None	-
2	0	1	-1	0	10	2	12	50%	2	3.5 yr	1	6 yr	3	6 mo.	None	-
1	0	1	0	0	28	2	42	48%	3	3 yr	7	5 yr	10	6 mo.	None	-
2	0	1	[illegible]	-1	11	2	9	44%	None	-	1	5 yr	3	9 mo.	None	-
1	1	0	0	0	24	2	107	44%	5	3 yr	18	5 yr	24	6 mo.	None	-
1	1	1	0	-1	19	2	32	43%	1	3 yr	10	5 yr	4	7.5 mo.	None	-
0	1	1	0	0	36	2	10	40%	2	5.5 yr	2	5 yr	None	-	None	-
TOTALS. ALL CELLS IN THIS SCORE						2	525	50%	65	4.5 yr	81	5 yr	113	7 mo.	6	18 mo.

SC-3- BX

numbers sent to the state prison, to indeterminate terms at the Yardville reformatory, to the county jail (for 12 months or less), and the county penitentiary (for 18 months or less). Median terms for those incarcerated are also shown. The bottom row of the guideline matrix gives, in effect, column sums and corresponding median terms; that is, it summarizes the information for all cases in a particular offense category (in this case, Breaking and Entering) with the same cell score (in this case, +2).

The New Jersey sentencing guidelines can thus be thought of as a huge contingency table, with rows corresponding to combinations of offense category and offender attributes, and five columns corresponding to dispositions (four different types of institutional setting, plus those not incarcerated at all -- which, though not actually shown in matrices like the one in Figure 4.3, can easily be obtained by subtraction). The number of rows differs, for different offenses (cf. Figure 4.1 above), but totals 396; thus the guidelines contain a total of 1,980 cells (396 x 5) in all. (The tidy-minded may prefer to think of a three-way crosstabulation -- offender characteristics by dispositions, within categories of offense --containing a total of 2,640 cells, 760 of which are structurally or definitionally empty since not all combinations of offender attributes are used for some offenses.) In practice, the sentencing judge locates the row corresponding to the offense category and offender attributes of the case being dealt with; the columns of that row tell the judge the proportion of cases of that kind (in the construction sample) incarcerated at all, the numbers of cases incarcerated in each of four types of institution, and the median terms of those four groups of cases.

Of course, even though the New Jersey guidelines are very detailed, they do not give information on all of the cases on which they are based; they could be unpacked still further. In Figure 4.4 we give a hypothetical distribution of cases which might underlie a single line of the matrix -- in this case, the configuration (+2 +1 0 0 -1), or "cell number" 7, in the Breaking and Entering matrix in Figure 4.3. The two bottom lines of Figure 4.4 restate, in a slightly different form, the information contained in the guidelines -- neglecting the 67 offenders in the construction sample who had the same configuration but were not incarcerated at all. The upper part of Figure 4.4 shows a hypothetical distribution of sentences in the 66 incarcerated cases, which might have given rise to that single line of the guidelines.

Insert Figure 4.4 here

Figure 4.4: New Jersey Sentencing Guidelines: An Overview for Breaking and Entering

1 Cell Configuration (Cell No.: 7)

	Criminal History = 2			
	Amenability to Supervision =1			
Controlling for:	Aggravating Factors=0			
	Community Background = 0			
	Actions Since Arrest = -1			
	Incarceration = Yes			
	Number of Cases Sent to Each Institution			
Sentence in Months	County to 12	County to 18	YRCC	NJSP
1- 6	12	1		
7-12	13	1		
13-18		2		4
19-24			3	6
			2	1
60			7	1
			2	1
96				5
				1
120				2
				1
300*				1
No. Incarcerated Cases (N=66)	25	4	14	23
Median Sentence (in months)	12	18	60	60

*Figure chosen solely for illustration

Although this distribution of cases is hypothetical, we do not think it is unreasonable. It will be seen that we have not distributed cases uniformly across all of the possible sentence lengths shown in Figure 4.4; this is because each of the four institutional settings in New Jersey -- county jail, county penitentiary, Yardville and the state prison system -- has different ranges of sentence lengths associated with it. Thus, county jails may only receive inmates with sentences of one year or less, while inmates may be sent to a county penitentiary or workhouse with a sentence of up to 18 months. Yardville, an institution that is part of the New Jersey Youth Correctional Center Complex, is limited to indeterminate sentences which usually have a maximum of five years. (However, an inmate may be given a state prison term to be served at Yardville as a young adult offender.) State prison sentences have minimum and maximum sentences in all cases, and the minimum sentence must be one year or more; what the guidelines show is median maximum terms. Our hypothetical distribution of cases across the various sentence lengths in Figure 4.4 has been constructed with these limitations in mind. We have also borne in mind the fact that in New Jersey (as in most other jurisdictions) judges tend to use only a few of the possible sentence lengths legally available to them, for particular institutions. Although we do not present the evidence here, it can be shown that most county jail sentences in New Jersey are for six months or twelve months (that is, most of the 56 hypothetical cases shown as having sentences "1-6 months" in Figure 4.4 probably had sentences of exactly six months); similarly, most Yardville sentences are for five years (60 months), though a substantial minority have a maximum of two years or 24 months; maximum state prison terms also tend to be for exact numbers of years, with numbers such as 2, 5, 8, 10 and 25 years (300 months) being common.

Part of the point of the hypothetical data in Figure 4.4 is to illustrate the fact that a single row of the guidelines (in this case, "cell number" 7, for Breaking and Entering) compresses, and thus conceals, a certain amount of variation in sentencing; in the case of sentences to the state prison, in particular, this variation is probably very great indeed. As Figure 4.4 also shows, there is some overlap between institutional settings; an offender may serve six or twelve months in the county jail or the county penitentiary, or may be sentenced to five years in Yardville or in state prison. But there is not much overlap of that kind. And if -- as we suspect -- a judge begins by deciding where an offender is to be sent (rather than for how long he is to be confined), then the range of possible durations of confinement is automatically curtailed. For the sake of completeness, we hammer this point home in Figure 4.5, which presents hypothetical data which might underlie the 265 incarcerated

cases in the construction data with a cell score of +2 and a current offense of Breaking and Entering. Again, the bottom two lines of Figure 4.5 present the information contained in the guidelines; the rows above show some sentences which might have given rise to those totals and median terms. Again, we believe that our figures, though purely hypothetical, are not at all unreasonable. It is probable that no offenders are sent to county jails for seven months, though that is (according to the guidelines) the median term. For county penitentiaries, it is not unreasonable to suppose (as we have) that 18 months is not only the median but also the modal term; it is also incidentally the legally permissible maximum. In the case of sentences to Yardville, five years (indeterminate) is the modal maximum term; it is also the median term, but that is mainly because deviations in either direction from that term are highly unusual. In the case of state prison sentences, the median of 54 months -- or four and one-half years -- for cases with cell scores of +2 may well not have been imposed on any offender in the construction sample. And because the frequency distribution of sentence lengths is almost certainly extremely skewed to the right (as shown), the variation around the median term is by no means symmetrical. These points have some implications to which we shall return later.

Insert Figure 4.5 here

Reconstructing the New Jersey Guidelines Data

As we noted earlier, the New Jersey sentencing guidelines are based on data relating to about 16,000 adult offenders sentenced in 1977; 10,629 of those offenders had committed an offense in one of the eleven categories with which we are concerned. Where available, data were collected (from pre-sentence reports) on a total of 842 variables; the resulting data set is obviously very rich, and contains an enormous amount of information which could be used by researchers to tackle a large number of questions about the criminal justice system in general, and sentencing in particular, in the state of New Jersey.

Unfortunately, as we explained in Chapter 1 of this report, access to these data has so far consistently been denied to researchers (other than those involved in constructing guidelines) by the New Jersey Administrative Office of the Courts. Thus, it has so far not been possible for us to carry out a thorough analysis of the guidelines, using the construction data; and there are many important questions which neither we nor any other researchers, to our

Figure 4.5: New Jersey Sentencing Guidelines: An Overview for Breaking and Entering

1 Cell Score

	Cell Score = 2			
	Incarceration = Yes			
	Number of Cases Sent to Each Institution			
Sentence in Months	County to 12	County to 18	YRCC	NJSP
1- 6	56	1		
7-12	57	1		
13-18		4		
19-24			21	25
			5	5
60			51	4
			1	5
96			3	9
				4
120				7
				3
300*				3
Total Incarcerated Cases (N = 265)	113	6	81	65
Median Sentence in Months	7	18	60	54

*Figure chosen solely for illustration

knowledge, have been able to address. This is extremely unfortunate -- the more so since the publications of the Administrative Office of the Courts, concerning the guidelines and other matters, are by no means as clear or complete as they might be.[2]

Fortunately, some cats can be skinned in more than one way. As we showed in the preceding section, the guidelines themselves present data relating to a few offense and offender variables, in an extremely disaggregated form. Using a simple Fortran program it was possible to generate case-level data[3] (N = 10,629) containing, for each case, information on the following items:

1) Offense Type
2) Criminal History Score
3) Amenability to Non-Custodial Supervision Score
4) Exacerbating Factor Score
5) Community Background Score
6) Actions Since Arrest Score
7) Incarcerative or Non-Incarceraive Disposition
8) Place of Incarceration (if incarcerated)
9) Median Sentence (the same median sentence is used for all offenders, within each configuration; who were sentenced to each particular institution).

The analysis in the rest of this chapter is based for the most part on the reconstructed data set just mentioned. It is "unauthorized", in the sense that we received absolutely no official cooperation from those responsible for making up the New Jersey guidelines. After a description of the data set, we will address the following issues:

(1) Can the New Jersey guidelines be simplified, without losing the degree of guidance which they might provide to sentencing judges? In particular, can the complicated mess of "configurations" be done away with, without losing information which might be useful to a judge who wants to know how his judicial brethren have sentenced in the past? Essentially, this is the question whether the predictive power of the guidelines would be weakened if the Burgess-type "cell score" were used (instead of configurations) to predict the "in-out" decision.

(2) What are the relations between the five offender attributes which are used in the guidelines? How well do they predict the decision to incarcerate, and what would be the effect on the predictive power of the guidelines if some or all of them were eliminated?

(3) Is it possible to combine offense types without predictive loss?

(4) What are the probable effects of using a point -- the median term -- instead of a range, as a sentencing guideline? And how do the guidelines serve to indicate the place in which an offender should be incarcerated?

(5) What are the implications of the way in which the New Jersey guidelines were constructed? In particular, why were the five offender attributes, as defined, included -- and on what principle could their inclusion be defended?

Description of the Reconstructed New Jersey Sentencing Guidelines Data.

The following is a preliminary analysis of the 10,629 cases which are included in the first volume of the New Jersey sentencing guidelines. This discussion will focus primarily on how the cases in the construction data are distributed within the variables of offense category, the decision to incarcerate, Criminal History, Amenability to Non-Custodial Supervision, Exacerbating Factors, Community Background, Actions Since Arrest, institution of confinement, and median sentence to incarceration.

Offense of Conviction

The largest offense group in the sample of 10,629 cases was that of Breaking and Entering (20 percent). None of the other offense groups included in the Statewide Guidelines possessed as large a number of cases. The second largest offense group was that convicted of Drug Possession, which included 12 percent of the sample's cases. Following the offenses of Drug Sales (12 percent) and Weapons (11 percent), Robbery accounted for the fifth largest offense group, with 10 percent of the cases convicted of this offense. Of the remaining offense groups, Assault, Larceny, and Fraud, each accounted for between nine and ten percent of the total sample of cases. The offenses of Forgery, Rape, and Lewdness included between two and four percent of the total sample of cases.

Criminal History Score

Of the 10,629 cases on which this preliminary analysis is based, only 28 percent were rated in terms of the sentencing guidelines as having a Criminal History score of 0. Thirty percent of the sample included those offenders listed as having a Criminal History rating of +1, and 41 percent of the sample were rated as having a Criminal History score of +2.

As we have noted, interpretations of the Criminal History score is complicated by the combinations of individual items

used to derive the score, which differ across offense groups. The individual items that are components of the Criminal History factor score are: adult convictions or juvenile petitions sustained for (a) any offense [4], or (b) any crime [5], or (c) similar crimes, or (d) prior incarcerations. With a few minor exceptions (Rape and Lewdness offenses), these four component items are included in the derivation of the Criminal History score for each offense group. The problem encountered during analysis, however, is that the number of any specific type of prior criminal behaviors (letters a, b, c, and d above) varies from offense to offense. While it might be plausible to assume that a Criminal History score of 0 would at least indicate the absence of any of items a, b, c, and d above, even this does not hold for the offenses of Breaking and Entering (score zero equals one prior offense of any type), Larceny (score 0 equals up to two prior offenses of any type), and Forgery (score 0 equals two prior offenses of any type or one prior conviction for any crime). However, it is the score 0 category that includes the least amount of variation across offenses in terms of its numerical derivation.

Insert Table 4.2 here

Amenability to Non-Custodial Supervision

Fifty-five percent of the 9,867 cases that were given scores on the Amenability to Non-Custodial Supervision variable [6], were given a rating of 0 on that variable. In terms of the specific items included in New Jersey guidelines as components of the Amenability to Non-Custodial Supervision score, a rating of 0 generally means that the defendant was (1) not under supervision at the time of the offense, and (2) had not had a prior probation negatively evaluated, and (3) was not indicated to be drug dependent by the information appearing on the pre-sentencee investigation report. Thus, 45 percent of the sample of offenders who were given a rating of +1 for this characteristic can be assumed to have any one (or more) of these above factors.

Insert Table 4.3 here

Exacerbating Factors

Due to the complex method of scoring used to derive the Exacerbating Factor score, it is very difficult to say specifically what a score of 0 or +1 means.

Table 4.2: Criminal History Factor Score

No. Adult Convictions or Juvenile Petitions For:	B & E		Robbery		Assault		Drug Sale		Drug Poss.		Larceny		Weapon		Fraud		Forgery		Rape	Lewdness
	+2	+1	+2	+1	+2	+1	+2	+1	+2	+1	+2	+1	+2	+1	+2	+1	+2	+1	+1	+1
Any Offense*	>5	2-5	>3	1-3	>3	1-3	>3	1-3	>3	1-3	>6	3-6	>3	1-3	>4	1-4	>3	3	≥1	≥1
or Any Crime*	>3	1-3	>1	1	>2	1-2	>2	1-2	>2	1-2	>3	1-3		-	>3	1-3	>2	2	-	-
or Similar Offenses	>2	1-2	>0		>2	1-2		-	>1	1	>2	1-2	>1	1	>1	1		-	-	-
or Incarcerations	>1	1	>0		≥1		>1	1	>1	1	>1	1	>1	1	>1	1	>1	1	-	-

Score "0" always indicates "none of the above"

*Cf. notes 4 and 5 in the text

Table 4.3: Amenability to Non-Custodial Supervision Factor Score

	Score "1" if any of the factors checked are present, for offenses of:										
	B & E	Robbery	Assault	Drug Sale	Drug Poss.	Larceny	Weapon	Fraud	Forgery	Rape	Lewdness
Under Supervision at Offense	✓	✓	✓	✓	✓	✓	✓	✓			
Prior Probation Negatively Evaluated*	✓			✓	✓	✓	✓	✓			
Drug Dependent According to PSI	✓	✓	✓		✓	✓	✓	✓			

"✓ " indicates those items included for each offense

Score "0" always indicates "none of the above"

*Negative evaluation could refer to either an unsatisfactory probation term, or to a probation revocation

First of all, not all of the exacerbating factors apply to all offenses. That there were multiple offenders involved in the instant offense, for example, would be explicitly considered as an exacerbating factor for the crime of Robbery, but would not be considered for the crime of Burglary. A glance at Table 4.4 clearly reveals that factors are sporadically considered for some offenses and not for others, and explicit consideration of any factor as an exacerbating item never counts to the favor of an offender.

Insert Table 4.4 here

The exacerbating factors that have been included are grouped into sets of items, relevant to specific offenses; drug offenses, for example, incorporate factors such as selling of drugs for profit or habit, the type and amount of the drug sold, etc. into the score. Of course, inclusion of items specific to drug offenses is reasonable; however, drug-related considerations may apply to other crimes as well, i.e., robbery to obtain money to support a habit. Yet, none of the factors included in the New Jersey sentencing guidelines appears in all of the offense groups.

A second problem encountered when attempting to interpret the concrete meaning of the exacerbating factor score is that the number of component factors that must be present to receive a score of +1 differs for each offense. A robber, for example, must have three of the component items present in order to receive a +1 score; a forger need have, and can have, only one item present -- that the forgery involve more than $1,000.

Aside from differences in the total number of items needed to receive +1 scores in each offense, in two cases -- for the offenses of Drug Sales and Larceny -- factors have been sub-grouped. (These two quirks in the system have been identified by single and double asterisks in Table 4.4.) To use the Larceny sub-group as an example, two factors are listed: goods valued over $500 and Motor Vehicle theft. <u>If either one, or both</u>, of these factors were present, the offender would receive a point. This procedure differs from that usually followed to compute the exacerbating factor score; for a robbery offense, for example, presence of any two items would total to two points, not one point.

Another problem confounding interpretation of the exacerbating factor score is that often the <u>degree</u> of the factor considered differs from one offense to another. For a Lewdness conviction, <u>any</u> injury to the victim is seen as an

Table 4.4: Exacerbating Factor Score

Score = 1 if:	1 or more	3 or more	1 or more	2 or more or sale $2000	1 or more	2 or more	2 or more	1 or more (Not Welfare Fraud)		1 or more	2 or more
FACTORS INCLUDED	B & E	ROBBERY	ASSAULT	DRUG SALE	DRUG POSS.	LARCENY	WEAPON	FRAUD	FORGERY	RAPE	LEWDNESS
Multiple Offenders		X	X				X				
Multiple Counts	X	X	X		X	X		X		X	X
Multiple Charges						X		X			X
Convicted on Weapon Charge	X	X		X		X				X	
Multiple Firearms							X				
Weapon was a Loaded Firearm							X				
Weapon Used to Injure or Threaten							X				
Injury - Caused by Weapon			X								
Any Injury											X
Injury - Emergency Treatment										X	
Injury - Serious			X								
Forced Sodomy										X	
Offender Used Force		X									
Offender - Leader of Ring						X					
Offender - Organized Operation				X*	X	X					
Offender - Habitual Con-artist								X			
Offender - Sold for profit, habit				X							
Access to unlimited drug amounts				X*	X						
Offender sells drugs (pusher)				X*	X						
Drug was heroin or opiates				X	X						
Victim (Purchaser) - Juvenile				X							X
Goods of Sentimental Value Stolen	X					X					
Stole for fun	X					X					
Person in Structure During Offense	X					X					
Property Damage over $100	X										
Cash (Drugs) Involved over $200		X		X							
Street/Commercial Robbery		X									
Forgery over $1000									X		
Possession of Burglary Tools						X					
Goods Valued over $500						X**					
Motor Vehicle Theft						X**					

*Presence of 1, 2, or all 3 of these factors would count as only 1 point.
**Presence of 1, or both, of these factors would count as only 1 point.

exacerbating item, while the injury caused by an Assault has to be either serious or caused by a weapon in order to be considered an exacerbating factor. Operational definitions are not supplied by the guidelines to aid the judge in determining whether a certain amount of injury was "serious".

However, to give a brief description of the frequency distribution for this factor, it suffices to say that 54 percent of the total sample of 10,629 cases were rated as having a score of 0, indicating -- depending on the offense -- that either none or few of the Exacerbating Factors were present. Forty-six percent of the sample received a score of +1 for the presence of some aggravating factors.

Community Background

Of the 10,629 cases included in the analysis of this factor, 40 percent of the cases were given a rating of -1.[7] Sixty percent of the cases were given a rating of 0. The component items upon which these scores were based were somewhat similar across categories. That the offender had a job, the military or a school to go to after sentencing, appeared in every offense category; items including whether the offender had been employed, in the military, or in school at the time of the offense, whether he contributed to the support of other persons, or had an alcohol or psychiatric problem, appeared in a majority of the offense groups; but none appeared in all offense groups. The fact that the offender had plead guilty is a component of the Community Background score only for the offense of Weapons. (It should be noted here that this item -- viz, having pled guilty is not included in the Actions Since Arrest score for Weapons offenses.) The method for determining the offender's rating of -1 or 0 is somewhat different for the Community Background score. A score of 0 always means that the offender did not have any of the above positive items present. However, a score of -1 is derived from combinations of the above items. For example, if an offender convicted of Breaking and Entering was employed, in the military, or in school at the time of the offense and had a job (military or school) to go to after sentencing, or contributed to the support of other persons, he/she would receive a score of -1. The score of -1 is not given unless the offender can satisfy either one or the other condition; and it should be noted that the first condition has two parts. Thus if the offender did not have a job, the military, or school to go to both before and after sentencing, or did not contribute to the support of other persons, he would receive a Community Background Score of "0". This unusual requirement, that complicated combinations of factors be present, hinders our ability to analyze this score category in terms of specific component items that applied to offenders in the original construction data.

Insert Table 4.5 here

Action Since Arrest

The Action Since Arrest Score appears for each of the following offenses: Breaking and Entering, Robbery, Assault, Drug Sales, Drug Possession, Larceny, Fraud, and Rape. The items included within the Action Since Arrest score are exactly the same across all offense groups listed above, with the exception of the offense category of Rape. These factors include: (a) entered drug/alcohol treatment, or (b) secured employment, or (c) made restitution, or (d) sought psychiatric help, or (e) entered school, or (f) sought skills/trades training, or (g) otherwise attempted to rectify past mistakes, and (h) entered a guilty plea. Again, as in the Community Background score, the presence of combinations of items is of utmost importance. Although we cannot generalize about the distribution of offenders across the first seven factors listed above, we can know that in order to get a score of -1 the offender must always have entered a plea of guilty. A score of 0 always can be expected, therefore, to mean that the offender did not meet both of the above conditions. For the offense of Rape, which is the only exception to the rule that one of the first seven factors just listed be present, the only requirement to receive a score of -1 is that the offender has entered a plea of guilty. Absence of the plea of guilty as a component item for the scoring of Weapons offenses is somewhat misleading, as mentioned before, as this item is included within the Community Background score. The distribution of cases for the Actions Since Arrest score, is as follows: 40 percent of those offenders included in the 10,629 case sample received a score of -1 indicating that these offenders had at least pleaded guilty to the offense of conviction (this need be the only factor present for Rape offenders) and (for other offenses) possessed at least one of the other seven supplemental condition items. Sixty-eight percent of the sample, on the other hand, received a score of 0 on the Action Since Arrest category, indicating that these offenders had not met both of the above conditions.

Insert Table 4.6 here

Institution of Confinement

Sixty-two percent of the 10,629 cases in the construction data received a non-custodial sentence. Information is not included in the New Jersey sentencing guidelines as to the

Table 4.5: Community Background Factor Score

	B & E	Robbery	Assault	Drug Sale	Drug Poss.	Larceny	Weapon	Fraud	Forgery	Rape	Lewdness
Employed, in military, in school at offense	X		X	X	X	X	X			X	
	AND		AND	AND	AND	AND	AND			AND	
Has job, military, school after sentencing	X	X	X	X	X	X	X	X	X	X	X
	OR		OR			OR	OR	OR	OR	OR	OR
Contributes to support of other persons	X		X			X	X	X	X	X	
Has neither alcohol nor psychiatric problem											X
							AND				
Plead guilty							X				

Score "0" always indicates "none of the above"

Table 4.6: Actions Since Arrest Factor Score

	B & E	Robbery	Assault	Drug Sale	Drug Poss.	Larceny	Weapon	Fraud	Forgery	Rape	Lewdness
Entered Drug/Alcohol Treatment, or	X	X	X	X	X	X		X			
Secured Employment, or	X	X	X	X	X	X		X			
Made Restitution, or	X	X	X	X	X	X		X			
Sought Psychiatric Help, or	X	X	X	X	X	X		X			
Entered School, or	X	X	X	X	X	X		X			
Sought Skill/Trades Training, or	X	X	X	X	X	X		X			
Attempted to Rectify Past Mistake	X	X	X	X	X	X		X			
	AND	AND	AND	AND	AND	AND		AND			
Entered Guilty Plea	X	X	X	X	X	X	*	X		X	

Score "0" always indicates "has not met both of the above conditions"
*Absence of this factor here is misleading as this factor is included under the "Community Background" Section

number of these offenders who received different types of non-custodial sentences --such as probation, restitution, fines, community mental health treatment, or drug treatment. In addition, assuming that a large proportion of these cases received some sentence to probation, information is not provided within the Sentencing Guidelines about the length of the probation supervision. New Jersey judges have commonly used the "split sentence" as a sentence alternative. Usually, a "split sentence" would combine a sentence to a county jail or penitentiary with a term of probation supervision to follow immediately thereafter. Those cases that may have been split sentences of this format are not identified by the New Jersey guidelines, though apparently (see McCarthy, 1978:18) they were treated as "in" decisions.

Of those cases incarcerated, 12 percent were sentenced to New Jersey State Prison, 10 percent were sentenced to the Youth Correctional Center Complex, 14 percent were sentenced to county jails, and 2 percent (259 cases) were sentenced to the three county penitentiaries presently existing in New Jersey. (These facilities are located in Essex, Mercer, and Hudson Counties.)

Median Sentence in Months

A cursory analysis of the lengths of the median sentence affixed to the cells of the guideline matrix indicates that median sentences can range from zero months of confinement for non-custodial sentences to 444 months imprisonment, or 37 years. The distribution of median sentence appears to be, on preliminary analysis, a multi-modal distribution. As we noted earlier, the usual term lengths to which offenders are sentenced in New Jersey are the lengths of 6 months, 9 months, 12 months, 18 months, 3 years, 4 years, 5 years, 7 years, and 12 years. Again, all 10,629 cases were included in the analysis. Offenders sentenced across all offense groups in the New Jersey sentencing guidelines had a mean sentence of 16.8 months, with a standard deviation of 32.2 months.

Do Configurations Matter?

As we have explained, for most offenses the New Jersey guidelines make use of information on five offender attributes; offenders are given a numerical score on each of these attributes, to provide what we have called a configuration, e.g. (+2 0 0 -1 0); the elements of such configurations can be added to produce a cell score, e.g. +1 for the configuration just mentioned. For seven of the eleven offense types, cell scores may range from -2 to +4; and except for those two extreme scores, they may be obtained from a number of different configurations or patterns of attribute scores.

The configurations are used to locate the row in the guidelines which is appropriate to the offender being sentenced; they are described in the guideline manual as identifying "the primary guideline which should be observed by the judge" (McCarthy, 1978:31). Cell scores are described as a "secondary guideline", most useful in evaluating rows in which there are extremely small numbers of cases; judges are cautioned "not to rely too heavily on the total cell score information, the secondary guideline" (McCarthy, 1978:32).

Reference back to Figure 4.3 will show that the different configurations giving rise to any cell score are listed, in the guidelines themselves, in rank order of percentage incarcerated. At one extreme, among those sentenced for Breaking and Entering with a cell score of +2, 63 percent of those with a configuration of (+2 +1 0 -1 0) were incarcerated, whereas only 40 percent of those with a configuration of (0 +1 +1 0 0) were incarcerated. The same thing is true for configurations within other cell scores; e.g. of those with configuration (0 +1 0 -1 -1), 42 percent were incarcerated; of those with configuration (+1 0 0 -1 -1), only 10 percent were incarcerated; both groups have cell scores of -1.

At first sight, then, it would seem that the configurations make (and are intended to make) an important difference to sentencing under the New Jersey guidelines; in the example just given, the chance of imprisonment increases by over 300 percent (from 10 percent to 42 percent), though the two groups of offenders have the same cell score of -1. In other words, it seems that what matters, so far as sentences of incarceration are concerned, is not merely the Burgess-type cell score obtained by adding together the five attribute scores, but also the way in which that cell score is obtained. This is one of the features of the New Jersey guidelines which makes them so complicated; since to cater for all possible configurations requires (for seven of the eleven offenses) a total of 48 cells.

Further analysis of the guidelines shows, however, that the configurations are not as important as they may at first sight seem. For one thing, the differences in percentages incarcerated, among different configurations with the same cell score, often are not really very great. The incarceration rates for the cell configurations shown in Figure 4.3, for example, do not show appreciable change across configurations. The suspicion arises that most of them would not be greater than might be expected by chance, through sampling variation.[8] If this is so, then the numbers incarcerated, given any configuration, should not differ appreciably from the number expected, given the percentage

incarcerated for the corresponding cell score and the number displaying the configuration in question. In other words, a "no configuration effects" model should fit the data. To assess this model, chi square tests were carried out on as many as possible of the groups of configurations (representing cell scores of -1 and +3) for each of the seven types of offense having 48-cell guideline matrices. In six of these 35 cases, there were too few cases imprisoned to make the chi square test possible; in one other -- cell score of +1, for larceny -- there is an apparent error in the guidelines, so the test was not carried out. In the other 29 cases, configurations with no cases were ignored; but cases with extremely small expected numbers (based on the overall percentage incarcerated for the particular cell score group) were included. This tends to inflate the computed value of chi square somewhat, so the test is conservative in the sense of maximizing Type I errors; it should if anything exaggerate the number of statistically significant differences.

Nonetheless, in only one of the 29 cases did the value of chi square attain a level equivalent to the .05 level of confidence, given the appropriate degrees of freedom. (At that level of confidence, of course, one such statistically significant result should be expected purely by chance; that is what the ".05 level of confidence" means.) In most of the other cases, the fit of the "no configuration effect" model was very close indeed (for eleven of the chi square tests, $p<.90$). In other words, given knowledge of the percentage incarcerated for a given cell score, and of the numbers of offenders displaying a particular configuration, one can almost always predict very closely how many of those offenders will be incarcerated. Almost invariably, rounding the expected value to the nearest integer produced an absolutely correct prediction.

Even though small, differences in percentages incarcerated among different configurations might perhaps be useful, if those differences were consistent; if, say, those with configuration (0 0 +1 -1 -1) were consistently more likely to be incarcerated than those with (0 +1 0 -1 -1), no matter what offense they were being sentenced for. Even this, however, is not the case. For each offense included in the guidelines, configurations are listed in rank order of percentage incarcerated; but those orders differ widely across offense types. For example, the configuration (0 +1 0 -1 -1) has the highest rate of incarceration among burglars (42 percent), but the lowest rate among those sentenced for assault (zero percent).

To investigate this matter, rank correlation coefficients (Kendall's τ) were calculated between configuration

orderings for the seven types of offenses with 48-cell matrices, for cell scores from -1 to +3. (Since cell scores of -2 and +4 can only be obtained in one way, the question of configuration order does not arise.) If the configurations were to make a consistent difference across types of offense, these coefficients should mostly be positive, and preferably fairly strong. Table 4.7 shows, however, that this is not the case. Of the 105 coefficients calculated, more are negative than positive; the median τ is about -.05. As the table shows, the 105 coefficients tend to cluster around zero, being mostly either low positive or low negative. Significance tests could be carried out for 63 of the 105 coefficients; of these 63, only six attained the .05 level of significance, and five of those were negative. There is thus no general ordering of configurations across offense types, for any cell score.

Insert Table 4.7 here

The conclusion seems clear: the configurations in the New Jersey sentencing guidelines add almost no useful information, and can safely be ignored without substantially reducing the descriptive or predictive power, or the utility, of the guidelines. As we shall see in the next section, differences in cell score do strongly and consistently affect the probability of incarceration, across all offense types. But it does not matter much, if at all, how a particular cell score is obtained.

Deletion of the detailed configurations would do much to simplify the guidelines themselves. Instead of 48-cell matrices (for most offenses), only seven cells would be needed, to convey substantially the same information to the sentencing judge.

What About Cell Scores?

The last section of this paper reported the finding that the cell configuration does not lend predictability to the sentence decision. Cell score, however, is a powerful predictor of this decision. The percentage of offenders that were incarcerated was computed for each cell score of each offense group. The entire 10,629 case sample was used as the data base for these computations.

Cell score was then used as the predictor of whether or not an offender would be incarcerated. It was found that as the cell score value increases, so too, does the likelihood of incarceration. This particular fact is verified out by the

Table 4.7: Rank Correlation Coefficients (Kendall's τ) Between Orderings of Configurations for Seven Offense Types in the New Jersey Sentencing Guidelines

Magnitude of Coefficient:	
+.91 to +1.00	
+.81 to +.90	
+.71 to +.80	
+.61 to +.70	II
+.51 to +.60	IIIIII
+.41 to +.50	
+.31 to +.40	IIIIIIIII
+.21 to +.30	IIII
+.11 to +.20	IIIIII
+.01 to +.10	IIIIIIIIIIIII
Zero	IIIIIIIII
-.01 to -.10	IIIIIIIIIIIIIIII
-.11 to -.20	IIIIIIIIIIIII
-.21 to -.30	IIIIII
-.31 to -.40	IIIIII
-.41 to -.50	II
-.41 to -.60	IIIII
-.51 to -.70	II
-.61 to -.80	I
-.81 to -.90	
-.91 to -1.00	I

Median value of τ = -.05.

(NOTE: Tests of significance could only be carried out for 63 of the 105 coefficients shown in the above diagram. Six of those 63 coefficients were statistically significant at the .05 level of confidence (see text).)

bivariate regression computed, the results of which are shown in Table 4.8. Cell score is an excellent predictor of the percentage of offenders who will be incarcerated within particular offenses.

Insert Table 4.8 here

A cursory reading of Table 4.8 reveals that the lowest R square, .599, was found for the regression of percent incarcerated on cell score for the total sample of cases. The R square coefficients for specific offense groupings were, as a matter of fact, higher than that found for the total sample in all instances. The lowest R square value found for a specific offense group was that of .791 for the Fraud offenses, while the two highest R square values, .989 and .984, correspond respectively to the Forgery and Assault offense groups. Each of the R square values computed, it should be noted, was significant at least at the .01 level. These regressions support the contention that one can very accurately predict the rate of incarceration, once the value of the cell score is known.

That the R square coefficient is much lower in value when offenses are combined and viewed as one group than when the coefficient is computed for each offense group separately is largely due to the different base incarcertion rate for different offense types. The A-intercept values listed for each offense's regression in Table 4.8 make this fact quite clear. Fifty-three percent of Robbery offenders who have rated the lowest possible cell score will be incarcerated as the norm; for persons convicted of Drug Possession and rating the lowest cell score, incarceration will result only for ten percent of the cases.

Simplifying Cell Score

The items of information incorporated into a predictive model usually are selected by following two general rules of thumb. One would like, first, for each of the items, or factors, to be highly associated with the decision that is being predicted; secondly, the factors should be relatively uncorrelated with each other. The model should, in other words, utilize the most efficient, independent factors to explain the variation in the decision, and thus, to predict that decision. The New Jersey sentencing guidelines are only moderately successful in fulfilling both of these requirements.

Table 4.8: Regression of Percent of Cases Incarcerated on Cell Score

Offense Group	R^2	Significance of R^2	Correlation Coefficient	A Intercept	B
Total Sample (N = 10,629)	.599	.00001	.77	26.47	10.88
Break & Enter (N = 2,152)	.959	.00006	.98	27.22	12.12
Robbery (N = 1,097)	.975	.00002	.98	52.76	11.03
Assault (N = 899)	.984	.00001	.99	28.13	14.29
Drug Sale (N = 1,236)	.884	.00081	.94	32.01	9.20
Drug Possession (N = 1,289)	.852	.00150	.92	9.68	7.94
Larceny (N = 993)	.915	.00037	.96	20.92	10.64
Weapon (N = 1,134)	.945	.00057	.97	12.82	13.97
Fraud (N = 1,067)	.791	.00367	.89	13.73	7.64
Forgery (N = 401)	.939	.00023	.99	15.44	16.84
Rape (N = 181)	.849	.01310	.92	52.94	19.70
Lewdness (N = 180)	.960	.01018	.98	16.70	18.18

The five offender variables' correlations with the incarceration decision usually ranged from low or negligible to moderate in magnitude (±.01 to ±.40) as shown in Table 4.9. When offenses were not differentiated, Criminal History was the most highly correlated variable; Amenability to Non-Custodial Supervision was the second highest correlate. All five of the offender attributes related significantly to the in/out decision. Criminal History was also the highest correlate for eight of the separate offenses; Amenability to Non-Custodial Supervision was the next strongest correlate for seven of the eleven offenses. Following the rule of thumb mentioned above, one would expect that these five offender scores be highly correlated with the decision; this was not found to be the case. The five attributes were, however, significantly associated with the decision for eight of the offense types (in part, a function of the Ns involved) even though their absolute value was low or moderate.

Insert Table 4.9 here

The second rule of thumb requires that the factors not be correlated with each other -- that is to say that the attributes should not be strongly associated with each other and, thus, should contribute independently to the decision. The correlations calculated were generally found to be negligible; thus, the guidelines were moderately successful in this respect. The only two scores that were highly correlated in the total sample were Criminal History and Amenabiliy to Non-Custodial Supervision (rpb = +.51). While the strength of the coefficients altered somewhat when offense types were analyzed separately, the general pattern was not refuted, Criminal History and Amenability to Non-Custodial Supervision were still the most highly intercorrelated scores. Two additional strong correlations were discovered when offense types were analyzed separately: for Robbery offenses, a moderate correlation appeared between Actions Since Arrest and Community Background; a similar moderate correlation existed between the Exacerbating Factor and Amenability variables in Drug Possession cases (see Table 4.10). Yet, while the coefficients appear to be of low magnitude, a large number were found to be significant -- 55 out of 88 possible coefficients were signficantly correlated.

Insert Table 4.10 here

The original cell score used in the New Jersey sentencing guidelines made use of information from <u>all five</u> offender

Table 4.9: New Jersey Sentencing Guidelines Correlation of Offender Attribute Scales with the Incarceration Decision

Coefficient	Total	B & E	Robbery	Assault	Drug Sale	Drug Poss.	Larceny	Weapon	Fraud	Forgery	Rape	Lewdness
+.51 to .60												
+.41 to .50												
+.31 to .40	1	1		1			11	11		1	1	
+.21 to .30	11	1	11	11	1	11			11	1	11	111
+.11 to .20	11	11	111	11	111	111	11	1	1			
+.01 to .10		1			1		1	1	11	1		
.00												
-.01 to -.10												
-.11 to -.20												
-.21 to -.30												
-.31 to -.40											1	
-.41 to -.50												
-.51 to -.60												
P ≤ .01	5	5	5	5	5	5	5	3	4	2	4	3
No. Correl.	5	5	5	5	5	5	5	4	5	3	4	3
N	10629	2152	1097	899	1236	1289	993	1134	1067	401	181	180

Table 4.10: Pearson Correlations Between Factors

Total Sample (N = 10,629)

Coefficient	Over All Offenses	B & E	Robbery	Assault	Drug Sale	Drug Poss.	Larceny	Weapon	Fraud	Forgery	Rape	Lewdness
+.51 to .60	1	1					1		1			
+.41 to .50			11		1	11		1				
+.31 to .40				1					1			
+.21 to .30	11		1	11		1		1	1		1	1
+.11 to .20			11		11	11	11	1	1111	1	11	1
+.01 to .10	1111111	11111111	111	1111	11111	111	11111					1
0		1	1				1	1				
-.01 to -.10			1	111	11	11	1	11	111	11		
-.11 to -.20											111	
-.21 to -.30												
-.31 to -.40												
-.41 to -.50												
-.51 to -.60												
P ≤ .01	9	6	6	5	9	7	4	3	8	1	4	2
No. Pairs	10	10	10	10	10	10	10	6	10	3	6	3
N	10,629	2152	1079	899	1236	1289	993	1134	1067	401	181	180

attribute scores. It is debatable, in light of the findings reported above, whether all five offender scores are required in order for the model to predict the incarceration decision efficiently. If the model is able to generate reasonably similar results when the offender scores are used either in different combinations or are deleted, the guidelines would gain simplicity.

The prediction of the incarceration decision for a random subsample of cases using the original cell score yielded a multiple correlation coefficient of .44. We then constructed five alternate cell scores (called Modified cell score No. 1-5). All of the modified cell scores differ not only from the original cell score, but also from each other. Different combinations of attributes have been included in the newly formed scores; attribute scales have been successively omitted from each of the scores, down to the final modified cell score (No. 5). Modified cell score No. 5 consists only of the Criminal History attribute score, and omits all other information about offenders.

Bivariate regression equations were computed to identify that modified cell score that would best account for variation in the decision to incarcerate. The results of these analyses, presented in Table 4.11, show that the predictive power of the model is only slightly compromised by the successive exclusion of certain items of information; in fact, the modified cell score consisting of the Criminal History attribute alone is able to explain the incarceration decision almost as well as the original five-factor cell score.

Insert Table 4.11 here

The original cell score, as well as the five modified versions that have been subsequently developed, can be viewed as predictive instruments -- devices to predict the decision to incarcerate. The Mean Cost Rating (MCR) statistic, originally proposed by Duncan, Ohlin, Reiss, and Stanton (1953), is used often in criminological research to test the strength of prediction instruments. The Mean Cost Rating computed for each of the predictive devices here, the original and modified cell scores, support our finding that little predictive power has been lost by the omission of attributes from the cell score. The MCR statistics that resulted from our computations ranged in value from .05 for the original five-factor score to .40 for the model based solely on the Criminal History score.

Table 4.11: Stepwise Regression of Original and Modified Cell Scores on the Incarceration Decision (N=2,118)

	Multiple Correlation Coefficient	R^2	Change in R^2	Correlation Coefficient	B	Bet
Original Cell Score	.44	.20	.20	.44	.139 constant .211	.44
Modified Cell Score #1	.44	.19	.19	.44	.168 constant .241	.44
Modified Cell Score #2	.42	.18	.18	.42	.172 constant .191	.42
Modified Cell Score #3	.42	.17	.17	.42	.185 constant .149	.42
Modified Cell Score #4	.40	.16	.16	.40	.196 constant .083	.40
Modified Cell Score #5	.39	.15	.15	.39	.227 constant .138	.39

Cell Score Factors

	Criminal History	Amenability to Supervision	Exacerbating Factors	Community Background	Actions Since Arrest
Original Cell Score	✓	✓	✓	✓	✓
Modified Cell Score #1	✓		✓	✓	✓
Modified Cell Score #2	✓		✓	✓	
Modified Cell Score #3	✓		✓		✓
Modified Cell Score #4	✓		✓		
Modified Cell Score #5	✓				

Subsequent work on the concept and proof of the MCR statistics by Lancucki and Tarling (1978) has shown that it is possible to transform the MCR statistic into a Z score. Thus, we next computed the associated Z scores for each of the predictive models. As the sample size was so outrageously large for this particular data set, the Z scores are somewhat unrealistic. More to the point, we were interested in determining whether or not the change in the value of Z, moving from a fully saturated to a single attribute model, would become significant. A formula has been derived to test the significance of the difference in Z from one model to the next.[9] Tables 4.12 through 4.17 show the values computed for each of these statistics: the Mean Cost Rating (MCR), the Z score (Z), the new Z score that represents the change in the Z score from the preceding model (Zz0-Z1), and the probability of the Z score associated with the change over models.

We mentioned before that the MCR statistics are only slightly reduced when the various offender attributes are removed from the saturated model. It also has been mentioned that the Z scores do not undergo drastic changes during the same process. The final statistics that have been computed -- those of the new Z scores representative of the change in the values of Z when moving from the original score to less complicated ones -- also do not evidence a dramatic reduction in predictive power when items of information are omitted. It is not until we reach the last stage of the analysis, when the Criminal History attribute is used as the sole predictive attribute, that a significant change in the Z score is discovered. The Z score of the change in predictive power from Modified cell score number 4 to Modified cell score number 5 is equal to 1.74, significant at the .041 level. These statistics are presented in Tables 4.12-4.17.

Insert Tables 4.12-4.17 here

The implications of these tables need spelling out. Of the total of 10,629 cases in the construction data, 61 percent were incarcerated. Thus, told merely that an offender was included in that group of 10,629 cases, one's best bet (and it would obviously not be very good) would be that that offender was sent to jail or prison. A more accurate prediction can be made if, in addition, the offender's cell score is known. As Table 4.12 shows, using the original cell scores, the construction data can be broken into sub-groups with incarceration rates ranging from 88 percent down to 23 percent. Predictions made with that extra knowledge would still be far from perfect, but they would be a substantial improvement over the base rate for the whole 10,629 cases;

Table 4.12: MCR Analysis of Original Cell Score

Sentence Imposed	Original Cell Score							Total
	-2	-1	0	1	2	3	4	
Non-Custodial	197	1,139	1,596	1,515	1,161	714	214	6,536
Incarceration	26	134	377	642	1,023	1,181	710	4,093
Total	223	1,273	1,973	2,157	2,184	1,895	924	10,629

Score Factors

Criminal History
Amenability to Supervision
Exacerbating Factors
Community Background
Actions Since Arrest

MCR Statistic = .50
Z_o = 43.9325

Table 4.13: MCR Analysis of Modified Cell Score No. 1

Sentence Imposed	Modified Cell Score No. 1						
	-2	-1	0	1	2	3	Total
Non-Custodial	215	1,241	1,884	1,788	1,122	286	6,536
Incarceration	30	158	511	993	1,473	928	4,093
Total	245	1,399	2,395	2,781	2,595	1,214	10,629

Score Factors

Criminal History
Exacerbating Factors
Community Background
Actions Since Arrest

MCR Statistic = .49
Z_1 = 43.2833
Zz_o-z_1 = .72
p = .236

Table 4.14: MCR Analysis of Modified Cell Score No. 2

Sentence Imposed	Modified Cell Score No. 2					
	-1	0	1	2	3	Total
Non-Custodial	954	1,775	1,914	1,422	471	6,536
Incarceration	97	366	917	1,609	1,104	4,093
Total	1,051	2,141	2,831	3,031	1,575	10,629

Score Factors

Criminal History
Exacerbating Factors
Community Background

MCR Statistic = .47
Z_2 = 42.0546
Zz_1-z_2 = .98
p = .164

Table 4.15: MCR Analysis of Modified Cell Score No. 3

Sentence Imposed	Modified Cell Score No. 3					
	-1	0	1	2	3	Total
Non-Custodial	410	1,879	2,148	1,602	497	6,536
Incarceration	60	323	931	1,643	1,154	4,093
Total	470	2,202	3,061	3,245	1,651	10,629

Score Factors

Criminal History
Exacerbating Factors
Actions Since Arrest

MCR Statistic = .45
Z_3 = 40.3415
Zz_2-z_3 = 1.38
p = .084

Table 4.16: MCR Analysis of Modified Cell Score No. 4

Sentence Imposed	Modified Cell Score No. 4				
	0	1	2	3	Total
Non-Custodial	1,609	2,209	1,932	786	6,536
Incarceration	213	716	1,762	1,402	4,093
Total	1,822	2,925	3,694	2,188	10,629

Score Factors

Criminal History
Exacerbating Factors

MCR Statistic = .43
Z_4 = 39.0739
Zz_3-z_4 = 1.07
p = .142

Table 4.17: MCR Analysis of Modified Cell Score No. 5

Sentence Imposed	Modified Cell Score No. 5			
	0	1	2	Total
Non-Custodial	2,506	2,209	1,821	6,536
Incarceration	502	1,009	2,582	4,093
Total	3,008	3,218	4,403	10,629

Score Factors

Criminal History

MCR Statistic = .40
Z_5 = 37.5692
Zz_4-z_5 = 1.74
p = .041

that is what the MCR statistic shows. Successively taking items out of the cell score reduces the number of sub-groups which it defines -- until we reach the irreducible minimum defined by Criminal History alone (values of 0, +1, and +2 -- see Table 4.17). Our predictions of incarcerations will necessarily be less accurate than those using the seven-category original cell score. But they will not be all that much less accurate: Criminal History by itself splits the construction data into groups with incarceration rates ranging from 83 percent to 41 percent. That is what the changes in MCRs show.

Combining Offense Categories

The presumed goal of any sentencing guidelines model is to provide as much information as is useful to the decision-maker. The principle involved here is that of maximum utility and minimum "information overload". Many of the guidelines that have previously been developed (Gottfredson, et al., 1976) have aspired to this goal by combining offenses into groups of similar offenses resulting in models that have been labelled "generic". Generic models for sentencing guidelines generally consist of four parts; grids are developed for violent, property, drug, and "all other" types of offenses. Another method that has been used to simplify guidelines matrices has been that of scaling offenses into groups according to the perceived severity of the offense. An example of this type of model is that used by the United States Parole Commission. As we explained earlier in Chapter 3 (Gottfredson, Wilkins, Hoffman and Singer, 1974).

The table that follows presents the results of our attempt to simplify still further the guidelines matrix of the New Jersey sentencing guidelines. The table contains the results of regression equations, computed for groups of offense categories (and using incarceration versus non-incarceration as the dependent variable). The groupings have been chosen on the basis of similar intercepts (base incarceration rates) and similar slopes. The logic underlying this method of offense combination is that offenses with similar base incarceration rates and similar slopes (i.e., increases in the base incarceration rate for each original cell score) are also similar in terms of their perceived seriousness. As is evidenced by the R square values for each of the groupings, the combinations of offenses chosen do not very substantially reduce the predictive power of the equations. There is merely a slight loss in the predictive ability of the model; but instead of eleven offense categories, we now have only five categories to work with.

Some people (in particular, some judges) might object to our "empirical" grouping of offense categories on the basis of similar patterns of disposition, rather than on some substantive similarity in the offenses thus grouped. For example, why should the offense category of Lewdness be grouped together with the category of Forgery in Table 4.18, rather than with (say) Rape? Why should Drug Sales be lumped in with Breaking and Entering (rather than Drug Possession), merely because of similar dispositional patterns? Our answer is that *if* guidelines are meant to give information about sentences in the past, and *if* two groups of otherwise dissimilar offenses have had very similar dispositional patterns, *then* those groups ought to be combined. We could of course provide an extra row for Drug Sales, or group it together with Drug Possession; but in the first case the resulting table would be more cumbersome, and in the second it would be less accurate in predicting incarceration (which is, of course, the only dependent variable we are now considering with this reconstructed data set).

Insert Table 4.18 here

What is the *maximum* simplification which we can obtain by rearranging the New Jersey guidelines, and what does it cost us in terms of accuracy in predicting incarceration? Tables 4.19 and 4.20 speak to this question. Table 4.19 uses all eleven offense groups, and (where possible) all seven levels of cell score; it thus has 77 cells, of which some are structurally empty. It will be seen that this table splits the construction data into groups with incarceration rates ranging from 0 percent to 91 percent; and that several of the cells have very similar rates. Table 4.20 uses our five combinations of offense groups as rows, and the three values of Criminal History alone as columns; this fifteen-cell table splits the data into sub-groups with incarceration rates ranging from 7 percent to 79 percent. Of course, there is an inevitable trade-off between simplicity and accuracy of prediction. But, first, both Tables 4.19 and 4.20 improve accuracy of prediction (of incarceration) over the base rate of 61 percent for the whole of the construction data; second, the fifteen-celled Table 4.20 is not *all that much* worse than the 77-celled Table 4.19 (MCR for Table 4.19 is .62; MCR for Table 4.20 is .55; a change of this size would be expected just because of the number of cells in the larger table). Third, both tables are a great deal easier to work with than the 792-celled table which would result from using the original New Jersey guidelines in a comparable format (eleven offense groups with all configurations, versus incarcerated or not incarcerated).

Table 4.18: Regression of the Percent of Cases Incarcerated for Grouped Offenses on Cell Score (N=10,629)

Offense Groups	R^2	Significance of R^2	Correlation Coefficient	A-Intercept	B
Breaking & Entering and Drug - Sales (N = 3,388)	.91	.00001	.96	29.61	10.66
Larceny and Weapons (N = 2,127)	.91	.00001	.95	18.07	11.79
Robbery Assault and Rape (N = 2,177)	.76	.00001	.87	43.48	13.01
Forgery and Lewdness (N = 581)	.98	.00001	.99	16.07	17.14
Drugs - Possession and Fraud (N = 2,356)	.81	.00001	.90	11.70	7.79

Insert Tables 4.19 and 4.20 here

Conclusions and Implications

The focus of our discussion thus far has been an analysis of the technical construction of the New Jersey sentencing guidelines. Cell configurations have proven to have little value in the prediction of the sentence decision; but cell scores are reasonably good predictors (or descriptors) of the "in-out" decision in the guidelines construction data. We have also shown that a reduction in the number of offender attributes included in the cell score does not drastically reduce the predictive ability of the guidelines. We have thus shown that the New Jersey sentencing guidelines could certainly be very much simplified without a loss of valuable information.

Such a statistical evaluation of a decision-making model is, we think, important in its own right. But it also raises a number of important substantive questions about the whole concept of sentencing guidelines. These questions can be put into three broad categories. The first relates to the construction of guidelines; the second to their content; and the third to their use. In conclusion, we briefly consider some issues in those three categories.

The Construction of the Guidelines

Why were the New Jersey guidelines constructed in the way that they were, and why were they presented in such a complicated form? Though we cannot prove this, we suspect that the decision resulted at least in part from beliefs about the marketability of sentencing guidelines, and a belief that (at least in New Jersey) they could not be "sold" to the judiciary unless they were "offense-specific", used several offender attributes, were highly disaggregated, were based on all cases sentenced in a year (rather than a sample), and so on.

Decision-makers in general, and judges in particular, seem to believe that elaborate models are required if their decision-making processes are to be accurately represented. A sentencing judge makes highly public decisions that can have drastic effects on people's lives; inevitably, criminal sentences restrict people's liberty to at least some degree. While a simpler model may describe and predict sentencing practice as well as, or better than, a more complex one, judges may feel more comfortable with the complex version. They may believe that it relies on more information about

Table 4.19: Percent of Cases Incarcerated by the Cell Score for Each Offense (N = 10,629)*

Offense	Cell Score						
	-2	-1	0	+1	+2	+3	+4
Fraud	8 (86)	7 (300)	9 (280)	13 (155)	21 (130)	34 (83)	58 (33)
Drug Possession	3 (31)	1 (151)	4 (230)	10 (253)	22 (245)	27 (248)	52 (131)
Weapon	-	6 (187)	11 (315)	21 (288)	36 (203)	51 (115)	77 (26)
Lewdness	-	3 (33)	10 (59)	34 (61)	56 (27)	-	-
Forgery	-	2 (56)	13 (75)	29 (111)	50 (131)	68 (28)	-
Larceny	11 (18)	5 (100)	17 (161)	23 (177)	39 (221)	58 (207)	68 (109)
Drug Sale	26 (27)	16 (165)	28 (286)	34 (296)	48 (239)	64 (140)	74 (83)
Breaking & Entering	10 (31)	16 (142)	24 (273)	30 (419)	50 (525)	66 (556)	81 (206)
Assault	0 (9)	20 (61)	23 (149)	38 (191)	57 (197)	76 (187)	87 (105)
Rape	0 (2)	42 (26)	71 (56)	77 (66)	81 (31)	-	-
Robbery	32 (19)	36 (52)	54 (89)	66 (140)	78 (235)	89 (331)	91 (231)

*N's appear in brackets below the percent incarcerated.

MCR Statistic = .62

Table 4.20: Percent of Cases Incarcerated by the Criminal History Score for Offense Groups (N = 10,629)*

Offense Group	Criminal History Score		
	0	+1	+2
Drug Possession Fraud	7 (959)	12 (716)	32 (681)
Larceny Weapon	9 (648)	24 (689)	50 (790)
Lewdness Forgery	9 (212)	32 (143)	46 (226)
Breaking & Entering Drug Sale	22 (722)	36 (1181)	61 (1485)
Robbery Rape Assault	42 (467)	59 (489)	79 (1221)

*N's appear in brackets below the percent incarcerated
MCR Statistic = .55

specific offenders and offenses, and therefore more accurately leads to specific sentences.[10] Confronted by a simple guideline matrix (like our Table 4.20 above), it is likely that judges -- in New Jersey and elsewhere -- would tend to feel that it was no more than a caricature of the elaborate deliberations which precede their decisions.

However that may be, it is clear that those responsible for constructing the New Jersey guidelines went out of their way to collect data on every conceivable item of information which might in some way or another have been relevant to sentences in the construction data, and to consider a great many ways in which relevant items of information could be combined. According to McCarthy (1978:21) "project staff experimented with perhaps hundreds of variables in multiple regression equations...of about 6-10 variables each." The 842 variables in the codebook were boiled down to "a list of consistently powerful variables" which had passed a "rigid test for statistical significance", i.e. the .05 level of confidence (McCarthy, 1978:22). These variables were combined into offender attribute scores which differed, as we have described, for different types of offenses; and different combinations of the scores themselves were used in some cases. Though the guidelines manual could be clearer about the exact procedures used, it seems clear that the objective was to maximize prediction of the decision to incarcerate. (See McCarthy, 1978:19-29.)

So far as we are able to tell, however, no effort of any kind was made to validate the results thus obtained, or to assess the stability of any of the relationships which appeared in the analyses of the 10,629-case construction data set. It is clearly not that they could not have sampled, with a data set of that size; indeed, for many of the analyses reported in this paper we worked with samples of a few hundred, without getting different results from those later obtained in analyses on the whole of the data set. (In view of the size of that data set, it is not surprising that so many relationships showed up as statistically significant at the .001 level; what is astonishing is that more did not.) Given the lack of validation, and the ad hoc method of constructing offender scores, it is impossible to stifle the suspicion that the data were badly overfitted, and that many of the associations found would not reappear if the analyses were repeated on another set of cases.

The Content of the Guidelines

Statistically-derived guidelines also raise a number of important legal and philosophical issues, of course, For instance, what sort of items should be excluded from

guidelines, even if found to be associated with sentences in the past? Race, sex, and ethnicity seem obvious candidates for exclusion; but what about items (such as employment history) which may be highly inter-correlated with an illegitimate item such as race? Items included should also be compatible with the dominant rationale for punishment espoused by the legislature and/or judiciary. Thus, "prognosis for recidivism" might be legitimate in a system which emphasized rehabilitation as an objective of sentencing; but it clearly would not be in a system aimed at "just deserts." It should be noted that (like most other sentencing guidelines so far developed), the New Jersey guidelines appear to take no account of such questions of principle, in their choice and definition of factors used. (There may be some principle according to which the differences in definition of e.g. Criminal History, which we noted earlier in this chapter, can be rationally justified; but it is not at all clear what that principle might be.) [11]

The Use of the Guidelines

The original aim of the developers of decision-making guidelines, as applied to sentencing, was to increase the equity of sentences through the articulation and use by judges of general sentencing criteria. To oversimplify the general strategy (which we outlined in great detail earlier in Chapter 3), a court (or jurisdiction) studied its past sentencing decisions, statistically derived functions that identified the likely predictors of these same decisions, and formulated a statement of sentencing policy consistent with the research results. The policy statement (with revision as required by periodic assessment of the reasons given for guideline deviations) would then serve as an aid to individual sentencing judges -- a tool designed to provide "guidance" to the decision-maker. (Cf. Gottfredson, Wilkins and Hoffman, 1978.)

The guidance in question has mainly been concentrated on only two questions: (1) Should this offender be incarcerated? (2) If incarcerated, for what period of time? An answer to the first question requires stipulating that sentences for certain groups of offenders (as defined by attributes such as current offense and prior record) should be either "In" or "Out"; an answer to the second requires a recommended range of sentence length (for example, 24 to 36 months), with a stipulation that sentences outside the range should be justified by special reasons. (See, for example, Gottfredson, Wilkins and Hoffman, 1978:119-127.)

But the New Jersey guidelines are not of the usual sort; and as a result, we suspect, they fail in several ways to

provide much guidance. Individual "cells" (i.e. rows) of the guidelines are not labelled as "In" or "Out"; nor do they stipulate a range of duration of confinement within which no special reason is needed for justification. Instead, they merely indicate the numbers (not even the percentages) of offenders in the construction data who were sent to each of four types of institutions (and, by implication, the percentage not incarcerated at all); and they give merely the median terms of those incarcerated. Those medians differ widely, for obvious reasons, among the four types of institution; and no overall median term is given. Thus, the length of term to be imposed is structured mainly by a choice as to where the offender should be sent (jail versus prison, for instance) and no guidance at all is given for this decision. Further, even though the median sentence among the four institutions would differ, the guidelines could still have constructed estimated sentence ranges had the actual time (sentence, minus parole elibibility, good time, and any known parole board rules) been used as the dependent variable. This sort of procedure was used by the Massachusetts guidelines project with some succes as we will discuss later in Chapters 7 and 8.

One may argue that, as the frequencies of various dispositions are indicated for each column of the guidelines, some guidance is given. But the interpretation of those figures is left entirely to the discretion of the judge. That interpretation may not be problematic where, say 70 or 80 percent of cases are incarcerated or are not. But surely it becomes a problem where the choices are more evenly split? Should a judge interpret a row in which 51 percent of offenders have in the past been incarcerated as an "In" row for which incarceration should be the normal disposition in the future? Moreover, what about choice of institution? This too may be clear, if, say, the overwhelming majority of those incarcerated in the past went to one institution (e.g. state prison). But this is often not the case. The bottom row of Table 4.3, for example, shows that of the 525 persons convicted of Breaking and Entering, exactly half were incarcerated; of those incarcerated, 43 percent received county jail terms of less than 12 months; about 30 percent (presumably young adult offenders) were sent to Yardville for indeterminate terms; and about 25 percent were sent to state prison.

Nor, assuming that a judge can be guided as to choice of institution, does the median by itself give much guidance as to length of term. Consider four offenders with sentences of 2, 49, 51, and 98 months; and four more with sentences of 47, 49, 51 and 53 months. For each group, the median term is 50 months (a figure which appears in neither set). But surely a

judge might want to know something about the relative variability of the two groups of cases -- in order to decide how much of a deviation from the median might be justified? Of course, the whole range might not be useful for that purpose; in our example, the sentences of 2 and 38 months might be of the unjustifiable variety that it is the purpose of the guidelines to get rid of. But the median by itself does not help much for that purpose either. (It is especially unhelpful in the case of the state prison sentences, for which the biggest range exists.)

The use of guidelines to make the "In"-"Out" decision raises a number of problems; in fact, suppose that in a certain row of the New Jersey guidelines, 70 percent were incarcerated. How is a sentencing judge to use that information in particular cases in the future? We might imagine him throwing a biased coin, which was arranged to turn up "Heads" 70 percent of the time; but such a procedure is not likely to appeal to anyone interested in doing justice, or structuring discretion, in individual cases. Alternatively, we might declare that that row was an "In" row; but if judges in the future were regularly to interpret it as such, the percentage incarcerated would surely tend to 100 percent. We might instruct judges that an In/Out ratio of 70:30 (or 80:20, or perhaps 90:10) would raise some sort of presumption in favor of the preponderant disposition; but -- apart from the vagueness of such an idea -- there is obviously no purely statistical method of determining when such a presumption should arise, what its strength should be or what sort of additional information will justify deviation. What would be needed, in such a case, would be further information about those attributes which tended to discriminate between the 70 percent who were "In" in the construction data, and the 30 percent which were "Out". But even that requirement is not guaranteed to do the trick, since there may be nothing at all which discriminates between the two groups. It may be that offenders (in the construction data) who found themselves in that row of the guidelines were absolutely homogeneous, in all relevant respects, and that the 70:30 split between "In" and "Out" was entirely a consequence of variation among judges rather than of differences between cases. Given this possibility, and the structure of the New Jersey guidelines, how will it be possible for judges, researchers or anyone else to know when judges have deviated from the guidelines in the future? [12]

The unauthorized analysis of the New Jersey guidelines presented in this chapter has, of course, its limitations. Without access to the original data from which the guidelines were constructed, we have no way of ascertaining how similar cases in each subgroup really are; we thus cannot know how

"similarly situated offenders" were treated before the guidelines were introduced. Nor do we know -- indeed, nor does anybody know -- how "similarly situated offenders" have been dealt with since the guidelines were implemented in October 1978. So far as we are aware, no attempts are being made to monitor the operation of the guidelines, or to see how often they are being complied with in practice (assuming one can define "compliance"). Nor is it clear how the guidelines have been affected by the new penal code introduced in New Jersey at the beginning of September 1979.[13]

What is clear is that the New Jersey guidelines -- over-hastily developed and implemented, using a crude statistical methodology, without any provision for policy statement, monitoring or revision -- are a particularly illustrative example of the problems and issues inherent in guidelines development. The actual effect of these guidelines on sentencing practice in the state could not be statistically estimated by this project (because of the lack of cooperation of the New Jersey judiciary and the guidelines project's staff), but that the guidelines pose at least a clear potential for an adverse effect because of the method of their development and because of the philosophical controversies raised by their content is abundantly clear. It is unlikely however that the potential adverse effects of these guidelines will ever be known since, as we earlier noted, the new penal code -- the Code of Criminal Justice -- went into effect in September 1979 and established more uniform classifications for crimes and penalties than had existed under the old code. In addition, the new code called for the establishment of a sentencing commission -- comprised of legislative, judicial, correctional, and public representatives -- to further refine the sentencing provisions of that legislation. At the time of this writing, new sentencing guidelines are supposedly being developed under the auspices of this commission with the assistance of staff from the earlier New Jersey guidelines project.[14] It is not yet known how the new guidelines will differ in development, content, or use from the guidelines discussed here.

Notes to Chapter 4

[1] A typographical convention must be noted. Whenever we refer to those attributes as they are defined in the guidelines, we write them with capital letters. This has the unfortunate effect of making the paper look rather like a chapter from Winnie the Pooh; but it is necessary since we wish to distinguish e.g., Criminal History as defined in the guidelines from criminal history in the ordinary sense of that term.

[2] The most recent study that used this extensive data base was released on September 4, 1979, Report of the Sentencing Guidelines Project to the Administrative Director of the Courts on the Relationship between Race and Sentencing (1979).

[3] The logic of the program is extremely simple. A single cell of the 1,980-cell guideline matrix contains n cases cross-classified by offense type and up to five offender attributes, plus information on disposition; thus, in Figure 4.4 above, there are 23 cases with an offense of Breaking and Entering, a configuration of (+2 +1 0 0 -1), and a sentence to state prison (median 5 years). The program takes that information as input, and generates 23 identical case-level records each containing that information. A listing of this program, and/or a copy of the case-level data set produced by it, are available from the authors on request.

[4] "Any Offense" is defined by the guidelines manual as being inclusive of "disorderly persons or J.I.N.S. but excluding traffic-related violations."

[5] "Any crime" is defined as being of "misdemeanor level or higher."

[6] Offenders convicted of Forgery, Rape, or Lewdness were not given scores for this variable.

[7] For each of the first three variables included in the sentencing guidelines -- Criminal History, Amenability to Non-Custodial Supervision, and Exacerbating Factors -- a score of 0 indicated the absence of negative offender attributes; positive scores of +1 and +2 indicated the presence of negative attributes. The final two variables of Community Background and Actions Since Arrest reverse this procedure. A negative integer (score -1) is assigned if the offender has done something that would mitigate his sentence, such as obtain employment, enroll in school, or plead guilty. In other words, a score of 0 now connotes the presence of

negative offender attributes (as only +1 or +2 did for the earlier attributes) rather than the absence of such attributes. Of course, the same result could have been effected while still retaining the earlier scoring system; score 0 would obtain for those cases where the offender had present positive factors, and score 1 would be reserved for negative factors.

[8] Some may object to the use of significance tests with what are essentially population data (all cases sentenced in 1977) rather than a sample drawn under a model appropriate to the test. We agree that for some purposes this would be inappropriate; there is no point in trying to set confidence intervals around an estimate of a population parameter, if one has population data. Nonetheless, it is useful to have some handle on the question "How big within-group difference is e.g. 6 percent?"; this can be done by using tests like chi square,, if it is not entirely ludicrous to suppose that the population data in question could have been obtained as a sample under some appropriate sampling model, e.g. Poisson or multinomial. Suppose that the New Jersey guidelines developers had drawn a sample from all cases sentenced during 1967-77, and that all cases in that sample had in fact been sentenced in 1977. That improbable result would of course lead one to suspect the sampling method used, but it would scarcely invalidate the use of significance tests.

[9] The authors are grateful to our colleague Donald M. Barry, for the development of this formula:

$$Zz_1 - z_2 = \frac{(S_1 - C_1) - (S_2 - C_2)}{\sqrt{\text{Var}\ (S_1) + \text{Var}\ (S_2)}} = \frac{(S_1 - S_2) - (C_1 - C_2)}{\sqrt{\text{Var}\ (S_1) + \text{Var}\ (S_2)}}$$

The formula yields a second Z score that is the Z score value of the difference in predictive ability of the two models being compared.

[10] There is something in this belief; but not much. It is true that given the wide variety of types of crimes to be sentenced, and the wide ranges of circumstances in which those crimes may take place, there will always be some factors relevant to sentences in a very small number of cases, which would not show up as significant in an aggregate statistical analysis. (It is presumably to cater for these rare but important factors that judges are allowed to go outside recommended ranges in special cases: cf. Gottfredson, Wilkins and Hoffman, 1978). But the statement that "every case is unique" is obviously either false or tautologous.

[11] This matter is even more complicated, since judges do not need to adhere to the offender variables included in the guidelines, however these may be defined. In the introduction to the guidelines (McCarthy, 1978:33) it is stated that "if a crime does not possess the exacerbating factors listed as most influential for that crime, and therefore receives an exacerbating factor score of '0' a judge may find some 'other' factor which he feels nevertheless makes the crime a more serious one. He then might assign the higher score and move on to the new cell." Would that be a deviation from the guidelines? Or merely a rather special use of them? And if the latter, what is the point of having guidelines at all?

[12] According to the guidelines manual, "...the guidelines are still advisory only, and a judge may deviate in any manner they (sic) choose" (McCarthy, 1978:33). But is any deviation from the median, or from exactly proportionate assignment of cases to institutions or to non-incarceration, a "deviation" for this purpose?

[13] The only real effect that the new code could have had on the guidelines would have been to prohibit their use in cases where the new statutory penalty for any particular offense would have not been one of the guidelines sentences. But since the guidelines themselves were not mandatory prior to the code and so were not referred to by all of the judges in the state, we have no way of knowing how many cases would actually have been affected.

[14] The American University Criminal Courts Technical Assistance Project's Summer 1981 bulletin included the following notice:

> Guidelines are being developed under the auspices of the Sentencing Commission created by the new Criminal Code. (The old guidelines were developed under the auspices of the Supreme Court.) The new sentencing Commission's membership includes representatives from the Legislature, Judiciary and the Bar.

This notice, to our knowledge, has been the only public announcement of the activities of this commission.

Chapter 5: Prisoners' Perceptions of Sentencing in New Jersey

The effects of the implementation of sentencing guidelines on the sentencing procedure in the State of New Jersey could be assessed by objective statistical analysis if only the necessary data were available. As indicated in the preceding chapter, we were denied access to raw data on pre- and post-guidelines sentence dispostions; therefore, the only information we have on sentencing in New Jersey is based on pre-guidelines data reconstructed from the New Jersey sentencing guidelines.[1] Without before and after data, we can say nothing about the "real" effect of guidelines on dispositions, and we are left without a clue as to whether guidelines actually reduced disparity, enhanced fairness, increased sentence terms, or whatever.[2] Aside from the question of whether sentencing guidelines made an actual change in sentencing practices, however, there is also the question of whether sentencing guidelines were perceived to make a change in sentencing. In order to answer this question, we interviewed two samples of prisoners at Rahway State Prison in New Jersey.

Prisoners notoriously have been known to complain about the lack of fairness in sentencing, and they also seem to have some rather concrete notions about what could be done to enhance the fairness of the sentencing process (or, at least, what could be done to enhance the fairness of their sentences).[3] Furthermore, prisoners talk to each other about sentencing and about their own sentences and, therefore, they typically have a very keen idea about variations in sentence dispositions. Since variation in sentences for like offenses and offenders is considered by both prisoners and decision-makers to be the paradigm of unfairness in sentencing, and as guidelines are ostensibly developed to reduce unjustified disparities, then it should be the case that if sentencing guidelines actually do reduce disparity this change would be noticed by those parties who express an interest in fairness. Therefore, given prisoners' inherent interest in the enhancement of fairness in sentencing, it should be the case that an objective change in sentencing practice would be reflected in a similar change in prisoners' subjective perceptions of sentencing.[4]

Whether the impact of guidelines can be determined from pre- and post-guidelines perceptual data depends, of course, on the amount of knowledge and the validity of the knowledge of the sampled population. It may be the case that guidelines have done much by way of reducing disparity in New Jersey, but that prisoners have yet to be made aware of the change. Even though we do have data from a sample of prisoners drawn in

1980, it is possible that information about guidelines and their effect on sentencing had yet to trickle down to our sampled population. Furthermore, even if a sufficient number of prisoners were aware of guidelines, their perceptions of the impact of guidelines could be wrong. It may well be the case that prisoners notice a change for the better in sentencing and wrongly attribute this change to the implementation of guidelines (if, say, guidelines per se have had no effect); or prisoners may perceive no change when in fact dispositional data show a marked decrease in disparity. For these reasons, the data presented in this chapter should not be construed as a substitute for a measure of "real" change, for it may be that the perceptual data presented are not even correlated with the objective reality of guidelines impact. Nonetheless, perceptual data about sentencing guidelines are all that we have from the State of New Jersey and it is unfortunate that we are not able to substantiate the claims of prisoners through dispositional data analysis.

Despite the fact that the perceptual data we have do not answer the same questions that dispositional data would, there is much that can be said about the perceived impact of sentencing guidelines; more importantly, there is much information on perceptions of the sentencing process in general, perceptions of the seriousness of offenses, and perceptions of the severity of sanctions, that was obtained from our two samples of prisoners. And additionally, even though our data show that most prisoners were totally ignorant of the existence of sentencing guidelines and of their supposed use, we do have data that allow us to estimate what prisoners would think about fairness in sentencing if the New Jersey guidelines were in full-fledged operation, and if prisoners were aware of their use.

Method

Since this aspect of the project was designed to assess prisoners' perceptions of the fairness of sentencing before and after guidelines implementation, two samples of prisoners were selected: one sample was drawn from the total inmate population as of June 1979; the other sample was based on the population of July 1980. The two samples, randomly drawn a year apart, necessarily included some overlap, as a number of those interviewed in 1979 were still around in 1980. As well, the two samples, also because of the random selection, are not "pure" samples; that is, the 1979 sample is not strictly pre-guidelines, as much as the 1980 sample is not strictly post-guidelines. Our concern with sample selection was not, however, with obtaining only those who had received dispositions before guidelines in order to compare with another sample of those who had received guidelines

dispositions; rather we were concerned with how information about guidelines came into the institution via newly sentenced offenders, and with how that information was assimilated into prisoners' perceptions of the fairness of the sentencing process.

Rahway State Prison was selected as the site for interviewing. Rahway State Prison typically houses around 900 men; and it is one of three state prisons in New Jersey.[5] Institutional policy allows for prisoners to be classified as either maximum, medium, or minimum security inmates; with minimum security inmates usually being near release and generally being housed at a satellite unit or "camp".[6] Medium security inmates are considered by most guards and the Classification Department at Rahway as being little different from maximum security inmates (they are allowed a few more privileges) and therefore we did not distinguish them from the maximum security men. Sample selection in both 1979 and 1980 was obtained by dividing men into either minimum or maximum/medium status, and then randomly sampling within these two frames, with a higher proportion of minimum men selected to compensate for their relative rarity. The total sample selected in 1979 was 226; in 1980 it was 292.

The 1979 questionnaire was administered to inmates during the month of June, by a group of six trained interviewers; nine interviewers participated during the survey that was conducted in July of 1980.

Characteristics of the 1979 Sample

Background data for all of the 226 prisoners in the 1979 sample were obtained from classification files. Diminishing funds and time precluded us from gathering comparable data for the 1980 sample; however, as there is no reason to think that the characteristics of the two samples differed markedly, we may assume that the background information we have on the 1979 sample could describe our 1980 sample equally as well.

Of the 226 prisoners in the 1979 sample, 153 inmates responded, at least in part, to our survey. A similar response rate was obtained in 1980; of the 292 inmates in the original sample, 166 participated. (The final N used for analysis in both the 1979 and 1980 sample dropped to 146 and 157, respectively. The remaining persons were excluded from the data analysis due to inability to understand English or incompleteness of the questionnaire.)

For the 1979 sample, nearly every conceivable comparison was made between the respondents and the non-respondents on the basis of background data.[7] Contrary to expectations,

the respondents and the non-respondents of the 1979 sample did not differ significantly or substantially on any variable.[8] (The respondents in the 1979 sample were slightly more likely to be employed at the time of arrest than the non-respondents; however, since employment at time of arrest is such a poorly coded variable in institutional files -- if the defendant was in jail prior to trial he invariably was designated unemployed at time of arrest in probation files -- this slight difference is probably an artifact of the data and does not represent any true difference between the two groups.)

Since the respondents and the non-respondents were substantially the same on the basis of the background data collected, it seems that there is every reason to believe that a representative sample of the prisoner population participated in the 1979 survey; and it is probable that a representative sample was obtained in 1980 as well.

The 1979 respondents had an age range of 19 to 60 with a median age of 31, and had been, on average, incarcerated at Rahway State Prison for the current offense a little over two years at the time of interviewing. They had a median of 11 prior arrests and six prior incarcerations. There were 18 men (12 percent of the respondents) serving life sentences; for the non-lifers, the mean sentence being served was 10-13 years. Consistent with these sentence lengths, the majority of the respondents were serving time for either armed robbery or homicide (each accounted for 28 percent of the total N).

The minimum security men (n=47) and maximum security men (n=99) differed in predictable ways: minimum security men tended to be older, had served more time on their current sentences, had received a shorter disposition, had fewer prior convictions, and were convicted on fewer charges for the present incarceration. Despite the differences in background characteristics, however, responses to survey questions did not vary by institutional security status.

Questionnaire Design

Before the findings of the research are reported, it is necessary to discuss the design of the two questionnaires employed, for the results have meaning only insofar as they are answers to specific questions. Readers uninterested in such methodological details are advised to skip to the next section.

The 1979 respondents and the 1980 respondents were asked to respond to different questionnaires, as the 1980 questionnaire was modified and substantially shortened after experience with the 1979 questionnaire indicated that a number

of questions produced undecipherable results.[9] The bulk of both questionnaires, however, contained the same core items; and those items, or sections, of the original questionnaire that were particularly successful in the 1979 survey were expanded for the 1980 survey. Since the two questionnaires have notable differences, and since the 1980 questionnaire is more complicated in design, the two questionnaires and their design will be discussed separately.

The 1979 Questionnaire

The 1979 questionnaire was comprised of eight sections. These sections are as follows: (A) Perceptions of sentencing in general; (B) Offense seriousness; (C) "Fair" sentences; (D) "Going" sentences; (E) Fairness of specific guidelines sentences; (F) Sentence equivalents; (G) Appropriateness of guidelines factors; and (H) Sanction severity. The majority of the respondents had little trouble answering the questions in each section, bar Section F which was excluded from analysis as prisoners were incapable of the abstraction and conceptualization needed to perform the task. In Section F, prisoners were given a sentence term, and then asked what that term "would be worth" if served in various other institutions. For example, one of the questions asked what a sentence of two years of probation would be equal to if served in (a) the county jail, (b) Yardville Correctional Center, or (c) prison. The questions were intended to give some idea of how place of incarceration and length of time incarcerated interact to affect the overall perceived severity of the sanction; reflection, however, leads us to believe that it would be rare to find anyone who could perform the task required in this section. Needless to say, Section F was not duplicated in the 1980 questionnaire; rather, another method was employed (successfully) to get at the issue of sanction severity.

Section A contained a number of questions designed to tap prisoners' views of the sentencing process in general, opinions on the appropriateness (or fairness) of their own sentences, and thoughts on what measures could be taken to improve the sentencing process. Nearly all questions in this section were open-ended and designed to elicit rapport with the respondent. (This section proved particularly useful for allowing the respondent to vent his anger and tell his war stories so that the following sections could move directly into specifics.)

Section B, the offense seriousness section, required prisoners to assign a score from 1 to 15 to 26 different offense descriptions similar to the ones used by a number of offense seriousness scaling researchers (cf. Sellin and Wolfgang, 1964; Figlio, 1975). Respondents were told to give

a score of 14 or 15 to offenses that they thought were very serious; and to give a low score, such as 1 or 2, to offenses that they thought were not very serious. Similar instructions were given in Section H, where prisoners were asked to assign scores to indicate their perceptions of the severity of 20 different penalties ranging from a fine of $50 to a life sentence. Both of these sections provided us with information on the relationship between the seriousness of the offense and the severity of the sanction, as defined by prisoners. [10]

Section C asked prisoners to assign "fair" penalties of their choice to the same 26 offense descriptions that appeared in Section B: in Section D, prisoners assigned sentences that represented the sentence they thought New Jersey judges typically gave for such offenses.

The questions in Sections E and G dealt specifically with issues surrounding the New Jersey sentencing guidelines. In Section E, respondents were given short descriptions of offenses and offenders, and then were asked whether the several sentences suggested by the New Jersey guidelines for those offenses and offenders were too heavy, fair, or too light for each case described. Section G asked respondents their opinions of the factors that the guidelines have designated appropriate for either exacerbation or mitigation of sentence disposition. These two sections provided us with the data that allow us to estimate what prisoners would think about the guidelines, if only they knew of them.

The 1980 Questionnaire

The 1980 questionnaire contained seven sections, and there were four different versions administered. The seven sections are as follows: (A) Perceptions of sentencing in general; (B) Offense seriousness; (C) "Fair" sentences; (D) "Going" sentences; (E) Appropriateness of guidelines factors; (F) Preference of sentence terms, or sentence place; and (G) Sanction severity. The four different versions varied interviewer instructions for Sections B and G. Section A of the 1980 questionnaire was a near duplicate of Section A from the 1979 questionnaire; and Sections C and D also resembled the 1979 version, although the number of items was reduced to 24 from 26.

Section F was a new section, and was designed to replace the old Section F of the 1979 questionnaire. In this section, respondents were asked to choose between two sentences, on the basis of which of the two sentences they would prefer to serve, and they were given 26 of these paired comparisons. [11] The 1980 version of Section F, contrary to its predecessor, elicited a show of understanding from the respondents; and,

also, gave us some indication of the weight that both length of sentence and place of sentence (i.e., jail, prison, probation, etc.) give to the overall perceived severity of a sanction. Section E -- which asked about guidelines factors -- is similar to Section G of the 1979 questionnaire; and although the actual questions differ somewhat, the guidelines factors are constant in both questionnaires.

The basic idea behind the offense seriousness section and the sanction severity section remained essentially the same in the two questionnaires; however, the 1980 version modified interviewer instructions to facilitate understanding on the part of the respondent. During the interviewing in 1979, a number of prisoners claimed that they found it somewhat difficult to assign numbers to offenses and sanctions (although data analysis showed that this difficulty was not as great as expected). In order to aid respondents in their number selection, the 1980 questionnaire utilized visual aids in Sections B and G; and two different interviewer instructions were employed in each of these sections.

All respondents were presented with scale cards, which had the digits 1 to 15 placed in equal increments across a straight line. It was hoped that if respondents had such a card in front of them, then they would be more likely to remember that 14 is two numbers bigger than 12, and that the 15 digits increase in equal increments -- this phenomenon being something that not many adults comprehend.[12] As well, we attempted to capitalize on the findings from the 1979 research by anchoring that scale, or setting moduli at the first and third quartiles, by telling respondents that their peers of the year before said that an injury during a fight was about a 3 or 4 on the scale, and an attempted armed robbery fell at about 11 or 12 on the scale -- similar instructions were given in the sanction severity section where the scale was anchored with two sanctions that were scored at the first and third quartiles by the 1979 sample. Since we were interested in seeing whether the views of the two samples were consistent between the two years, we experimented with the anchoring instructions for half of the respondents and used instructions similar to the 1979 questionnaire for the other half of the respondents. The two versions of the interviewer instructions for Section B appear below; Section G used similar instructions.

Version 1

> Now I'm going to read out some descriptions of crimes, and I'd like you to tell me how serious you think each one is by giving it a score from 1 to 15. If you think the offense I describe is very serious, give it a

high score like 14 or 15; if you think it is not very serious, give it a low score like 1 or 2; and if you think it is about average, give it a score in the middle like 7 or 8. OK? (GIVE R BLANK SCALE CARD AND MAKE SURE HE UNDERSTANDS.) For all the offenses, the offender is an adult male with no prior record and what he has done is very much like what most people do when they commit these offenses. OK? Now on this scale, how serious do you think it is if the offender injures a person in a fight? (WRITE IN VERBATIM HERE AND ON THE SCALE.) What score do you give this offense: the offender demands money of a person and assaults the victim with a weapon; the victim does not give him any money? (WRITE IN VERBATIM HERE AND ON THE SCALE.)

Version 2

Now I'm going to read out some descriptions of crimes; and I'd like you to tell me how serious you think each one is by giving it a score from 1 to 15. If you think the offense I describe is very serious, give it a high score like 14 or 15; if you think it is not very serious, give it a low score like 1 or 2; and if you think it is about average, give it a score in the middle like 7 or 8. OK? (SHOW R SCALE CARD WITH MARKINGS.) See, last year inmates told us that if an offender injures another person in a fight, then that offense got a score around 4 or 5 on this scale. And if the offender demands money of a person and assaults the victim with a weapon; and the victim does not give him any money, then that offense got a score around 11 or 12 on this scale. Now for each offense that I am going to read to you, I want you to assume that what he has done is very much like what most people do when they commit these offenses. OK?

About half of the respondents received Version 1 instructions, in which they were allowed to anchor their own scale with the modulus offenses; the second half were allowed no freedom in selecting seriousness scores for the modulus offenses (and few indicated disagreement with the placement of the offenses on the scale). Similarly, roughly half of the respondents to Section G were allowed to select their own scores for the modulus sanctions. The two versions of both sections were mixed to form the four different versions of the 1980 questionnaire, and respondents were randomly allocated to one of the four questionnaire versions.

Findings

The two questionnaires, both of which took on average one hour to administer, provided responses to a wide variety of sentencing issues. The results of the data analysis have been divided into three categories, reflecting the major areas of interest. The first division of this section will discuss results from the open-ended section of both questionnaires, and will report prisoners' perceptions of sentencing in New Jersey. The second division relies primarily on 1980 data; and it discusses prisoners' perceptions of offense seriousness, sanction severity, and the relationship, as they perceive it, between the seriousness of the offense and the severity of the sanction. Finally, the third division will report results from the specific guidelines sections of the two questionnaires.

Analysis of Sentencing Perceptions

Clearly, asking prisoners their opinions of the sentencing procedure in a jurisdiction is bound to produce responses indicating something less than complete pleasure with how things currently operate. The respondents at Rahway State Prison certainly felt quite strongly that justice was not the word they would use to describe criminal procedure in New Jersey. Of the 1979 respondents, 66 percent said that the sentences in the State of New Jersey were unfair, and another 23 percent thought that at least some of the sentences were unfair; thus bringing the percent dissatisfied to a total of 89 percent. Similarly, in 1980, 72 percent thought all sentences were unfair, and an additional 12 percent indicated that at least some sentences did not match up to their definition of fairness.

That the majority of the respondents in both years felt that sentencing was unfair should not be surprising. We asked prisoners whether they thought sentencing in New Jersey was generally fair or unfair; and forced choices of this sort are unlikely to produce answers that suggest unconditional acceptance. Prisoners do, however, have reasons for thinking that sentencing is unfair, and these reasons indicate a number of concerns. Table 5.1 presents the reasons that prisoners gave to explain why they thought sentencing was generally unfair. Data are presented for both 1979 and 1980; and the percentages are based on the overall number of responses given to the question (as opposed to the number of respondents answering). (It should be noted that in 1979, respondents were allowed to give three answers to this question, although only 7 respondents did so; in 1980, only the first two responses were coded.)

Insert Table 5.1 here

As this table indicates, the fact that sentences were perceived as being inconsistent accounted to a large degree for the dissatisfaction in 1979. Furthermore, the fact that sentences were perceived as not matching the seriousness of the crime was mentioned with a relatively high frequency in 1979. (The category of "other", for both years, contains a number of responses that could not possibly be taken as the answer to the question asked, or the response category was filled by an n of one.) In 1980, dissatisfaction shifted to a concern with the length of the sentence, and a concern with racial injustice. Interestingly, those factors that were mentioned with the greatest frequency in 1979 -- i.e., inconsistency and lack of fit between seriousness and severity -- are two major concerns that sentencing guidelines purport to address. And concern with inconsistency dropped by almost 21 percent by 1980; furthermore, no one in 1980 said that sentences do not match the crimes.

Whether this finding is conclusive evidence that the guidelines did in fact achieve their aim in just one short year is, of course, doubtful. For one thing, if guidelines were truly in operation, and operating as intended (and if prisoners were cognizant of this), then it is unlikely that a concern with variation by judge or attorney would increase after guidelines implementation, since guidelines are also intended to reduce disparity in judicial behavior. The more likely explanation for this shift in concern from 1979 to 1980 is not that guidelines affected perceptions, but that something else affected perceptions. A new sentencing code (Ch. 2C. of the New Jersey Statutes Annotated) was implemented in September, 1979; and during the interviewing of 1980, numerous prisoners expressed concern that this new sentencing code would increase sentence terms dramatically. It is not known yet what effect the new code has actually had on sentence terms[13]; but inmates at Rahway thought they knew exactly what effect the new code would have -- thus it should be expected that they would be more concerned with sentence length in 1980. As well, it may be the case that neither guidelines nor the new sentencing code had any impact on prisoners' perceptions; but that something independent of a change in sentencing practices affected perceptions; or indeed, the percentage differences between the two years may be misleading.[14] This table does, however, capture those areas that prisoners think contribute to the overall unfairness of sentencing in New Jersey.

Table 5.1: Percentage of Responses Indicating Dissatisfication with Sentencing in New Jersey.

Response	Questionnaire Version	
	1979	1980
Inconsistent	25.5% (39)	4.9% (10)
Forces Pleas	3.2% (5)	3.4% (7)
Too Long	9.2% (14)	22.3% (46)
Prejudicial	11.1% (17)	11.7% (24)
Based on Prior Record	2.0% (3)	7.3% (15)
Varies by Judge or Attorney	2.6% (4)	8.7% (18)
Don't Match Seriousness of Crime	13.1% (20)	-0- (0)
Don't Take in all Circumstances	8.5% (13)	5.8% (12)
R's own Sentence Unfair	----------	6.3% (13)
Other	24.8% (38)	29.6% (61)
TOTALS	100.0% (153)	100.0% (206)

The 1980 questionnaire added a new coding category to the question on what makes sentencing unfair: this category -- respondent said his own sentence was unfair -- was intended to reduce the size of the "other" category; clearly the intended result did not occur. Prisoners, however, do tend to think in somewhat selfish terms, and 13 respondents in 1980 said that because their own sentence was unfair, the whole system was unfair. Presumably, though, all respondents are basing their perceptions of sentencing in New Jersey on their own experience with sentencing, so the suggestion that "if mine's unfair then they're all unfair" may not be altogether unreasonable. At least those that believe that they were treated unjustly have some reason for thinking that injustice may be a component of the system.

As expected, when prisoners were asked directly whether they thought the sentence they received was a fair sentence, 77 percent in 1979 and 72 percent in 1980 said that they had received an unfair sentence. A number of men were willing to suggest that the sentence they received was unfair because they were innocent of any wrongdoing (which might lead one to question the veracity of at least some of the respondents); but the majority of the men stated more specific reasons for why they believed they did not receive a fair sentence, indicating at least a marginal acceptance of, or belief in, their guilt.

Table 5.2 presents the reasons that the 1979 and 1980 respondents used to justify why they believed that they had received unfair sentences. Again, the percentages in this table are based on the total number of responses to the question, rather than the total number of respondents. (In 1979, one response per respondent was coded, so the percentages in this table also represent percentages based on the number of respondents. In 1980 two responses were coded.)

Insert Table 5.2 here

For the 1979 respondents, 39 percent felt that their own sentence was unfair because it was too long, followed by 17 percent who thought that they did not deserve any sentence whatsoever. The 1980 respondents indicated that they too thought their sentences were too long, and a number of respondents felt that not all the circumstances surrounding their cases were considered fully. Respondents for both samples also said that their sentence was unfair either because it did not match the sentence friends received for similar offenses, or because it did not match the sentence that their codefendants received. This concern with what

Table 5.2: Percentage of Responses Indicating Dissatisfaction with Own Sentence.

Response	Questionnaire Version	
	1979	1980
I Was Innocent	17.4% (20)	12.8% (22)
Too Long	39.0% (45)	22.7% (39)
Codef./Friends Got Different Terms	12.2% (14)	10.5% (18)
Not all Circumstances Considered	9.6% (11)	12.8% (22)
Forced into Plea	5.2% (6)	5.8% (10)
Prejudicial	0.9% (1)	-0- (0)
Based on Prior Record	1.7% (2)	7.5% (13)
Other	14.0% (16)	27.9% (48)
TOTALS	100.0% (115)	100.0% (172)

others received suggests that, to a large extent, prisoners define the fairness of their own sentences in relation to what they know others have received for similar offenses. This, of course, should not be surprising; it is difficult to imagine how inmates could realistically evaluate the fairness of their own sentence without conducting some rough comparative analysis.

Interestingly, the majority of respondents said that the other prisoners serving time at Rahway believed that they had received unfair sentences (73 percent); but only 49 percent of the respondents agreed that their fellow inmates received unfair dispositions. So while it may be true that, in part, respondents define the fairness of their own sentence in terms of what other inmates received, they seem reluctant to believe that their fellow inmates also received unfair dispositions.

In summary, it seems that respondents either felt that they should have received no time at all, or they felt that they should have at least received a lighter sentence than they did receive. In fact the 1980 respondents said, on average (55 percent), that the sentence they received would not be fair for others convicted of the same charge because (a) the sentence was too long (25 percent), or (b) the circumstances surrounding all crimes vary so much (33 percent) -- and they were willing to suggest these reasons despite the fact that they did not think that all of those sentenced received unfair dispositions. [15]

Furthermore, respondents felt that sentencing in New Jersey in general was unfair, and that the majority of this unfairness was exhibited (especially for 1980 respondents) in lengthy sentences. And when the 1980 respondents were asked specifically about the length of sentences handed out by the New Jersey judiciary, 57 percent said that the sentences were generally too heavy, while only 21 percent said that the sentences were generally fair. Not one respondent said that the sentences in New Jersey were generally too light. (23 percent of the respondents were unwilling to commit themselves to any of the forced choices for this question; they, therefore, comprise the "other" category.)

Analysis of Seriousness and Severity Perceptions

One way of finding out whether prisoners think that sentencing is fair or unfair is by asking them directly. In many ways, the questions discussed in the preceding section did just that. It seems, however, that given the way in which these questions were formatted, it would be unusual to find anything but negative responses to the questions posed, considering the subject matter, and the respondents. Another

way of assessing perceptions of the fairness or unfairness of sentencing is to take a round-about route, and ask prisoners what they think are fair sentences for specific crimes, and then compare these suggested fair sentences with their idea of what sentences New Jersey judges typically give for such offenses. Another measure of perceptions of unfairness, therefore, will be reflected in the amount of disagreement between perceived fair sentences and perceived going sentences for specific offenses. Any comparison of "fair verses going" sentences for specific offenses, however, requires some knowledge of how prisoners perceive the seriousness of the offenses and the severity of the sanctions, for one of the things that makes sentencing fair is consistency between the seriousness of the offense and the severity of the sanction. [16]

Respondents were first asked to give a number from 1 to 15, to each of the offense descriptions in Section B of the questionnaire. These offense descriptions appeared on cards and the interviewer was instructed to shuffle the cards, in order to insure that the respondents received the offense descriptions in a random order, thereby reducing the chance that any of the scores would be a function of an order effect. [17] These same offense descriptions were then given to the respondents, in random order, two more times: one time respondents were asked to indicate what they thought would be a fair sentence; the other time they were asked to indicate what sentence they thought New Jersey judges typically give for such offenses.

Respondents were also asked to give a number from 1 to 15 to each of the sanctions in Section G of the 1980 questionnaire, and Section H of the 1979 questionnaire. In 1979, sanctions were not presented in a different order to each respondent, and analysis shows that the failure to insure that these sanctions were presented in a random order evidently produced an order effect. [18] As well, the 1979 data on sanction severity did not contain as many stimuli as the 1980 data, nor were the 1979 data coded as finely as the 1980 data. For these reasons, this portion of the data analysis will utilize 1980 data only; those readers interested in 1979 data on the seriousness of the offense and the severity of the sanction are referred to Shelly and Sparks (1980).

Perceived Seriousness of Offenses

For these data, the median was used to represent the central tendency of the assigned scores, as it is a more resistent measure of central tendency than the arithmetic mean. The 1979 data for offense seriousness scores show that

for each of the offense descriptions a response of 1 and a response of 15 was obtained. In 1980, responses did not always fall at both ends of the scale for each offense description; however, there still remained wide variation in responses. If the mean were to be used with this data, the more serious offenses would (because of low deviant responses) appear to be less serious than was really the case; similarly, the less serious offenses would move up the scale and claim higher mean scores than they really deserved. The median, by contrast, is not nearly as susceptible to influence by extreme outliers (cf. Mosteller and Tukey, 1977:207-208; Lax, 1975).

Table 5.3 presents the median scores, and SAM values for each of the 24 offense descriptions in Section B of the 1980 questionnaire, in rank order of perceived seriousness. SAM, or the Standardized median Absolute deviations from the Median, is a measure of spread around the median.[19] This statistic is not widely used in data analysis; but it is a measure more suited to the data here than other measures of spread. SAM is very much like the coefficient of variation; however, unlike the coefficient of variation, it does not rely on the mean for its computation. The SAM statistic is derived by computing the absolute deviations of score values around the median -- which is the median absolute deviation, or MAD estimate[20] -- and then dividing the MAD by the median. MAD measures the spread of values around the median, and the division by the median itself allows for comparisons to be made between stimuli having different medians. (The division by the median, in essense, cancels out the effect that the magnitude of the median has on the measure of spread -- it, therefore, standardizes the statistic.)

Insert Table 5.3 here

Comparisons between Version 1 data and Version 2 data indicated that different respondent instructions did not, as anticipated, produce different values for the medians and SAM's; therefore, the two sets of data have been aggregated for analysis. As this table displays, prisoners too think that incest, rape, and intentional homicide are very serious crimes. In fact, this rank order suggests that prisoners' perceptions of the seriousness of crimes vary little from what would be expected if the sample was comprised of non-deviant persons. The one offense that may, perhaps, seem to fall lower than would be expected if the respondents were of the general population is that of assault of a police officer (median = 5.27).[21] Prisoners, however, do seem to feel quite strongly about this offense; many prisoners indicated that they believed that an assault of a police officer was

Table 5.3: Medians and SAM Values for Offense Descriptions

Offense	Median	SAM
Incest	14.84	.01
Rape	14.79	.01
Intentional Homicide	14.73	.02
Arson	13.91	.08
Homicide	13.50	.11
Attempted Homicide	11.80	.15
Armed Robbery (elderly victim)	10.47	.24
Assault with Intent to Commit Robbery	10.31	.26
Aggravated Assault and Battery	9.94	.21
Sale of CDS	8.37	.43
Sale of CDS to a minor (marijuana)	8.22	.46
Armed Robbery	7.82	.28
Burglary (commercial)	5.33	.43
Assault	5.27	.62
Burglary (residential)	4.66	.49
Possession of a Dangerous Weapon	4.32	.53
Embezzlement	3.91	.49
Receiving Stolen Property (credit card)	3.46	.43
Assault	3.36	.71
Possession of CDS	3.00	.67
Forgery	2.96	.65
Auto Theft	2.82	.43
Larceny from the Person	2.77	.64
Receiving Stolen Property (stereo)	2.26	.57

never unprovoked, was never anything but a reaction to the violence of the police officer himself, and was in many instances to be considered the duty of every citizen. Given this rather peculiar definition of the offense, it perhaps should be considered surprising that this offense received such a high score.

The degree of consensus (as measured by SAM) clearly varied across offenses; inspection of the table shows that consensus is quite closely related, negatively and in a linear fashion, to median scale scores. Thus, the less serious (on average) the offense was scored, the greater the disagreement on its seriousness value.

There are, however, notable exceptions to this linear increase in consensus as the seriousness of the offense increases. The largest amount of dissensus is shown in the offense of simple assault (median = 3.36; SAM = .71). It should be remembered, though, that all the respondents to Version 1 received this offense as the first stimulus; therefore, the high value of SAM may indicate that not only did respondents disagree as to the seriousness of the offense but also, respondents had yet to understand clearly the section instructions. The second stimulus received by respondents (attempted robbery) displays a much smaller value of SAM (.26), which might be taken as an indication that after respondents were presented with their second stimulus, they were able to make the comparisons between offenses that are necessary for making relative seriousness judgements.

A word of caution is advised in interpreting both the median scale values and the SAM values for the very serious offenses. One of the major criticisms of the scaling technique used here is that of "end-point truncation effect". Figlio (1979) has argued quite convincingly that restricting responses with the use of a categorical scale produces a scale that has a disproportionate number of stimuli falling at both the upper and lower ends of the scale. Other researchers, therefore, have allowed respondents to utilize the entire integer scale (known to them, of course) when selecting scores to represent their perceptions of seriousness.[22] Table 5.3 shows that, by restricting respondents to the first fifteen digits of the natural number system, a truncation effect did occur at the high end of the scale; therefore, it is erroneous to assume that incest is perceived as being more serious than rape, and rape is perceived as being more serious than intentional homicide -- rather, all three of these offenses are perceived as being <u>very</u> serious, and the variation in median score values (due to the interpolation of the SPSS program) represent little more than variation in scores given

by one or two respondents. Also, given the objective seriousness of these offenses, it should not be surprising that the variation shown in SAM is so low. It may well be the case that if respondents were allowed to use the full range of integer values, then the clumping at the upper end of the scale would disappear, and the values of SAM for the very serious offenses would increase.

Perceived Severity of Sanctions

Table 5.4 presents the medians and SAM values for the 24 sanctions from Section G of the 1980 questionnaire; the 1979 sanction data will not be analyzed here.[23] Both versions of data have been aggregated for analysis as, similar to the offense seriousness data, variant instructions did not affect medians and SAM's for the two versions.

Insert Table 5.4 here

As this table displays, prisoners felt that capital punishment was the most severe punishment, while a fine of $50 was the least severe punishment; this should not be surprising. Also, within each of the categories of type of disposition (i.e., prison, jail, probation, and fine), the larger the time or dollar dimension of the punishment, the greater the perceived severity. This positive relationship between time and severity does not, however, take a straight linear form; rather, severity increases logarithmically the with time or dollar amount of the sanction.

This rank order of sanctions also indicates that prisoners do not perceive prison as always being more severe than jail, jail as being more severe than probation, and probation as being more severe than a fine. This suggests that neither time nor place of punishment (or degree of loss of freedom or money in the case of probation and fines) is the sole determinant of perceived sanction severity; rather it seems that prisoners weigh type of punishment and time of punishment together in order to come up with an overall severity score. Severity of sanction is thus not a "unidimensional" concept -- but then neither is seriousness of crime (see e.g., Gottfredson, 1979).

The SAM values on this table indicate that not all prisoners were in agreement with the overall median severity scores. Consensus as to severity is quite high at both ends of the scale; a relatively large amount of dissensus occurs at those points on the scale where type of penalty switches from prison to probation, from probation to jail, and from

Table 5.4: Medians and SAM Values for Sanctions

Sanction	Median	Sam
Capital Punishment	14.97	.002
Natural Life	14.95	.01
Life with Parole	14.69	.02
25-30 Years Prison	14.44	.04
20-25 Years Prison	13.85	.09
15-20 Years Prison	12.34	.14
12-15 Years Prison	11.71	.15
10-12 Years Prison	9.81	.12
7-10 Years Prison	7.95	.14
5-7 Years Prison	6.26	.21
3-5 Years Prison	5.14	.23
1-3 Years Prison	3.50	.43
5 Years Probation	2.92	.38
18 Months Jail	2.86	.38
12 Months Jail	2.43	.58
6 Months Jail	1.90	.47
$500 Fine	1.84	.44
2 Years Probation	1.65	.38
1 Year Probation	1.28	.23
6 Months Probation	1.23	.17
$100 Fine	1.18	.17
$50 Fine	1.09	.09

probation to a fine. Thus, while there is little disagreement as to the severity of those very serious penalties and the least serious penalties, the middle-of-the-range sanctions do not display much consensus across respondents. So while it may be true that the majority of prisoners perceive a sanction of 6 months in jail as being more severe than a $500 fine, there are also a large number of prisoners who believe that the fine is more severe than the jail term. We suspect that, if such data on the 1980 sample were available, those prisoners who are quite used to doing time and believe that they could do a 6 month jail term "standing on their head" would be more likely to perceive the fine as being a more severe sanction for them; while those individuals who cannot do time in such a leisurely fashion would find that the jail term was the more severe sanction.[24] Therefore, the high values of SAM for the middle of the scale sanctions probably represent differences in sanction preferences among the sampled population.

The Relationship Between the Two Scales

If it is assumed that the median values for the offense descriptions and the sanctions represent perceived seriousness and severity, then it is possible to use the two scales, in conjunction with the data from Sections C and D of the 1980 questionnaire, to determine quantitatively the differences between perceptions of what would be the fair sentencing practice and perceptions of present sentencing practice. If there is a difference between the sentence that prisoners chose as fair for an offense description and the sentence that they said New Jersey judges typically handed out for those offenses, then this will indicate that prisoners do not think that current sentencing practice is fair; the amount of difference between stated fair and perceived going sentences will also indicate how far from fair current sentencing practice is, according to prisoners.

As a measure of the perceived disparity between what prisoners think should be the fair sentences for specific offense descriptions and what they think current practice is, the median fair sentence response can be subtracted from the median going sentence, thus giving some measure of the amount of perceived disparity. Since severity of sanctions does not increase in a linear manner with the time dimension though, subtracting number of years or months for the two variables would give a misleading picture of perceived differences in severity between "what is" and "what should be"; therefore, in order to treat the fair and going sentence responses as interval data that increase in severity in a specified way, the data must be recoded so that they reflect their perceived severity. Thus, all responses to Sections C and D were

recoded with the median perceived severity values of the response, and the median fair and going sentences were then recalculated for each of the offense descriptions.

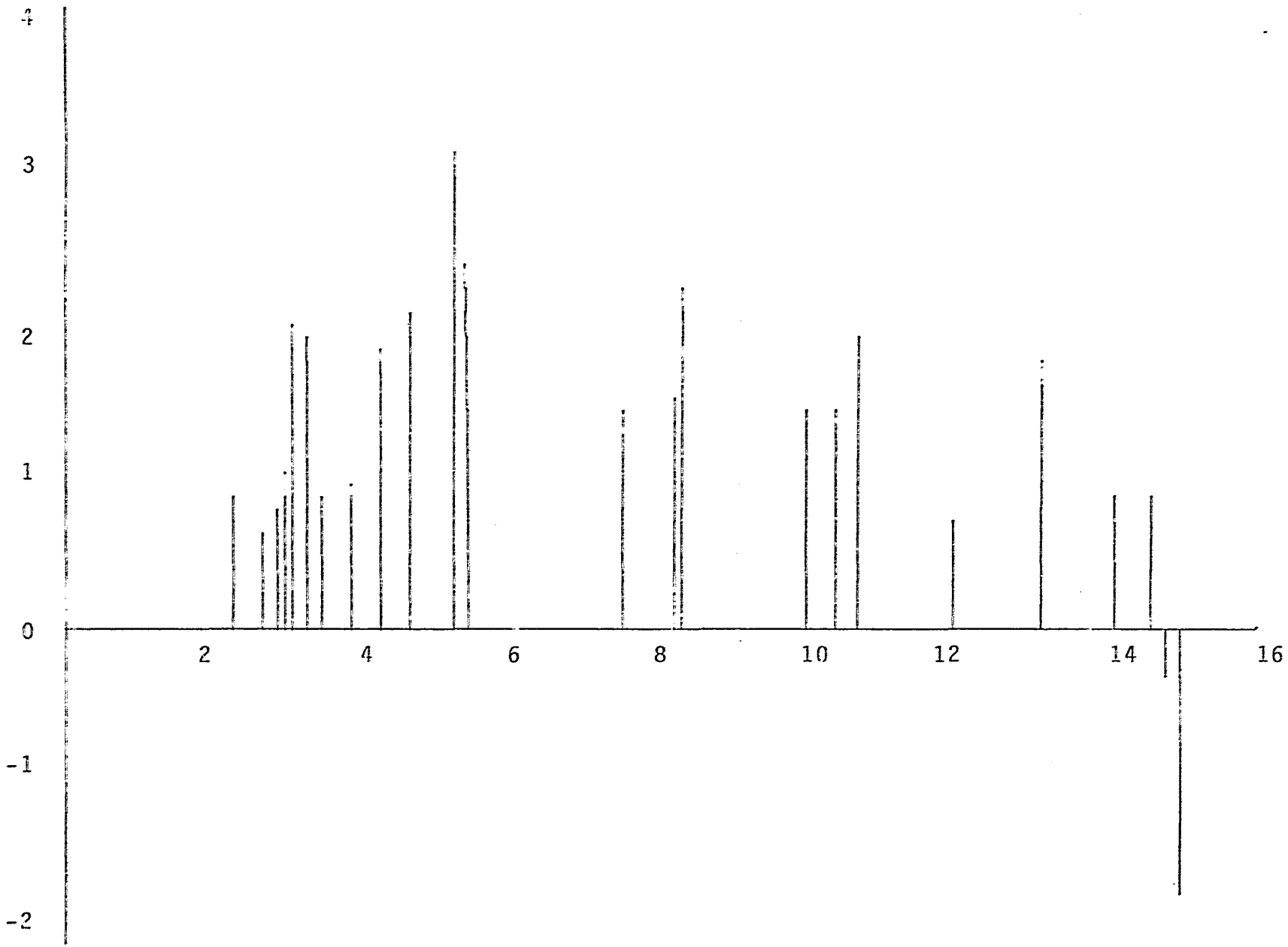
Figure 5.1 presents a plot of the perceived disparity between median fair sentence and median going sentence for each of the offense descriptions. The horizontal axis on this plot is the offense seriousness scale; values above and below the horizontal axis represent the magnitude of perceived difference between perceived fair sentences and perceived going sentences. If it were the case that prisoners perceived no difference between current sentencing practice and their definition of fair sentencing, then there would be no values above or below the line, as these values represent the difference when the median perceived fair sentence is subtracted from the median going sentence; therefore, those values above the line indicate a positive value (or that the going sentence is heavier than the fair) and those values below the line indicate a negative value (or that the going sentence is lighter than the fair).

Insert Figure 5.1 here

These data substantiate the claims of prisoners that sentencing is not fair in New Jersey; and that current sentencing is generally too heavy for the seriousness of the offense. For all but two of the offense descriptions, prisoners felt that the going sentence was more severe than the fair sentence; prisoners did feel, however, that the fair sentence for the two most serious offenses (rape and incest) should be more severe than current practice. If these two extreme values are disregarded, however, it seems that unfairness, and degree of unfairness, are unrelated to the perceived seriousness of the offense. Thus, disparity between "what is" and "what should be" does not increase or decrease with the seriousness of the offense; but, rather, perceived disparity is randomly distributed across offenses. (One offense shows a high degree of disparity and distorts the distribution: this is the offense of assault on a police officer -- clearly, if any offense is perceived to get too much time, this is that offense.)

Asking prisoners directly about sentencing practice produces responses that show that current practice is perceived as being far from fair; interestingly, asking prisoners the same question in a round-about way produces the same response. At the very least, then, prisoners are committed to the belief that sentences in New Jersey are unfair and too long, and they can substantiate this belief by

Figure 5.1: Disparity Between Perceived Fair and Going Sentences by Seriousness of the Offense



responding to a series of questions that do not directly relate to perceptions of fairness in sentencing. Of course, this does not mean that sentencing really is unfair in New Jersey; but there is strong evidence to support the hypothesis that prisoners think that sentencing is unfair.

Analysis of Guidelines Perceptions

The New Jersey sentencing guidelines were developed, and implemented, for the purpose of enhancing the fairness of the sentencing process by reducing unwarranted disparities in dispositions. The guidelines intent is to structure judicial sentencing behavior uniformly statewide for specific offenses, while simultaneously allowing variation within each offense group for aggravating or mitigating factors. We would guess that the Administrative Office of the Courts (which we noted earlier was responsible for the construction of the guidelines) was not concerned with the perceptions of prisoners when they proposed structuring sentencing; but prisoners are the consumers of a sentencing policy, and one indicator of the success of a new sentencing policy can be found in the satisfaction of its consumers. [25]

Since the prisoners at Rahway State Prison were not aware that a new sentencing policy was in effect [26], they could not, of course, comment intelligently on the impact of sentencing guidelines. They did, however, agree with the concept of sentencing guidelines, and believed that such a policy would help to make sentences uniform for like offenders, while still allowing latitude in dispositions for individual and case variants. [27] But even though Rahway prisoners were in support of guidelines in principle, they were dissatisfied with the suggested sentences for the specific offenses, and with the factors designated appropriate for aggravation or mitigation of a disposition.

In order to find out prisoners' perceptions of the suggested guidelines sentences, we gave them a series of vignettes, which described an offense and an offender. A list of penalties followed each vignette, and prisoners were asked to indicate whether they thought each of the penalties was too heavy, fair, or too light for the case described. For example, the first description read as follows:

> The offender has been convicted of robbery. He has one prior adult conviction and he was on probation at the time of arrest. He stole more than $200. The offender has a job, and will still have it after sentencing.

For this offense and offender, the New Jersey sentencing guidelines allow any of the following dispositions: eight and a half years in prison, seven years indeterminate at Yardville, 12 months in the county jail, 18 months in the county penitentiary, or no sentence to incarceration. [28] Any of these dispositions would be "fair" if imposed in the foregoing case; but as these appropriate dispositions vary in severity to such an extent, we expected that prisoners would find some of these dispositions "fairer" than others.

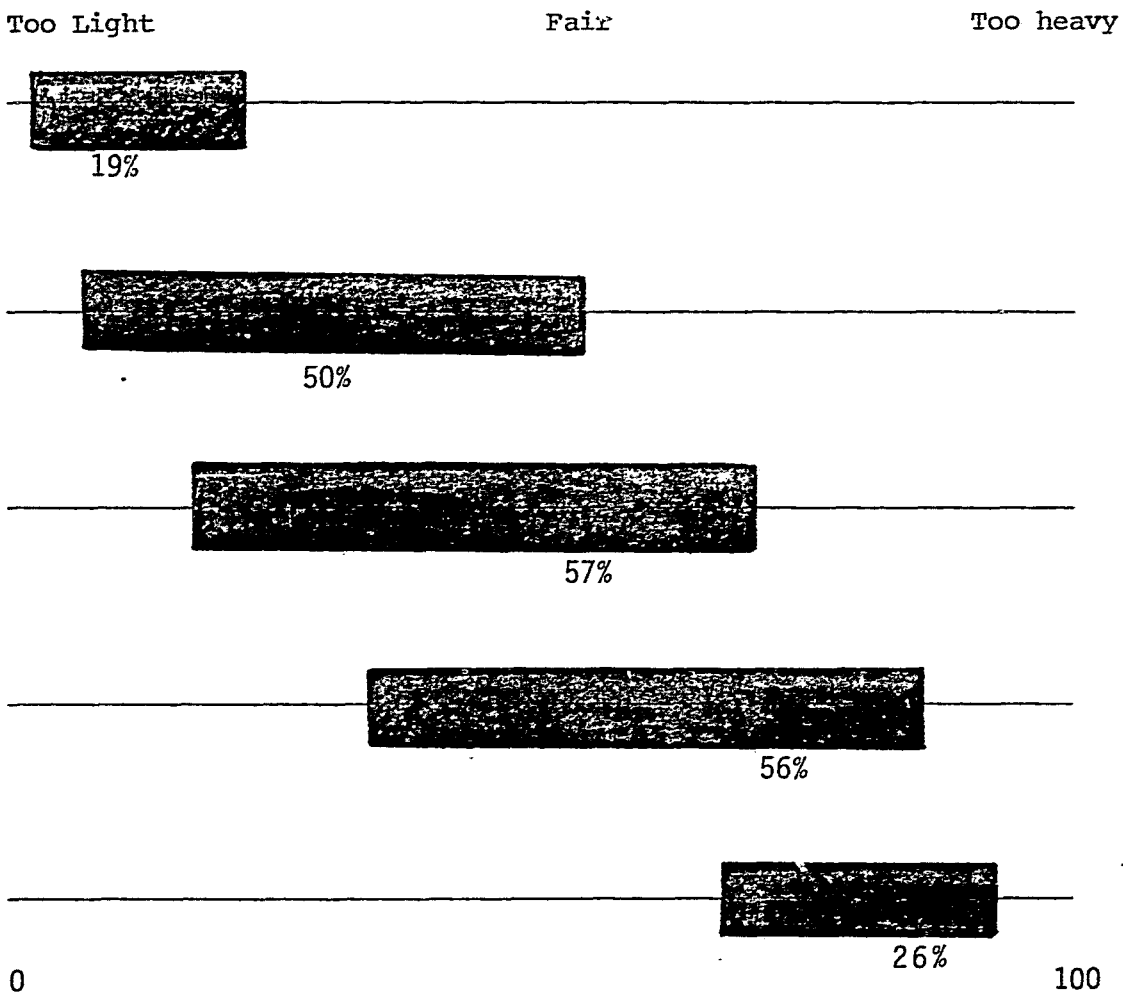
Figure 5.2 presents prisoners' perceptions of the five different guidelines sentences that are designated appropriate for the robbery case. The width of the shaded boxes on this plot indicate the proportion of respondents who thought that the respective penalty was fair; the length of the lines to the left of each box display the proportion of respondents who thought that the suggested guidelines sentence was too light; and the lines to the right of the boxes display the proportion of respondents who thought that the sentence was too heavy for the case described.[29]

Insert Figure 5.2 here

Not one of the respondents felt that no sentence to incarceration was too heavy for the robber in the vignette; nor did many think that eight and a half years in prison was too light of a sanction. Therefore, the appropriate sentence must be perceived to fall somewhere in between these two extremes. The majority of the respondents felt that either a term of 18 months in the county penitentiary or 12 months in jail would be fair; and half of the respondents thought an indeterminate term would be appropriate. For those respondents who disagreed with the majority as to the appropriateness of these three sanctions, most who disagreed with the indeterminate sentence thought that the sentence was too heavy, and most who disagreed with the 12 month jail sentence thought that it was too light. The term of 18 months in the county penitentiary split those who disagreed pretty much evenly between the too heavy and too light extremes.

Generally, not one of the suggested penalties for the robber was considered fair by a large number of prisoners. A little over half of the respondents said that two of the sentences would be fair, but a substantial percentage disagreed, and they disagreed as to whether the sentence was too heavy or too light. Similar patterns of agreement and disagreement as to the fairness of a particular penalty were found in responses to five other vignettes. In fact, the highest percentage of agreement as to a <u>fair</u> sentence, for all

Figure 5.2: Prisoners' Perceptions of Guidelines Sentences for Robbery Case*



*(The total N across sanction stimuli ranged form 140 to 143)

six cases, was for the 18 month penitentiary sentence for robbery; the sentence that respondents thought was most unfair was no sentence at all for rape (only one respondent felt this would be fair). It would have been more encouraging if at least one of the penalties suggested for a case showed a high degree of consensus as to its fairness, for then it could be argued that that penalty was perceived to be the fair penalty for the case. It appears however that, just as the New Jersey judiciary perceives the need to allow widely different sentences for a specific offense and offender, some prisoners also perceive that five different sanctions can all be fair for one offender.

The New Jersey sentencing guidelines not only suggest appropriate dispositions for specific offense types, but they also suggest terms for a specific offender within each of the offense types. Evidently, the Administrative Office of the Courts felt that certain offender characteristics were in the past, and should be in the future, significant for the dispositional decision; and, therefore, those influential characteristics of the offender were built into the actual guidelines. Some of these offender characteristics reflect favorably on the offender (such as being employed), and are used to mitigate the disposition; other characteristics (such as use of a weapon) are used to exacerbate the disposition. The guidelines recognize that any one defendant may have some mitigating and some exacerbating characteristics, so they allow for specified dispositions for "unique" offenders within each of the offense groups. [30]

As the perceived fairness of the New Jersey guidelines sentences cannot be assessed without understanding their relation to the specific offender characteristics in the guidelines, we asked both samples of prisoners their perceptions of the various mitigating and exacerbating factors deemed appropriate for consideration in disposition. We also asked prisoners if other factors that are not built into the guidelines (such as race and sex) should be considered by a judge before making a dispositional decision. By and large, both samples felt that offender characteristics should and should not be considered -- most offenders were very quick to respond with "same crime, same time" when asked whether different offenders should receive different dispositions if they have committed the same crime; but they simultaneously felt that the judge should take everything known about the offender into account before sentencing, including such things as vague as "community background". Certainly, these two viewpoints, expressed by one person, seem to be polar opposites. Prisoners in both our samples, however, did not appear to see the inconsistency between these two statements; and it is our guess that the glib "same crime, same time"

statement was in many cases used simply because it is standard prison jargon, and was not truly believed by all who used it. In fact, when prisoners were asked directly about the appropriateness of specific guidelines factors (rather than asking them if offender characteristics in general should be considered) the number of respondents claiming to believe that solely the crime should determine the severity of the penalty decreased markedly.

We used two different methods to try to get at prisoners' perceptions of the designated guidelines factors. In 1979 we asked prisoners which of two offenders should get more or less time, on the basis of a certain characteristic, given that they had committed the same crime. (For a few of the factors, we simply asked the 1979 respondents if the factor should be considered in sentencing.) For example, one question asked: "Do you think a married man who supports a family should get less time for the same offense than a single man, or a man who doesn't support his family?". In 1980, by contrast, we formatted the question this way: "What if one [offender] is married and the other [offender] is single. Should they get the same sentence?". Those respondents who said that the two offenders should get different sentences were then asked which of the two offenders should receive the harsher disposition.

Although in both samples we asked prisoners relatively similar questions about the same guidelines factors, the subtle difference in wording produced different responses between the two years. It appears, however, that the different responses between the two samples do not represent a difference in aggregate perceptions; but, rather, are a function of the question format.[31] In 1979 we gave respondents only one choice: should one offender receive less time than another. Forcing the response in this way may have precluded responses such as "same crime, same time", and may have limited responses suggesting that the offender should not receive less time than his counterpart, but more time. Of course, 1980 respondents were forced into agreeing or disagreeing with the "same crime, same time" proposal; but they seemed to be quite willing to do so. Nevertheless, because of the forced nature of both sets of questions, responses indicating agreement with one offender receiving less time in 1979, and both offenders receiving the same sentence in 1980 may be more prevalent than should be the case. With this caution in mind, we turn first to prisoners' perceptions of the mitigating and exacerbating guidelines factors, followed by prisoners' perceptions of the factors not contained in the sentencing guidelines.

Of all the guidelines factors we asked prisoners about, marital status produced the largest percentage of

disagreement, amongst the 1980 respondents, with the New Jersey guidelines. Although the guidelines do not specifically state that a married man should receive less time than a single man, they do allow mitigation of disposition if the offender is supporting a family. The majority of the 1980 respondents (76 percent) felt that the married man and the single man should receive the same disposition. Being married, however, does not necessarily imply support of a family; and this may explain in part why the 1979 respondents, when asked whether a married man who supports his family should get less time, were more likely (45 percent) to think that the married man should receive a lighter disposition.[32] Had the 1980 question included the element of support, it is suspected that the proportion of respondents advocating similar dispositions for the two offenders would drop. But since it is impossible to separate out the effect of including the element of support on the 1979 responses, it is difficult to tell whether the 1979 respondents agreed with the Administrative Office of the Courts that a man who supports a family deserves less time, or whether the 1979 respondents felt that marital status alone was a significant aspect to be considered by judges. It appears, however, that a number of prisoners believed that a judge should consider marital status as it relates to the support of a family, as of the 21 percent of the 1980 respondents who said that the married man should get less time, nearly all said that marital status was important because a married man has responsibilities to support a family.

The guidelines also allow for mitigation of a sentence if the offender displays other characteristics that suggest a strong community background.[33] Along with support of a family, the guidelines have allowed for less severe dispositions for the offender who has a job, or is attending school. Prisoners did not seem to be suitably impressed with the sentiment that an offender in school should receive a less severe disposition than an offender not in school -- 72 percent of the 1980 respondents said that these two offenders should get the same sentence. Since very few of the prisoners in Rahway were attending school at the time they committed the offense for which they were currently incarcerated, perhaps education does not mean much to the prisoners, nor may it be perceived to alter an offender's character. Employment, as well, does not appear to be perceived as a factor worthy of consideration for disposition. In 1979, 45 percent of the respondents said that the employed offender and the unemployed offender should get the same sentence; 28 percent said that the employed offender should receive a less severe disposition; and 15 percent said that the unemployed offender should receive a less severe disposition. The 1980 respondents also tended to think that both offenders should

receive the same sentence (60 percent); however, a much larger proportion of the 1980 respondents who felt that the two offenders should receive different dispositions, believed that the offender who is employed deserves a more severe disposition than the unemployed offender[34]. This represents a substantial amount of disagreement with the guidelines, both in terms of whether employment should be considered at all, and in terms of what effect it should have on sentences. Thus, the majority of the respondents in both surveys felt that education and employment status should not be considered; if employment status is taken into account, then it should be used in exactly the opposite way that the guidelines suggest. Rather than considering employment as a factor that should decrease a sentence because it represents strong community ties, prisoners seem to believe that employment is a factor that increases culpability; it is the offender without the job that has the need to commit crime, whereas the employed offender does not have such need.

Prior record is one variable that some of our prisoners seemed to agree was important for the dispositional decision. The 1979 respondents were fairly well split on their opinion of prior record, with 42 percent saying it should be considered, and 46 percent saying it should not be considered; the majority of the 1980 respondents, however, felt that the offender with the prior record should receive more time.[35] Although the percentage of respondents in agreement with the sentencing guidelines is not large in this case, we at least found that when prisoners said that a factor should affect the disposition, they believed that the effect should be as the guidelines intend.

When we asked prisoners whether they thought a plea of guilty reflected favorably on an offender, and whether they thought it should be considered in the dispositional decision, we again received an even split of opinions. 56 percent of the 1980 respondents felt that the offender who goes to trial and the offender who pleads guilty should receive the same sentence; and 43 percent felt that the two offenders deserved different dispositions. And of the 43 percent who did not think that the two offenders should receive the same sentence, only 58 percent said that the offender who goes to trial deserves the longer disposition. Thus, roughly half of the respondents thought the two offenders should receive the same sentence; and approximately one-quarter of all respondents agreed with the policy established in the sentencing guidelines.

If an offender does not use drugs or alcohol, or has joined a drug program since his arrest, the guidelines allow for mitigation of the disposition. Of the 1980 respondents,

63 percent felt that the offender who is addicted to drugs deserves a different disposition than the non-addicted offender; but rather than believing that the user should receive the longer disposition, 70 percent of those who said that the addict and non-addict should receive different dispositions felt that the non-addict should receive the lighter disposition. Prisoners justified this opinion on the grounds that the non-user is fully aware of his actions, and the addict is forced to commit crime because of his habit. As with employment status, then, prisoners tend to take the underdog position; and appear to believe that if an offender does not have a job, or has a drug habit to support, the only option available to insure survival is the commission of crime.

The race, sex, and age of the offender are not factors that the New Jersey sentencing guidelines consider appropriate for exacerbation or mitigation of a disposition in the guidelines. We did ask prisoners, though, whether they thought these things should be considered by a judge. Our respondents agreed that race was not appropriate for consideration, were ambivalent about an offender's age, and appeared (reluctantly) to agree that an offender's sex should not affect the disposition. 95 percent of the 1979 respondents said that an offender's race should not affect a disposition; and because of the high amount of agreement on this question, and because of the outrage produced by the suggestion that a black and white offender should receive different dispositions, we did not bother to ask our 1980 respondents the same question. As for the age of the offender, 60 percent in 1979 said that the older offender should get the lengthier disposition because knowledge of right and wrong increases with age; but 56 percent of the 1980 respondents felt the age should not matter in the dispositional decision. Since we do not know what our respondents had in mind when they were commenting on variation in dispositions for old and young offenders (were they thinking of an 18 year old and a 60 year old, or were they thinking of a 25 verses a 30 year old?), it is difficult to say what is the perceived importance of an offender's age for these two samples. Sex, however, is another matter. In 1979, 53 percent of the respondents said that the male and female offender should receive the same disposition; and in 1980, 72 percent said that these two offenders should receive the same disposition. Even though the majority of these responses indicate support of the guidelines, interviewers seemed to feel that this "same crime, same time" response for this question was far from prisoners' true sentiment. Whether our interviewers are right, it is interesting (and not at all surprising) that those respondents who felt that the female should receive a lighter disposition than her male

counterpart, believed that this should be for typically "macho" reasons such as: "women are weaker than men" or "I just like women".[36]

In summary, it would be misleading, if not plain wrong, to suggest that our two samples of prisoners were in agreement with the decisions made by the Administrative Office of the Courts. The various sentences suggested by the guidelines for offense and offender specific cases showed highly mixed perceptions as to their fairness; and the guidelines factors as well were not highly regarded as appropriate for consideration by a substantial majority of the samples. Race and a non-incarcerative sentence for a rapist are two exceptions; but it is still discouraging to note that only two aspects of the sentencing guidelines were perceived by many to be important for the dispositional decision. Nonetheless, this lack of support and praise for the sentencing guidelines does not necessarily mean that prisoners perceive that guidelines will not enhance the fairness of sentencing; it just seems that the New Jersey guidelines are not going to be perceived to do enough to enhance the fairness of sentencing in New Jersey.

Discussion

The original intent of the two surveys of prisoners at Rahway State Prison was to assess their perceptions of the New Jersey sentencing guidelines. Whether it is because prisoners at Rahway were uninformed of the implementation and use of the sentencing guidelines, or whether guidelines were in fact not in use, it is clear that the perceptions of prisoners of sentencing guidelines, as discussed in this chapter, are based not on their actual knowledge of the guidelines, but on our perceptions of what prisoners' perceptions of the guidelines are likely to be. Neither of our samples was in a position where they could discuss the impact of guidelines intelligently with us, so only by presenting them with hypothetical cases, and by asking them their perceptions of sentencing before guidelines, were we able to glean some notion of what prisoners are likely to think about the sentencing guidelines.

If we assume that prisoners' perceptions of the fairness of sentencing and of the virtues of sentencing guidelines are important, then the picture is grim. Prisoners do not think that sentencing was fair before guidelines, and they will not in all likelihood think that sentences are fair under sentencing guidelines. Of course, it must be remembered that consensus among prisoners was not great; therefore, it might be the case that guidelines will be perceived to enhance the

fairness of sentencing for at least some chunk of the prisoner population. But exactly how, and to what extent, guidelines will be perceived to enhance the fairness of sentencing is another question; it is our guess that, short of decarcerating all of the prisoners at Rahway State Prison, promising them acquittal if they transgress in the future, and ensuring them that all those "others" who deserve to serve time will do so, prisoners as a whole will continue to perceive of sentencing in New Jersey as unfair.

Notes to Chapter 5

[1] Chapter 4 to this report explains the method used to reconstruct pre-guidelines sentencing data from the New Jersey sentencing guidelines.

[2] It would be possible to predict what dispositions would look like after guidelines implementation; but in order to predict with some degree of accuracy, it is necessary to know if the guidelines are systematically applied in dispositional decisions -- and this is not known.

[3] Rahway prisoners generally agreed that sentencing was unfair, but they were especially critical of the "unfairness" of their own sentences. In fact, 76 percent of the 1979 sample felt that at least some prisoners deserved the sentence that they received; but of course very few men believed that they deserved the sentence that they themselves received.

[4] This is not to suggest that subjective perceptions will invariably change in the case of an objective change.

[5] The two other state institutions are located in Trenton and Leesburg. Trenton State Prison houses only maximum security inmates; Rahway and Leesburg contain minimum, median, and maximum security inmates.

[6] During both the 1979 and the 1980 surveys, minimum security inmates were housed at either Rahway Camp, Rahway Trailer Park (both of which are adjacent to the main institution), or Marlboro Camp. Rahway Trailer Park was subsequently eliminated because of a high escape rate.

[7] Between group comparisons were made across the following background variables: race, age, resident county, county where sentenced, offense of present commitment, sentence length, prior record variables, amount of time served on present sentence, and various aggravating and mitigating characteristics (e.g. use of weapon during offense, employment status, etc.).

[8] We anticipated that the guards responsible for giving passes to prisoners, so that they could be interviewed, would exercise their discretion, and that they would select out for interviewing only those persons that they thought we might like to see. Our respondent group, however, contained roughly the same proportion of rapists, murderers, and non-English speaking persons as the non-respondent group. We, therefore, thank the administration at Rahway State Prison, and the correctional officers.

[9] Even though the 1979 questionnaire was pretested on a group of graduate students, modifications and simplification in the 1980 questionnaire were necessary so that prisoners could understand fully the questions.

[10] It is not necessarily true that prisoners perceive a commensurate relationship between the seriousness of the offense and the severity of the sanction; but as it is somewhat more difficult to justify theoretically some alternative relationship, it is hypothesized that prisoners do think in terms of commensurability.

[11] See Thurstone (1927) for a discussion of the use of paired comparisons methods for comparative analysis.

[12] The Texas Study of Adult Competency (1976) provides data on the literacy rate of the general public.

[13] Data published by the New Jersey Department of Corrections suggest that the new penal code did not affect the severity of dispositions, at least within the first year after its implementation.

[14] As the bases from which the percentages were derived are not close in value (153 vs. 206), some of the difference in percentages between the two years is directly attributable to the size of the base, rather than representing a difference in perceptions.

[15] See note 3.

[16] von Hirsch (1976) and Singer (1979) would argue that commensurability between the seriousness of the offense and the severity of the sanction is not <u>one</u> of the things that makes sentencing fair, but is <u>the</u> thing that makes sentencing fair. We allow here for other considerations of fairness that prisoners might hold.

[17] Pepitone and DiNubile (1976), using an experimental design, found that the order in which stimuli were presented affected perceptions. For example, if a stimulus of homicide were to follow a stimulus of bicycle theft, the homicide stimulus would be "perceived" as being much more serious than if it had followed a stimulus of arson. Presenting stimuli in a random order, however, randomizes the order effect across a sample, resulting in a more accurate picture of perceived seriousness across a sample as a whole.

[18] For the 1979 sanction severity data, the term of six months probation was "perceived" as being less severe than two years probation but more severe than one year probation.

Since there is no sound reason why a term of six months probation could possibly be considered more severe than a term of one year probation, this "error" is most likely attributable to the order in which the stimuli were presented. When sanction stimuli were randomized in the 1980 survey, problems such as the above did not occur.

[19] We are indebted to Howard Wainer for instruction on this and other points.

[20] See Mosteller and Tukey (1977), and Tukey (1977) for further information on the use of MAD, and other innovative summary measures.

[21] In the National Survey of Crime Severity, Census Bureau interviewers were instructed to probe whenever respondents gave seemingly strange scores to offenses, in order to determine that they had not misunderstood the scoring procedure. It appears, however, that in most cases respondents had reasons for assigning unusual scores, but that these were unusual reasons. (Personal communication from Linda Murphy.)

[22] One thing that researchers who use magnitude estimation must assume, however, is that respondents understand the equal increase in increments property of large numbers; we were unwilling to make that assumption with our population. See Sellin and Wolfgang (1966) for the classic study of offense seriousness perceptions using magnitude estimation.

[23] For those readers interested in 1979 severity data, see Shelly and Sparks (1980).

[24] This hypothesis is based on conversations with prisoners; and it appears to make sense.

[25] The judiciary of New Jersey are also the consumers of sentencing guidelines, but as no one spoke with members of the judiciary, their level of satisfaction cannot be documented. (It should be noted that refusal to communicate came from the judiciary; staff were more than willing to discuss the impact of sentencing guidelines.)

[26] Only a small number of prisoners claimed to know of the guidelines; and it is believed that those who said they were familiar with sentencing guidelines were really familiar only with the new penal code.

[27] 78 percent of the 1979 respondents said that there was a need for sentencing guidelines. Although we cannot be

sure that the respondents and the researchers had the same definition of sentencing guidelines in mind, 38 percent did say that guidelines would help to make sentences uniform.

[28] A "no sentence to incarceration" term does not mean that the defendant received no punishment whatsoever. The severity data discussed earlier show that probation or fines may be even more severe than an incarcerative sentence; so it is erroneous to assume that "no sentence to incarceration" is the least severe penalty offered in the guidelines.

[29] This exercise was presented only in the 1979 survey.

[30] Of course, all aspects of a given offender could not possibly be contained within the sentencing guidelines (they are cumbersome enough already); therefore, only unique offenders within the guidelines factors that were designated appropriate for consideration are enumerated in the actual guidelines.

[31] A comparison of the size of the "other" categories between the two surveys indicates that respondents had much more difficulty with the questions in the 1979 survey. The percentage of responses that fell into the "other" category in 1979 ranged from 3 percent to 18 percent; in 1980; however, the percentage of "other" responses ranged only from 1 percent to 3 percent.

[32] This represents the majority, as only 4 percent of the 1979 respondents said that the single offender should get less time; 36 percent said both offenders should receive the same sentence; and 15 percent responded with such bizarre answers that they comprise the "other" category.

[33] Under the heading of "community background", the guidelines include the following factors: drug or alcohol use, employment status, whether the offender is in school, and whether the offender contributes to the support of others. See Chapter 4 for a complete description of this section of the New Jersey guidelines.

[34] 72 percent of the sample felt that the offender who is unemployed deserves a less severe disposition than the employed offender; but this percentage is based on those respondents who said that the offender should receive different dispositions.

[35] Of the 62 percent of the 1980 respondents who said that the first offender and the repeater should receive different dispositions, 91 percent said that the offender with the prior record should receive the lengthier disposition.

[36] It is possible that some of the responses to the question on the appropriateness of considering the sex of the offender were induced by the sex of the interviewer. No differences across responses were noticed when sex of interviewer was controlled for; however, the effect of the interviewer's sex on the responses could conceivably vary by respondent characteristics.

Chapter 6: Sentencing in Massachusetts: The Climate for Sentencing Guidelines

The Massachusetts sentencing guidelines purported to be primarily[1] the product of a statistical analysis of past sentence dispositions by all of the Superior Court judges throughout that state. As we have seen, the claim that sentencing guidelines describe past sentencing practice, rather than prescribe it, has been characteristic of most of the guidelines development projects -- especially up until the time when the Massachusetts guidelines project began to construct statewide sentencing guidelines for the Superior Court.[2] Whether the Massachusetts guidelines actually live up to that claim -- i.e., of accurately describing past sentencing practice through the use of a statistical model -- will be considered by us in later sections of this report. This chapter will first describe the context in which sentencing in Massachusetts took place, prior to the development and use of sentencing guidelines. It will also try to shed some light on the social and political climate that led to the Superior Court judiciary's decision to develop statewide sentencing guidelines.

As a prerequisite to an evaluation of sentencing guidelines -- especially those that purport to be based, even in part, on past sentencing practice -- some notion of what past sentencing practice actually was, is necessary;[3] and data on past sentences -- like the data used to construct guidelines -- will not by themselves adequately describe that practice or its organizational context. To that end -- of understanding the nature of the sentencing process in Massachusetts before sentencing guidelines were created -- the legal structure of sentencing in Massachusetts in 1978[4], and the informal sentencing policies of judges, prosecutors, defense counsel and others in the state's criminal justice system at that time, will be described.

Finally, the perceptions of the sentencing guidelines concept by criminal justice system personnel -- judges, prosecutors and defense counsel, among others -- as well as the projected operation of guidelines and their expected impact on sentencing in Massachusetts, will be discussed in some detail. Of course, this last section can only describe the system's personnel's notions of what statewide guidelines would do to the then usual sentencing policies, in a hypothetical sense; for at the time of our interviews with the various system actors, guidelines had not yet been implemented. Even so, the concerns of criminal justice system personnel about the guidelines' structure and use provided additional direction to the evaluation of those guidelines, by

pointing out significant "system" areas of concern in addition to the methodological and statistical aspects of our evaluation.[5]

Origins of the Massachusetts Guidelines

We went to some lengths in Chapter 2 of this report to point out that the trend toward the control or "structuring" of judicial discretion in sentencing is far from recent, and that sentencing guidelines are merely one relatively new approach to sentencing reform. This section's concern is with the conception and development of a sentencing guidelines reform effort in a single state, namely Massachusetts. What influences make a local or state jurisdiction interested enough to consider sentencing reform in the first place? And what factors, in the final analysis, lead a reforming body to choose sentencing guidelines as the specific mechanism of reform? Does the type of reforming body (e.g. the legislature vs. the judiciary), or do the reasons behind the reform movement, affect the structure or content of the product -- in this case, sentencing guidelines? The origins and the development process of the Massachusetts sentencing guidelines provide a complex and illustrative case study suggesting answers to questions such as these.

The history of the Massachusetts sentencing system parallels, in part, the cyclical nature of movements to control judicial discretion seen nationally and abroad, as discussed in Chapter 2. In our description of the Massachusetts general laws dealing with the sentencing of criminal offenders in the following section, we note that the sentencing structure in Massachusetts is principally indeterminate. Judges are presented, usually, with a fairly wide range of sentencing alternatives that they may choose from in dealing with offenders. Over the years, however, revisions to the sentencing statutes have stipulated that sentences may not be suspended for specific types of offenses, and have established mandatory sentences for second offenses or for certain specific types of offenses, such as weapons violations.

Concern about sentencing reform in Massachusetts followed in the wake of a general movement, which appears to have started in the 1960's, to modernize or streamline the organization of and procedures followed in the Massachusetts criminal justice system. The revision of statutes regulating the setting of bail and the granting of the defendant's release on recognizance (ROR) appears to be the first concerted judicial, legislative, and bar organization effort toward some sort of procedural modernization (Spangenberg, 1967). The Committee on Bail of the Massachusetts Council on

Crime and Delinquency carried out a small research study (statistically examining the influence of about 15 variables) which concluded that some "lack of uniformity" existed in the use of bail which could not be explained by the "certain highly relevant factors" that they were able to examine (Spangenberg et al., 1967:143). One early product of the bail reform effort -- The Massachusetts Bail Reform Act, Chapter 681 of the Acts of 1966 -- provided judges with a set of factors[6] to consider in the determination of bail or ROR. The bail reform effort itself is not what concerns us in our case study of Massachusetts; what is important to us is that the writings about that effort are the first to note concern about disparity -- or "lack of uniformity" -- in a type of disposition, and to establish some sort of decision-making criteria even though loosely defined. Later on, these two issues would be cited again, by Massachusetts reformers, as reasons for developing statewide sentencing guidelines.

Just as the work of the Bail Committee was ending, a commission whose purpose it was to "help 'modernize the criminal law'" was established on the recommendation of the Governor's Committee on Law Enforcement and Administration of Criminal Justice. This commission -- the Massachusetts Criminal Law Revision Commission (MCLRC) -- began operating on 27 June 1968 (MCLRC, 1972:vii-viii). Included among the tasks with which the Commission was charged[7] was the development of "a rational grading and classification of crimes, with an attempt to develop criteria for sentencing." (p. viii, emphasis added). The MCLRC's work continued until 20 November 1971, when what became known as S.200 -- An Act Establishing the Massachusetts Criminal Code -- was filed in the 1972 Massachusetts legislative session. Even though S.200 was never enacted, the sentencing scheme that it would have provided -- with the consent of the MCLRC's 52 members (including legislators, judges, district attorneys, public officials, scholars, and citizens) -- deserves some discussion.

The definition of crimes and penalties provided by S.200 was similar in structure to that proposed by the Final Report of the National Commission on Reform of the Criminal Laws and contains the standard sentencing provisions embodied by the Model Penal Code (American Law Institute, 1962). The code would have classified offenses into seven categories: 4 classes of felony offenses (A to D), 2 of misdemeanors (A, B) and 1 class of "violations". The corresponding sanctions prescribed the maximum sentence allowable for each class of offense. As the offense class decreased (from felony A to D, or from misdemeanor A to B) so did the allowable maximum sentence.[8] For the two most serious classes of offenses (Class A and B felonies), the judge would be allowed to set a

minimum length of time that the offender would have to spend in confinement before parole.[9]

The code specified that the presumption in most cases would be not to incarcerate an offender[10] and provided general (and non-exhaustive) criteria to guide this decision[11]; in cases where incarceration was thought the appropriate disposition, the judge would be free to sentence an offender to any term of incarceration up to the maximum stipulated for each offense class. Criteria that would guide the judge in the choice of sentence length within the maximum range allowable were not provided in S.200. It was presumably thought that the maximum term alone would provide enough direction or limitation to that decision.

Clearly, the sentencing "criteria" of S.200 provide only guidance of a very general nature. It is doubtful, though, whether the Commission ever intended to do more than this. The Governor's Committee report, The Revision of the Massachusetts Criminal Code, expressed the concern that "Under our statutes, similar crimes often carry dissimilar sentences, and dissimilar crimes often carry identical sentences" (1968:10). When given the task of reforming the criminal law in answer to the concerns of that Governor's Committee Report, the Criminal Law Revision Commission proposed a structure in which the seriousness of crimes would be matched by the severity of maximum permissible penalties; thus, similar crimes would now carry similar sentences (assuming that judges would sentence similarly within the allowable range).

Not expressed, however, by either the Governor's Committee report, or by the MCLRC, is a concern with disparity -- which, in this instance, can be rephrased as offenders dissimilar (in certain relevant characteristics) receiving similar sentences for similar crimes. The Governor's Committee report did note that sentencing criteria "had not been articulated in Massachusetts, either by the legislature or by the courts" (1968:10), and the MCLRC's proposed code did provide general criteria for the incarceration decision; but the concern about lack of sentencing criteria does not appear to have been linked to "sentencing disparity" or the "abuse of judicial discretion." In fact, the opposite seems to be the case: the sentencing criteria included in S.200 were not intended to limit judicial discretion; rather, as Professor Livingston Hall (the Chairman of the Criminal Law Revision Commission) noted, "the sentencing provisions of the Code emphasize flexibility" (Hall, 1972:63); in a later report that castigated "new mandatory sentences which would unduly limit the discretion of the sentencing judge" (Hall, 1973:341, emphasis added).

It is clear from Livingston Hall's comment that the subject of mandatory sentences had begun to cause some controversy in the Massachusetts legislature even before the debate over S.200 had been concluded. S.200 was never to be enacted, as we noted earlier; but a mandatory sentence bill -- specifically dealing with offenders caught possessing and/or carrying a handgun without the appropriate registration licenses -- received considerable legislative backing. The Bartley-Fox bill, as the legislation came to be known, was passed by the legislature in July 1974 and took effect on 1 April 1975 (Massachusetts General Laws, C.269, s.10, as amended). Within a year, the Massachusetts Supreme Judicial Court had opportunity to review the constitutionality of the Bartley-Fox law in the case of Comm. v. Jackson (344 N.E.2d 166, 1976).[12] The defendant had challenged the constitutionality of the statute on a number of grounds; one contention that deserves discussion here was that the statute's mandatory minimum sentence requirement infringed on judicial discretion in violation of the separation of powers doctrine of the Massachusetts constitution.[13] In holding that the operation of the statute's sentence provisions did not "directly affect the capacity of the judicial department of function" (344 N.E.2d 177, 1976), the court concluded that the judiciary did not possess any discretion in light of the legislative mandate in the Bartley-Fox law.

> Although it is the court's function to impose sentences upon conviction, it is for the legislature to establish criminal sanctions and, as one of its options, it may prescribe a mandatory minimum term of imprisonment. (344 N.E.2d at 177, emphasis added).

The belief that it is the perogative of the legislature, and not the judiciary, to determine criminal sanctions -- as discussed in this decision and often reiterated since -- influenced dramatically the eventual structure and use of the Massachusetts sentencing guidelines. A commentary on the decision remarked that "The Court's opinion is a careful exercise of judicial restraint...the opinion notes the extent of serious debate over the efficacy of mandatory sentencing, but the court restrains itself from entering what it deems to be a purely legislative function..." (Mandatory Sentencing, 1976:158).[14]

The Bartley-Fox law was the first mandatory minimum sentence[15] to be passed by the Massachusetts legislature, and its passage did not mark the end of the sentencing reform debate.[16] Rather, the bill, in retrospect seems to have focussed the attention of the legislature, judiciary, and the public squarely on the topic of sentencing disparity. The publicity accorded to the enactment of the mandatory minimum

gun law -- with its heralded theme that "nobody can get you out" -- may in part be responsible for rising public outrage over the extent of violent crime.[17]

The Bartley-Fox gun law completely eliminated judicial discretion in the sentencing of gun law violators. A proposal that requires that offenders serve one year in jail for a gun law violation clearly announces that such a sentence was not the rule in the past; if a one year sentence had been the standard, there would not have been any great need to pass a new law. Rather, what was clearly implied was that a one year sentence was not the standard gun violation sanction -- due, at least in part, to the fact that judges, in exercising their discretion, did not impose such a sentence in all (if any) such cases. (Also present was the implication that, had judges previously applied a one year sentence for all gun law violations it would have made a difference to the violent crime rate. This is, of course, a question open for considerable debate, but one beyond the scope of this report.)

In any event, our interviews with judges throughout the Commonwealth (which we describe in more detail below) confirmed that judges were not enthusiastic about the Bartley-Fox gun legislation. Among the reasons cited for the displeasure was an indication that the removal of the judge's sentencing discretion led, at times, to a miscarriage of justice -- such as when a judge was forced to sentence someone who was either unaware of the law, or who was not otherwise a "real offender", to a one year jail term. Faced with rising public support for the idea of "locking up" potential criminals through the use of new mandatory minimum sentences for other offenses as well, it does not seem unreasonable that the judiciary might conclude that there was some need to protect their sentencing discretion. In his first report as Chief Justice of the Massachusetts Supreme Judicial Court in 1977, the Honorable Mr. Justice Edward F. Hennessey described the problem in very similar terms.

> Because of the pervasive public fear and anger induced by the wave of violent crime, there is a danger that disparity in sentencing -- a problem which has been with us for centuries -- may become more acute. There is a demand that judges must deter crime by adopting new sentencing policies. (Hennessey, 1977:14-15).

Mr. Justice Hennessey admitted in this address that some "sentencing disparity" did exist in the Commonwealth, but he urged that the judiciary explore techniques (such as attending sentencing councils and institutes, and continued use of appellate review of excessive state prison sentences) to combat the problem. And he noted:

> I refer to another disparity problem: the fixed sentence set by statute and its first cousin, the mandatory minimum sentence set by statute.
>
> ...While the fixed and mandatory minimum sentences seem at first blush to eliminate disparity, it is also argued that, to the contrary, they create disparity in an insidious form: that they treat the crime, and not the criminal, with all his individuality and need for individual treatment. (emphasis in original)
>
> ...It is of basic importance to deal with the criminal as well as the crime. To do so, discretion must continue to reside in the sentencing judge. Justice will not be consistently served by inflexible statutory standards. (Hennessey, 1977:17-18) (emphasis added)

One of Chief Justice Hennessey's suggestions to reduce sentencing disparity echos the theme of the Governor's Committee report, The Revision of the Massachusetts Criminal Code, of ten years earlier. Hennessey's major proposition that "statutory standards or criteria for sentencing should be established" also included the stipulation that

> ...As to the more serious cases, detailed written findings by the judge should be required, subsidiary to the judge's conclusion as to the sentence to be imposed. Those findings should not be pro forma but must be specifically addressed to the facts of the background of the individual defendant, and show full consideration for the statutory standards. (Hennessey, 1977:18)

The judiciary was not the only group to express concern about hasty revision of the sentencing process. The availability of federal financing of sentencing research projects during this time[18] probably played some part in shaping the responses to the legislative mandatory sentence proposals which were being introduced for additional types of offenses. On 29 April 1976, the Crime and Justice Foundation in Boston had applied to the Law Enforcement Assistance Administration's regional office for funding to "review and improve sentencing in Massachusetts".[19] That proposal called for a review of current approaches to sentencing reform -- including mandatory sentences. The authors of that proposal noted:

> ...Sentencing is already an area of concern to the General Court of Massachusetts. In the 1976 session seventy-two bills were filed requiring mandatory sentences be imposed upon offenders convicted of certain offenses. (Crime and Justice Foundation, 1976:10)

The research proposal submitted by the Crime and Justice Foundation in 1976 was not funded; though the reasons for the denial of funding are not known to us, clearly it was not because of any lack of interest in the subject of sentencing reform, especially in mandatory sentences. During the following legislative sessions, proposed sentencing reforms included a code providing mandatory sentences for drug offenders,[20] legislation that would increase prosecution under the Bartley-Fox gun law,[21] legislation that would mandate minimum sentences for sexual assault,[22] auto theft,[23] and prostitution,[24] and, most recently, a code that calls for complete statutory revision of the criminal laws and the establishment of presumptive sentences.[25] Although these proposals met with varying degrees of support and success,[26] that the Massachusetts legislature continued to wrestle with sentencing reform through the 1981 term certainly implied that interest in the topic had not abated. The sequence of this legislative action must be described in detail here.

In September 1977, the Massachusetts legislature established a "joint special committee to make an investigation and study of the revision of the criminal law statutes for the purpose of establishing a criminal code with uniform sentencing procedure."[27] The committee was chaired initially by Senator Arthur J. Lewis, Jr. and Representative James Segel and, more recently, by Senator Lewis and Representative W. Paul White. The initial reports of the committee (First Interim Report (1978) and Second Interim Report (1979)) proposed presumptive penalties for drug offenses. Both reports of the committee met with a good deal of public attention. In a commentary on the Joint Committee's First Interim Report, dated 22 March 1979, the Massachusetts Bar Association labelled the presumptive drug law revisions "piecemeal recodification" and noted that the Joint Committee had not addressed a number of very important issues, such as: what would be the presumptive sentence ranges, who would set them, and what effects would the presumptive sentences have on the court system and on corrections. Though the presumptive scheme proposed by the committee was never enacted, the legislature did finally, in 1980, revise the sentencing structure for drug offenses that involved the manufacture or sale of dangerous drugs. The legislation that was ultimately adopted called for mandatory minimum sentences for both first and second offenders with the length of the minimum term of imprisonment variable depending upon the class of the substance.[28]

We should note, at this juncture, that in September, 1977 -- at about the same time that the Special Joint Committee was appointed and prior to the enactment of the mandatory minimum

drug legislation -- the Massachusetts Superior Court judiciary requested and received funding from the Massachusetts Committee on Criminal Justice to conduct a thorough study of sentencing practice throughout the state with the ultimate aim of developing statewide sentencing guidelines. By the time that the legislature finally approved mandatory minimum sentence legislation for certain drug offenses, the state's sentencing guidelines project had been in operation for over two years and, in fact, was completing development of a working sentencing guidelines model. (The activities of the project were not a secret, although they were not well publicized. In fact, in the same commentary that rejected the Joint Committee's First Interim Report, the Massachusetts Bar Association recommended that the judiciary be allowed the opportunity to develop its own sentencing guidelines.) This fact is relevant to our current discussion because terminology common to discussions about sentencing guidelines is found in a 1979 opinion of the Massachusetts Supreme Judicial Court on the constitutionality of the then pending mandatory minimum drug offense legislation.[29]

In May 1979, the Massachusetts House of Representatives requested that the Supreme Judicial Court review the constitutionality of several pieces of legislation on the topic of drug offenses pending in the legislature.[30] In particular, the Court was asked if some of the language used in the legislation could be construed as "unconstitutionally vague" and, more importantly, if the sanctions proposed by the various bills were "cruel and unusual punishment". In upholding[31] the constitutionality of the rather substantial mandatory minimum sanctions, the Court used a three-pronged test.

> ...Three factors compose the gist of the analysis: (1) the nature of the offense and of the offender, (2) a comparison with penalties of other jurisdictions for the same offense, and (3) a comparison with penalties for more serious crimes within the jurisdiction. (Opinion of the Justices to House of Representatives, 393 N.E.2d 319, 1979)

The language of the court, although used in the context of upholding the constitutionality of a very restrictive type of sentence, yet employs phrases and concepts often used to describe sentencing guidelines. Guidelines, we noted earlier in Chapter 3, are often said to take account of both the offender and the offense, for example. Also, the principal, and most visibly apparent, feature of most guidelines matrices to date has been the comparative severity of penalties for crimes of comparative seriousness.

In both this opinion and in at least one other opinion of a more recent date,[32] the Supreme Judicial Court's language pertinent to sentencing in general appears to have a dual aim: to bow to the legislative role in defining crimes and penalties, yet also not to close the door on the judicial exercise of sentencing discretion. In other words, the opinion urges restraint in the legislation of new criminal sanctions. In light of then pending legislation, the caution appears appropriate. During the 1980 session,[33] mandatory minimum sentences were also prescribed by the legislature for the offenses of: indecent assault and battery on a child under age 14; rape; assault with intent to rape; auto theft; auto theft to defraud; false statements alleging motor vehicle theft; maintaining a house of prostitution; and deriving support from an inmate of a house of prostitution. It is evident from glancing at the above list that, at that time, the legislature was proceeding, piecemeal, to revise the criminal code -- and that all of the statutory revisions encroached upon judicial sentencing discretion.

The 1981 Massacusetts legislative agenda continued the debate about sentencing reform. The Third Interim Report (1981) of the Special Joint Committee on Uniform Sentencing and Revision of the Criminal Law Statues submitted legislation on January 12, 1981[34] that proposed a presumptive sentencing structure for the sentencing of all criminal offenders to be added as a separate chapter to the General Laws.[35] Of course, the Special Joint Committee had proposed a presumptive penalty scheme (applicable only to drug offenders) in its First and Second Interim Reports (which were not adopted for enactment by the legislature); however, the presumptive structure proposed by the Third Interim Report differed from those earlier proposals in a number of respects: the ranges of the presumptive sentences for each class of offense were not as wide, and the sentence lengths proposed were, in general, much shorter and more realistic. Most importantly, for our purposes, the general structure and intent of the proposed presumptive penalty structure closely resembles the traditional structure of sentencing guidelines. (It should be noted that, by January 1981, the legislature had gained at least some familiarity with the concept of sentencing guidelines, for the Superior Court judges had been testing the second sentencing guidelines model developed by their project since May 1980.) For example, after forcefully rejecting the continued use of mandatory sentencing,[36] the Committee states:

> ...the proposed presumptive sentencing system will insure more certain and equitable sentencing than is present practice and, at the same time, will allow for the thoughtful exercise of judicial discretion by:...

> (2) establishing a narrow imprisonment range for each class of offenses from within which the judge will impose a specific imprisonment term;...
> (3) authorizing the judge to consider the aggravating and mitigating factors of each case;...
> (4) allowing the judge to sentence outside the presumptive imprisonment range if detailed written findings are provided. (House, No. 6304:15-16; emphasis added)

Of course, it may be happenstance that the language we emphasize above from the proposed legislation mirrors that of a working guidelines model, but in light of the Committee's prior reports such a conclusion seems unlikely. In any event, it is still uncertain whether the Massachusetts legislature will adopt this legislation (as it is equally uncertain what would happen to the Superior Court sentencing guidelines if it did adopt it); as this final report was in preparation, legislation drafted by the Governor's Office was yet to be introduced in the fall of 1981 that would again propose mandatory sentences for all offenses.[37]

What should be most clear from the above discussion is that, over a period of a considerable number of years, dissatisfaction with the system of sentencing offenders was quite prevalent in Massachusetts -- and that dissatisfaction was widespread among the various political forces (judges, legislators, and the prosecution and defense bar) throughout the state. To better understand the reasons behind everyone's apparent unhappiness with the existing sanctioning system, it is important to understand the complexity of the laws governing that system and its operation.

The Legal Structure of Sentencing in Massachusetts

The principal characteristic of the Massachusetts sentencing structure, despite the many abortive legislative attempts toward reform, is indeterminacy. The Massachusetts General Laws typically provide that a criminal offense may be punished by a sentence of any term up to a prescribed maximum number of years.[38] Although the stated statutory maximum constrains judicial discretion as to the length of sentence to some extent, the available range of sentence options usually is still quite broad. The sanction for the commonly high-volume offense of robbery, for example, is thus stated in the Annotated Laws of Massachusetts:

> Whoever, being armed with a dangerous weapon, assaults another and robs, steals or takes from his person money or other property which may be the subject of larceny shall be punished by imprisonment in the state

prison for life or for any term of years....(ALM GL C.265, s.17)

In this instance, a judge may sentence a robbery offender to any number of years of imprisonment up to the maximum allowed sentence of life and, additionally, may exercise his discretion by suspending any portion or all of the stated sentence. The judge also has the option of sentencing the offender to a term of probation supervision either instead of, or after, a term of incarceration.

While not all of the statutory definitions attached to criminal offenses in the Massachusetts General Laws have quite so broad a range of sentence options as our cited example, considerable judicial discretion as to the choice of sanction does continue to exist. Generally, the range of sentence options available to the sentencing judge for any particular criminal offense is bounded by four distinct considerations: (1) the jurisdiction of the court of conviction, (2) the statutory classification of the offense as either felony or misdemeanor, (3) the statutory maximum sentence that may be imposed for that particular offense, and (4) statutory mandated minimum sentences for certain types of crimes or for certain types of offenders. These four limitations on the type and length of sentence that may be imposed for any particular case often overlap in practice, as we will explain further in the sections that follow; however, the constraint applied by force of these rules is in practice often mimimal, considering the range of sentence alternatives available.

Court Jurisdiction and Offense Classification

The Superior and District Courts have jurisdiction over adult criminal cases in Massachusetts. The Superior Court generally handles all serious felony offenses, as well as appeals from the District Courts. The District Court, by comparison, disposes of less serious offenses including violations of public ordinances, misdemeanors, and less serious felony offenses (punishable by up to 5 years), in addition to several other felony offenses that share concurrent jurisdiction with the Superior Court.[39] It is clear from even this brief synopsis[40] of each court's jurisdiction that considerable overlap exists -- a number of offenses may be disposed of in either District or Superior Court.

With the exception of those offenses that may be handled by either court, the distinguishing characteristic of a Superior Court case is that it involves at least one serious felony offense.[41] What offenses are felonies in Massachusetts? The classification of offenses as either

felony or misdemeanor in Massachusetts is determined by the maximum sanction that may be imposed upon conviction. A felony is thus any offense which may be punished with a state prison sentence.[42] A misdemeanor offense may only be punished with a jail (in Massachusetts, House of Correction) sentence of up to 2 1/2 years, or a lesser penalty, such as a fine.

Unless a criminal case is initiated by indictment in the grand jury of the Superior Court, cases are first heard in the District Court, regardless of offense classification. Felony offenses punishable by more than 5 years imprisonment in the state prison and over which the District Court does not have concurrent jurisdiction, are automatically "bound over" to the grand jury for prosecution in, and ultimate disposition by, the Superior Court after an initial hearing in the District Court to establish "probable guilt". Should the offense be one over which the court does have concurrent jurisdiction, the District Court may choose to dispose of the case in the first instance or, again, may decline jurisdiction and "bind over" the case for handling by the Superior Court.[43] The Massachusetts Supreme Court has repeatedly said that a number of factors should be considered by the District Court judge before jurisdiction is declined.[44] The major items are (1) the circumstances of the particular case, (2) the penalty permissible for the offense, and (3) the need for grand jury review of the case and its potential ramifications.[45]

As we noted, felony offenses punishable by up to 5 years imprisonment in the state prison are within the jurisdiction of the District Court. Additional felony offenses that have concurrent jurisdiction in both the Superior and District Courts usually may be punished by sentences even longer than a 5 year state prison sentence. Regardless of the maximum allowable statutory penalty, however, the District Court cannot impose a state prison sentence for any offense upon conviction.[46] This provision restricts the sentencing discretion of the District Court judge; technically, he may only impose a House of Correction sentence on a convicted offender, and House of Correction sentences have a maximum time limitation of 2 1/2 years.[47]

The Superior Court is not restricted in its choice of incarcerative institutions when sentencing an offender; thus, a Superior Court judge may impose any sentence authorized by statute for an offense.[48] Clearly, therefore, the decision as to whether a case should be bound over to the Superior Court can have important ramifications on the length of an offender's sentence, since the length of the sentence could be much more severe in that court. This point is, in a sense, tangential to the main focus of our study -- that of

evaluating the adequacy of the Superior Court sentencing guidelines that were developed to reduce sentence disparity, rather than comparing Superior and District Court sentencing policies. However, should similar offenders, convicted of similar crimes, receive very different sentences in the Superior Court from what they would have received in the District Court but for the bind-over decision, the ability of the sentencing guidelines to reduce disparity would clearly be adversely affected.[49]

A hypothetical example may make this point clearer. Let us say that offenders convicted of breaking and entering in the daytime (an offense of concurrent jurisdiction) usually receive a lighter sentence if the case is disposed of in the District Court rather than in Superior Court. The Superior Court sentencing guidelines would incorporate only the sentence information of the Superior Court cases and would thus have a higher average sentence for breaking and entering in the daytime than the true average (which would be based on sentences for this offense from both courts).

If this hypothetical example is correct, sentencing guidelines might affect the bind-over decision and, thus, sentencing practice, in a number of unintended ways. For one thing, if the sentencing guidelines provide for the higher average Superior Court sentence, the guidelines would be institutionalizing sentence disparity, not within the Superior Court, but between the Superior and District Courts.[50] A second possible effect of sentencing guidelines that call for a higher sentence would be that they could change the way in which bind-over decisions are made. Though the District Court judge is ultimately responsible for the bind-over decision, it would be naive to ignore the influence of other court personnel, especially the prosecutorial staff, in that decision. Prosecutors may request that the District Court bind over a case to the Superior Court citing, of course, the special nature of the case or the unusual history of the defendant. Alternately, a prosecutor may, on his own initiative, bring the case to the grand jury for indictment.[51] Speer (1979:6) notes that if an indictment in the Superior Court is obtained, "the entire case is removed to the Superior Court and the District Court complaint will be dismissed".[52] If the prosecution has reason to believe, because of the sentence stipulated in sentencing guidelines, that a more severe sanction will be imposed if the case is disposed of in the Superior Court, there would be an incentive for the state to urge that the case be bound over or to proceed with a direct grand jury indictment. While both of these possibilities are only that, it should be recognized that the entire dual court structure and case load could be affected by sentencing guidelines and, as well, could lead to new types of unanticipated sentence disparities.

Minimum and Maximum Penalties

The maximum sanction allowed by statute for any specific offense is usually the only guidance a judge is given in determining the appropriate penalty for any case; the minimum statutory penalty is only of importance to sentence length for second offenders of certain types of crimes and for, at least at the time of the sentencing guidelines construction, violations of weapon offense legislation. (As of this writing, minimum sentences of a particular length have recently been mandated for a handful of other offenses as well.)[53]

Even though statutory minimum sentences are not required for most Massachusetts offenses, the judge is forced, by statutory limitations on minimum sentence length for certain institutions, to take account of the minimum sentence when imposing penalties. The process is rather complicated. Massachusetts judges state not only a length of time, but also the place of incarceration, in their sentence dispositions. Offenders sentenced to the state prison in Massachusetts must serve at least a one year term.[54] This statutory provision, in effect, imposes a mandatory minimum sentence of three years on state prison sentences, taking into account, of course, the one-third parole eligibility rule that applies to sentences for most offenses. A judge could not, therefore, impose a six month state prison sentence; in light of the statutory requirements about state prison sentences, a six month sentence would be illegal. (This example reverses the logic that must govern the judge's determination of the length of a House of Correction sentence where the maximum cannot exceed 2 1/2 years: in setting a state prison sentence, the judge must keep the 1 year minimum in mind, while in setting a House sentence, he must keep the 2 1/2 year maximum in mind.) Thus, the minimum sentence must be considered by judges in conjunction with the choice of place of incarceration before sentence is imposed.

Appellate Review of Sentence

Criminal justice practitioners, philosophers, and others who have expressed concern about disparity in the sentencing of offenders note that, more often than not, criminal sanctions are not subject to review. We discussed several of the proposed remedies that would provide for a review of the use of judges' discretion in sentencing, such as sentencing institutes and councils and appellate review of sentence, in Chapter 2. One of those remedies -- appellate review of sentence -- is available in Massachusetts to offenders sentenced to the state prison.

The Appellate Division of the Superior Court,[55] a tribunal appointed by the Chief Justice of that court, is authorized to hear appeals from any state prison sentence given to male offenders[56] and from sentences to Framingham (the female penititentiary) of five years or more given to female offenders.[57] The Appellate Division usually sits twice a year (in Wrentham, Massachusetts) for a period of several weeks to review the sentences of offenders who have requested appellate review within 10 days of their conviction.[58]

Sentences of either excessive harshness or excessive leniency may be the subject of appellate review.[59] In practice, however, since only the defendant, and not the government, may request review by the Appellate Division, and since offenders do not appeal a sentence they considered lenient, the job of the Appellate Division is narrowed somewhat to the review of excessive sentences. The power of the Appellate Division to adjust sentences, however, is not so restricted. An offender who appeals his sentence to the Division may have the sentence either decreased or increased or the Appellate Division may leave the original sentence unchanged.[60] Although few statistics are available to prove that the power of the Appellate Division to increase sentence deters some offenders from appealing their sentences, the fact that, on rare occasions in the past, the Division has actually imposed a more severe sentence probably has had some effect on the number of cases appealed.[61]

Parole in Massachusetts

We have already noted that the place of an offender's incarceration is tied to the length of the offender's sentence. A House of Correction sentence, for example, is bounded on the upper end of the scale by a maximum sentence of 2 1/2 years; a state prison sentence is bounded at the lower end of the same scale by a minimum time to be served of at least 1 year -- and more for certain types of offenses.

Our interviews with judges and others involved in the sentencing process in Massachusetts confirmed our belief that the sentences imposed were based on projections by judges of the amount of time that an offender would actually spend incarcerated in the particular institution to which the offender had been sentenced. A robbery offender sentenced to 2 1/2 years in the state prison would not serve the same amount of time "inside" as he would if sentenced to the House of Correction; and judges not only know this, but they also base the sentences they impose on their understanding of the rules that govern parole eligibility for institutions and certain types of offenders and offenses in the state.

During our reanalyses of the Massachusetts sentencing data, we will employ the rules governing parole eligibility to generate comparable "time to be served until parole" estimates; so that our readers will understand how the various time limitations and parole computations affect sentences in Massachusetts, however, those rules are described in detail in the sections to follow and are summarized in Table 6.1 below.

Insert Table 6.1 here

Institution of Confinement

In general, the institution where the offender is sentenced to serve time has the most differential effect on the actual length of time he will be incarcerated. Parole eligibility at Houses of Correction in Massachusetts is computed at one half of the sentence pronounced.[62] State prison sentences are reduced, for purposes of parole eligibility, by one-third for violent offenses, or by two-thirds for all other offenses, of the imposed sentence.[63] Should the offender be incarcerated for several consecutive sentences, the rules above still apply, however, parole eligibility is based on the aggregate minimum term for each type (i.e., one-third or two-thirds) of sentence.

We noted above that it is the institution named on the sentence disposition that governs parole eligibility. That an offender is sentenced to a certain institution, however, does not ensure that he will actually serve his time at that institution. The Department of Corrections has the power, after sentence is pronounced, to determine the final place of the offender's incarceration. But this power does not negate the judge's sentence -- parole eligibility is still based on the institution of sentence, regardless of where the offender eventually is incarcerated.

While there is no stipulated minimum amount of time that must be served on a House of Correction sentence, an offender sentenced to the state prison must serve at least one year (or 2 years for violent offenses) before eligible for parole.[64] And, as we mentioned earlier, while there is no stipulated maximum limit on a state prison sentence (apart from that which applies to the offenses), a House of Correction sentence may not be for any length of time that is more than 2 1/2 years for a single sentence.[65]

Parole eligibility for offenders sentenced to the Massachusetts reformatory at Concord -- for terms commonly known as "Concord sentences" -- is determined by a different

Table 6.1: Massachusetts Parole Eligibility - Sentence Conversion Rules

Institution	Offense/Offender Type	Sentence Minimum or Maximum	Parole Eligibility
House of Correction	Any	2½ year maximum per single sentence	1/2 of minimum sentence
State Prison-Walpole	Non Persons; not committed while on probation or parole	1 year minimum	1/3 of minimum; not less than 1 year
	Persons; committed while on probation or parole	2 year minimum	2/3 of minimum; not less than 2 years
	Persons	Life sentence-not Murder 1st Degree	15 Years
	Persons	Murder 1st Degree	life
Concord	No Prior Incarcerations	6 years	6 months
	No Prior Incarcerations	12 years	12 months
	No Prior Incarcerations	18 years	18 months
	No Prior Incarcerations	24 years or more	24 months
	Prior Incarcerations	6 years	12 months
	Prior Incarcerations	12 years	18 months
	Prior Incarcerations	18 years	24 months
	Prior Incarcerations	24 years or more	24 months

set of time "translation" rules.[66] Generally, offenders serve six months of every six years of their stated sentence. Thus, an offender sentenced to six years at Concord will serve a six month sentence, while a offender sentenced to 24 years or more will serve approximately 24 months. Also, 24 months is the maximum amount of time that usually can be served at the Concord correctional facility.

Whether the offender has previously been incarcerated is an important factor, however, in computing "time to be served before parole" for a Concord sentence. If the offender has previously been incarcerated, the minimum amount of time that he will have to serve at Concord is 12 months. Otherwise, the 6 months per 6 years rule applies. Thus an offender who has a prior record, and who is sentenced to 6 years at Concord, will serve 12 months, while his counterpart who has not previously been incarcerated will serve only 6 months. However, the maximum amount of time served at Concord is still 24 months, whether or not an offender has been incarcerated before; thus if the offender has previously been incarcerationed, it does not much matter if the sentence he was given was 18 years or 24 years, since he will still serve 24 months at Concord.

Parole Eligibility for Special Offenses

The second most influential determinant of differences in the amount of time that an offender will serve before becoming eligible for parole is the type of offense that the offender is convicted of. We mentioned above that most offenders sentenced to the state prison in Massachusetts are eligible for parole at one-third (or, at not less than 1 year) of their sentence. Offenders convicted of several specific offenses, however, are not eligible for parole until two-thirds -- or at a minimum, 2 years -- of their sentence has expired.[67] These offenses, commonly referred to as "persons offenses" -- a label descriptive of the violent nature of the offense involved[68] -- numbered 23 at the time of the guidelines development in 1978.[69] For the purposes of determining parole eligibility at two-thirds of the sentence of the minimum sentence (subject to a minimum of two years), the Annotated Laws of Massachusetts define a "persons offense" as being one of the following crimes:

Offense	Statute Ch./Sect.
Manslaughter	265/13
Indecent Assault (and Battery) on Child Under 14	13B
Mayhem	14
Assault with Intent to Murder or Maim	15
Assault and Battery with a Dangerous Weapon	15A
Assault with a Dangerous Weapon	15B

Attempt to Murder by Poison	16
Robbery Armed with a Dangerous Weapon and Assault while Masked	17
Assault with Intent to Rob while Armed	18
Armed Assault in a Dwelling	18A
Robbery, Unarmed	19
Assault with Intent to Rob, Unarmed	20
Confining/Putting in Fear for Purpose of Stealing	21
Rape	22
Rape of Child under 16	22A
Rape of Child	23
Assault with Intent to Rape	24
Assault on a Child under 16 with intent to Rape	24B
Attempt to Extort by Threat	25
Kidnapping	26
Incest	272/17
Unnatural Acts	35
Unnatural Acts with a Child under 16	35A

The most notable violent "persons offense", first degree murder, is absent from the above list because statute mandates that offenders convicted of it receive a life sentence and are not eligible for parole.[70] Life sentences allowed by statute and imposed for other types of offenses are parolable, however, after fifteen years.[71]

Finally, the Massachusetts statutes delineate one further circumstance when two-thirds of the minimum sentence (rather than one-third) is to be the basis for determining parole eligibility. Offenders committed to the state prison for offenses committed while on parole must also serve two-thirds of the minimum sentence (or, at least, 2 years).[72] Note, however, that while it is possible to advance the two-thirds eligibility to one-third for certain cases, offenders convicted of crimes committed while on parole do not have this option.

An Overview of Sentencing and Parole: Sentencing Guidelines in a Climate of Reform

The system of rules we described above that judges must use to set sentences and that the parole board must abide by in releasing offenders appears to be an extremely intricate one. Yet, the experience of practiced professionals (judges, prosecutors, defense counsel, and others) in the Massachusetts criminal justice system has given them an intuitive understanding of the workings of the sentencing and parole process. This is not unusual: in much the same way that inmates from a New Jersey state prison could estimate the

typical or "going" sentence for certain offenses and offenders, and the actual amount of time an offender would serve on a particular sentence (as we described in detail in Chapter 5 of this report), criminal justice system staff in Massachusetts could also provide sentence and time served estimates. (No doubt Massachusetts inmates could do the same but, unfortunately, we did not have time or funds to find that out.) The point here is that, although the statutory provisions about sentencing and parole are complex, the rules have been incorporated into the decision-making processes of district attorneys and defense counsel who offer sentence recommendations, and judges who impose the ultimate sanctions. And, unfortunately, each of these decision-makers (to say nothing of the legislature) had somewhat different translations of these rules into actual sentencing policies; hence, the source of some of the dissatisfaction with the current sentencing system.

Few of the sentencing and parole regulations are new; although considerable debate about reform of the sentencing structure occupied the Massachusetts legislature (and, on occasion, some aspects of that structure have been changed), the basic structure of the system had not been altered since the mid-1800's. Criminal justice personnel, thus, had reasonable expectations about how the sentencing and parole systems operated -- based on their perceptions of the past operation of that system. Those personnel, because they had a feel for the way the system operates, also provided an excellent resource group that had the ability to project, or hypothetically simulate, how new changes -- such as sentencing guidelines -- would affect the operation of the sentencing and parole systems. Perhaps more valuably, these same judges, prosecutors, and defense counsel had an understanding of the local political climate that we described earlier, and that importantly affects why changes in the sentencing and parole systems occur and, as in comparison with the Pennsylvania guidelines described in Chapter 9 of this report, why some changes are less effective than others. Thus they could try to project not only the ways in which guidelines might change the sentencing process but, also, they could project how the concept of guidelines would be received and how the political climate would affect a guidelines model's structure.

We noted earlier that in the midst of all the legislative reform of various sentencing statutes, the Massachusetts Superior Court judiciary were developing their own version of sentencing reform -- Superior Court sentencing guidelines. In the summer of 1979, the construction of those guidelines had been almost completed and the judges of the Superior Court's Committee on Probation and Parole anticipated proposing the guidelines to the Superior Court for formal adoption in the

fall of that year. Our project staff had by then some general notion of what the guidelines to be proposed were going to be like, and we wanted to find out how the guidelines would be received by judges, prosecutors, defense counsel, and other personnel in the Massachusetts criminal justice system who would be using them, or who would be affected by, the guidelines operation. In light of the political climate about sentencing reform in the legislature, our project staff believed that the initial reactions of those in the Massachusetts criminal justice system to the general concept of sentencing guidelines would have a tremendous effect on the acceptance of the final guidelines model that the Court developed, and on the use of that model in the sentencing structure we described in the last section.

To find out what judges, prosecutors, defense attorneys, and others thought about sentencing practices throughout Massachusetts, and about the concept of sentencing guidelines in particular, staff of the Evaluation of Statewide Sentencing Guidelines project conducted interviews in selected counties in that state during the summer months of 1979. Beyond the selection of those counties in which interviews were to take place,[73] the selection of persons to be interviewed[74] and, indeed, the interviews themselves, were relatively unstructured.[75] In general, we interviewed every available Superior Court judge then assigned to the counties that we had selected,[76] several District Court judges,[77] the district attorney and the chief attorney for the Massachusetts Defenders Committee in each of those same counties, a sizable number of assistant district attorneys[78] and defense counsel (both public and privately retained) and, finally, the chief and assistant probation officers of each county. During that same summer, we also spoke with the Chief Justice of the Supreme Judicial Court and a number of additional court personnel at the state level. Tables 6.2 and 6.3 detail the number and type of personnel interviewed during 1979. (The table also includes comparable figures for interviews conducted during the summer months of 1980 after the release of the sentencing guidelines model. We will have more to say about the results of those interviews in later chapters.)

Insert Tables 6.2 and 6.3 here

The interviews focussed on a number of topic areas: how criminal cases were processed in Massachusetts, what sentencing in Massachusetts was generally like, what factors seemed to be important (at least in their past experience) to judges' sentencing decisions, did much sentencing disparity exist, what did they think of the concept of sentencing

Table 6.2: Massachusetts Criminal Justice System Staff Interviews in Four Counties: 1979 and 1980

County	Prosecution 1979	Prosecution 1980	Public/Private Defense 1979	Public/Private Defense 1980	Superior Court Probation 1979	Superior Court Probation 1980	Total
Hampden	11	4	6	4	11	6	42
Plymouth	6	5	2	3	6	5	27
Suffolk	17	13	16	10	4	10	70
Worcester	1	5	5	5	9	2	27
Total	35	27	29	22	30	23	166

Table 6.3: Massachusetts Interviews with Judicial, Legislative and Correctional Personnel: 1979 and 1980

Year	Superior Court Judges	Judicial Court Personnel	Corrections/ Parole Staff	Legislative	Total
1979	24	4	—	—	28
1980	32	7	3	3	45
Total	56	11	3	3	73

guidelines, would sentencing guidelines have any effect on sentencing disparity (to the extent it existed) in the Massachusetts system, and how that effect would be manifested.[79]

In their assessment of current sentencing practices, and the extent of disparity evident in those practices, judges, prosecutors, defense counsel and others held very similar views. All three groups claimed that the nature of the offense and certain background factors about the offender were important judicial considerations in the choice of sentence disposition. In addition, judges acknowledged that the recommendations of prosecutors and defense counsel did, indeed, hold some implications for their final sentence decisions. These same major types of factors, i.e., offense and offender factors, were also claimed by prosecutors and defense counsel to weigh heavily in their recommendations offered to sentencing judges as to what sentence to impose.

What were those offender and offense factors supposedly accorded weight in sentencing decisions or recommendations? They were not very much different from those consistently hypothesized, and often found to exist, in sentence decision-making research.[80] Offense factors mentioned frequently in our interviews with judges (and others involved in the sentencing process) involved the numbers and types of charge(s) against the defendant and the background history of the offense(s), an assessment of the strength of the evidence, specific victim risk factors present during the commission of the offense, the role of the offender in the offense and in a "criminal network" generally.

The "background history of the offense", of course, concerned the circumstances peculiar to the present offense. The judges, and others posed questions such as: What factors precipitated this particular offense? Did the offender know the victim? What was their relationship?, Was this an isolated event, or the culmination of a series of events? What was the physical state of the involved parties -- e.g., was this a bar fight where everyone, both victim(s) and defendants were drunk? What did the defendant do specifically? How does what the defendant did relate to this (or these) charges against him? -- and other similar questions. In other words, judges (and others) were interested in discovering what criminal events actually took place and what caused them.[81] And the determination of what events actually took place could not be made simply from a review of the charge(s) against the defendant. Massachusetts judges referred often during the course of our interviews to the "charge package" and to the "sentence package." These are two distinct items; the "charge package" represents the

district attorney's choice of charges to be prosecuted against a defendant, while the "sentence package" represents the type of sanction imposed corresponding to the charge package.

The charge package is composed of all of the offenses with which the defendant is charged or convicted, and there appear to be several types of charge packages, which can be classified, more or less, into the following groups:

1. A group of distinct and similar offenses, for example, three robberies committed on different occasions;

2. A group of distinct but not similar offenses, for example, two robberies and one rape; and

3. A major offense, and charges that can be said to be secondary, or included, in the major offense, plus the possible addition of trivial offenses, such as the following group of offenses: armed robbery, assault with a dangerous weapon, and running a red light (called, in Massachusetts, a red lens violation).

Most Massachusetts offenders who are charged with more than one offense can be classified, by the types and numbers of charges pressed, into one of these three groups. The important point to keep in mind, however, when assessing the validity of the charge package, and the corresponding sentences, is the question of how closely the actual criminal behavior is mirrored by the charges pressed. If each charge is indicative of a separate criminal act stemming from a separate criminal event, then the importance of each charge may be mirrored in the final sentence accorded to that charge. However, if the package of charges represents really one criminal act -- usually taking place on the same occasion -- the judge must determine which charge is the really important one.

Why would the package of charges not represent accurately the actual behavior involved? The question here concerns the district attorney's purpose[82] (underlying or manifest) in settling for a set of specific charges. As far as we could decipher these purposes, they too could be classified into three groups:

1. The charges may reflect the actual behavior of the defendant, i.e., the defendant did commit three separate robberies;

2. The charges may be a result of overzealous charging practices and the prosecutor's desire to have

some charge to "hang" on the defendant; and if this is the case, a particular charge will be used for negotiation while others are filed; or

3. The charges may be present to allow the state flexibility in terms of prosecution and sentencing. In such a case, the district attorney will push the charge that best fits either the offense itself (usually that offense that can most easily be proved in evidentiary terms), or the desired disposition.

One might question what the relation is between the charge package and the district attorney's desired disposition -- one factor of importance in point 3 above -- especially since the district attorney can offer a sentence recommendation in court as to what would be the preferable sentence anyway. Actually, though, there is a relation between the desired disposition and the charge package, and it is a very important one. We noted in an earlier section of this chapter that, although the Massachusetts sentencing system is principally indeterminate, some statutes of that system require that a judge not suspend sentence for particular crimes.[83] In operation, the requirements of those statutes mean that a judge must either sentence the offender convicted of such a crime to incarceration[84] or place the person on probation. And in some cases, neither sort of disposition standing alone, judges believed, would be appropriate. As an example, a sentence may not be suspended for the offense or armed burglary. If a young first offender broke into a house armed with a jackknife, that offender's sentence would, by statute, have to be either incarceration or probation if that were the only offense that he were charged with. A judge might think the incarceration sentence too severe for the case, yet also think that the probation sentence would not adequately reflect the gravity of the conduct. If, however, there were additional charges that did not require this either/or choice, but would allow for a suspended sentence or a jail sentence, such as a charge of being armed with a dangerous weapon, the judge could sentence the offender to probation for the armed burglary, and stipulate a suspended sentence for the second charge of armed with a dangerous weapon.

What is needed, therefore, is some way of assessing the correlation between the real behavior involved in an offense and the charges that have been pressed, in light of the statutory requirements as to sentence. Judges, in Massachusetts, claim to be able to perform this assessment and view charges as "packages" of one of the three earlier mentioned kinds. The particular class of the charge package will, apart from other offender and offense variables, help to

explain the final sentence given to an offender -- and that sentence is mirrored in the "sentence package." And, in the same way that charges must literally be viewed as a "package", so must sentences -- in other words, a judge may not intend a sentence he attaches to a particular charge as the sentence that that charge, on its face, deserves; rather, the sentence is part of a package, including other offenses and other charges, and while there is not one-to-one correspondence in severity between particular charges and sentences, such correspondence does (according to judges) exist between packages of each. (For the above example, the sentences should really be reversed to correspond to the appropriate charges.) A sort of a matching exercise thus takes place. But it is a matching exercise that all actors (i.e., defense and prosecution as well as judges) in the Massachusetts system are aware of and use daily. District attorneys, after all, bring the charges -- and the point here is that they do so with the intricacies of the Massachusetts legal sentence structure in mind.

The charge package of offenses appeared, from our interviews, to be the major offense factor considered by judges and others in determining sentence and sentence recommendations; but it was not the only item of importance. The second item mentioned -- an assessment of the evidence -- sounds like a factor that should influence the prosecutor's decision to bring a case to court at all, more than it should influence the judge's decision as to the appropriate sentence. And, of course, for prosecutors (as well as for defense counsel) this is a major consideration: both must balance the likelihood of conviction at trial against the possibility of a guilty plea. It is a factor that judges also mentioned that they consider, however, for several reasons. First, their consideration of the evidence is necessary for a determination of guilt where there is a trial, but the jury has been waived: in such cases, the judge is the sole arbirator of the ability of the evidence to prove the offender's guilt. Secondly, and perhaps more importantly, the majority of cases in Massachusetts, as in most jurisdictions, are disposed of by means of a guilty plea; a judicial consideration of the strength of the evidence, judges believe, acts as a control on the prosecutors charging discretion and also gives them some information about the facts of the case that can be proved.[85] And many of the facts of the case involve additional factors about the offense and the offender which they also claimed to consider when imposing sentence.

The third category of offense items the judges claimed would be found to be relevant to their sentence decisions involved a "victim at risk" assessment. Items considered along this dimension were those such as the following: Had

the offender used a weapon during the crime? What type of weapon was it -- in other words, was it a type of weapon that could cause extensive injury or death to the victim? Did any injury to the victim occur? Was the victim particularly vulnerable to injury? Judges also repeatedly emphasized the fact that actual injury did not have to occur for that item to weigh in their decisions -- rather, if the offender merely placed a victim in a situation were there was a substantial likelihood of injury then, even though the injury did not occur, the risk that it could have would have weighed in the sentence determination.

A final category that judges claimed influenced their decisions was that of the role of the defendant in the particular crime: was he the principal actor in the offense, or did he play only a minor accessory role? Also, the judges said that the role of the offender in any one single offense was, at times, not so important as his known involvement in a criminal network in the past.

Several items of information about the offender, in addition to the offense items we just discussed, were consistently said to be important in the sentence choice. Almost without exception these items include and give dominant weight to:

1. prior criminal history
2. age
3. history of substance abuse
4. family stability/involvement in criminal culture
5. poverty
6. employment
7. offender repentence -- evidenced by a guilty plea.

Prior criminal history, as could have been predicted, was the major item cited by judges as having a major role in their sentence decisions. But how did these judges define "prior criminal history"? It appeared, from our interview responses, that offenders' prior criminal histories are examined by judges first, to determine the number and the seriousness of the prior offenses and, secondly, to determine whether a pattern exists in the prior criminal conduct of the offender that has bearing on the current charge. There were, evidently, several types of prior criminal history patterns discernible to judges. A consistent pattern would be evidenced by a series of prior offenses similar to the current offense. An escalating pattern would be one where the offender, during the start of his "criminal career," had been convicted of minor offenses and, as his involvement with criminal activity continued, the offenses became progressively more serious. The final type of pattern is not really a

pattern at all -- it involves the first time offender who might deserve, in many judges' opinions, a mitigated sentence.

It may seem odd that the judges often indicated that they did not look to patterns mentioned above to indicate a higher (or lower) culpability level, or to automaticaly increase the length of the sentence; rather, the pattern of prior offenses evidenced by an offender was seen by judges as a sort of "flag" to inform them as to the character of the individual to be sentenced. There was an apparent tendency, however, for judges to weigh serious prior arrests and convictions more heavily and to retain the assessments of these offenders in mind when, at a later stage in the process, equating type and length of current offense and sentence. To stress the point, the number, the offense type, and the pattern of prior offenses, in combination, are of importance to all of the judges.

The other offender factors listed above -- age, substance abuse, family stability, poverty, and employment -- really played a secondary role in terms of importance to the sentence outcome compared to the prior record dimension of the offender's background. This is not, however, to say that they did not play a role at all, or that that role did not vary in terms of relative importance depending on the offender's prior record and the nature of the current charge package. These factors -- with the exception of the final one -- have often been labelled, especially by guidelines developers, as "social status" variables, and a good deal of controversy has been generated about the appropriateness of their consideration in reaching sentencing decisions. We believe them to be self-explanatory, and so will not waste much time here discussing them -- except to say that exactly how these items should figure in sentencing decisions -- in other words, whether the item should mitigate or aggravate sentence -- was not agreed upon uniformly either by judges, or by any of the other system actors that we interviewed.

The final factor -- offender repentence, usually as evidenced by a plea of guilty -- deserves somewhat more discussion, however. Of course, criminal justice literature has consistently identified the mitigating influence of the plea of guilty on sentences, and these findings have never been unexpected. The reason usually given for reducing the sentence after a guilty plea is that the plea saves the state the cost, and the criminal justice system the delay, of a trial; and it saves the victim from public scrutiny (and sometimes humiliation) on the witness stand. Massachusetts judges, by and large, did not mention these as the primary justifications for mitigation;[86] rather, with only slight exception, the Massachusetts judges viewed the guilty plea as

a verbal sign of the offender's repentence. And if the offender was already repentant, it was argued that he would be an excellent candidate for rehabilitation -- a goal of sentencing that these judges also adamantly believed could not be effected in prison; hence the justification for a mitigated sentence (and often one of non-incarceration).

Following (or perhaps concurrent to)[87] the consideration of important offender and offense information, judges (as well as the prosecutors and defense counsel) claimed to need additional information related to prior criminal justice system responses (if any). In other words, judges claimed to incorporate not only the prior record of the defendant, but also the prior dispositions, into their decision as to what would, at this stage in the particular offender's career, be an appropriate sanction for this charge. Thus judges, prosecutors, and defense counsel wanted to know both the type and the length of sentences for previous convictions; it seemed to us that these decision-makers were interested in finding out what previous decision-makers involved with a particular defendant had thought would be appropriate sentences for his earlier offenses. Secondly, the decision-makers were concerned with whether the offender had previously been on probation (with either successful completion of that sentence, or with eventual revocation of it[88]). Finally, judges in particular, but the other actors as well, were given at this juncture to making predictions about the "shock value" effect of a possible heavy current sentence -- a sentence that would initially be on the heavy side, but one that would later be followed by a judicial order to "revise and revoke" the sentence to something lighter. The purpose of such a procedure is first to scare the wits out of the offender and then later adjust the sentence to what would be equitable for the offense and offender facts.[89]

Discussion of offender variables brings up another inter-related topic -- the decision method used by judges (as well as by those making recommendations) to determine the appropriate sentence and the sentence's relation to the charge/conviction package. We have already brought up the example of a case where, because of statutory requirements, the sentences had to be given on charges other than the ones on which they really should have been given, so as to reach what the judge thought of as the appropriate result.[90] There are, in addition, several other factors both about the offense (similar to the problem of the charge package above) and the offender, that are also considered by judges -- so they claimed -- in reaching appropriate sentences in particular cases.

It was quite clear to us, both from our discussion with judges[91] and from our independent observation of sentencing hearings and dispositions, that the following method (though not necessarily in order of presentation) is used by judges in determining sentence: judges decide

1. the actual length of time that they want an offender to serve (if any);

2. the place of confinement, either in agreement or not with the actual amount of time imposed;[92] and

3. how to dispose of the charges -- by fitting the sentence they want to impose to the particular charge that will allow that sentence, and by filing,[93] dismissing, or imposing concurrent sentences on the remainder of the charges pending sentence.

One point deserves particular emphasis at this juncture: judges do, indeed, sentence on the basis of their estimates of precisely how much time the offender will serve on the sentence. Of course, the Massachusetts Department of Corrections has considerable statutory authority to release prisoners in advance of the usual eligibility date (i.e., the date found by using the one-third, two-thirds rules described in the last section) and judges are aware that the Department does exercise its early releasing power. However, judges do not have any knowledge, at the time they sentence, that an offender might serve less than one- or two-thirds time; thus the sentence they impose is based on a reasonable estimate of the amount of time the offender would serve. Also, since judges do not have any power to stop the Department of Corrections from exercising its early release authority, there really is very little -- other than basing the sentence on such a reasonable estimate -- that a judge could do.

A final aspect of sentencing in Massachusetts is one that may have the most influence[94] on offenders' dispositions: this is the sentence recommendation process. The recommendations of district attorneys and defense counsel play a part in the judge's determination of the appropriate sentence. The majority of the judges that we interviewed indicated that these recommendations, along with the recommendations of probation officers as to the defendant's suitability for probation, were given some consideration. The importance of the recommendation was enhanced if the defendant had pleaded guilty to the charges (or to some of them); this was based on the ground that these system personnel, who had professionally advised the client as to the best course of action, knew the offender and the real facts of the case much better than did the judge in the plea situation. (Judges

repeatedly stated that sentence recommendations did not play as much of a sentence-determining role in trial cases, for they believed that they knew as much about the defendant and the crime as anyone else did after a trial.)

The degree of reliance by the judges on the recommendations of defense, prosecution, and probation was far from uniform both among judges, and across areas of the state. A number of factors extraneous to the actual case at hand (and to whether or not the case involved a guilty plea) appeared to have significant bearing on the actual weight accorded to sentence recommendations. These additional factors are quite diverse, as is evident from the follwing list:

1. the background of the sentencing judge, i.e., the judge's experience in practicing criminal law prior to appointment to the bench, and his or her length of term on the bench;

2. the reputation of the district attorney's office in a particular county, and the background and expertise of the office's staff;

3. the reputation of the Massachusetts Defender's Committee (or, alternatively, the specific private or appointed defense counsel) in a particular county, and the background and expertise of this staff;

4. the reputation of the probation office in a particular county and the background and expertise of this staff (and it should be mentioned that assessments of probation department expertise levels were especially varied across the state); and

5. the office policies developed regarding the handling of certain types of offenses for each of these three offices in any particular area; this includes: the policies governing the filing of charges, with or without a plea of guilty; and the policies governing the binding over of cases under concurrent jurisdiction from District to Superior Court.

Perhaps the most important influence on the judicial evaluation of the proferred sentence recommendation from the list above has to do with its quality and utility, as estimated by the judges' evaluation of whether or not the recommendation is offered by someone with adequate experience and perspective to assess the case at hand, and available sentence options, realistically. At the time of our interviews, Massachusetts judges freely acknowledged that sentence recommendations were treated with the professional

respect they deserved. Thus, suppose that a judge, thought a the recommendation was "out of line" from what a "normal" recommendation would be -- bearing in mind that what is judged as "normal" is different for prosecutors, who recommend on the "high" side of a sentence, than it is for defenders, who would, conversely, recommend on the "low" side of the same sentence. If the judge thought that the reason that the recommendation was out of line was due to the relative inexperience or otherworldliness of the offerer, then the recommendation would be accorded substantially less weight. While judicial perceptions of the quality and experience of system actors is not an easy item to capture on paper, these perceptions apparently exerted considerable influence on sentences, insofar as those were affected by sentence recommendations.

It is pertinent to mention here that the judge in a case is not the sole determinant of the sentence recommendation offered; nor is the particular judge's evaluation of the proffered recommendations the sole determinant of the sentence outcome. One additional assessment process precedes the judge's entire involvement with a set of defense and prosecutor recommendations. Defense counsel and prosecutors, during the pre-court case conference process, assess each other's recommendations and, apparently, try to pull in the slack between them where possible. The eventual importance of the sentence recommendations results from a balancing of the separate interests of all of the parties involved.

One final point deserves mention at this juncture: although sentence recommendations are offered by defense and district attorney's in almost every case (and by probation staff in some cases), judges, by and large, did not feel that they were "bound" by any of the offered recommendations. If the recommendation was "agreed" -- which is to say that the defense counsel and the prosecutor acknowledged to the judge that the recommendation offered was agreed to by both parties -- judges felt slightly more inclined to follow the suggested sentence than not. However, if the recommendation was not "agreed", judges indicated to us that they usually would sentence as they thought appropriate. It also occurred, however, that the district attorney's recommendation provided the "upper limit" for their sentence disposition. This tendency -- to use the district attorney's recommendation as the "upper limit", or highest sentence that they would impose -- also was enhanced if the recommendation had been agreed to by both counsel. It may seem odd, given that similar general types of factors were considered by each of the groups, that all three groups also acknowledged the existence of disparity in the sentences imposed in Massachusetts. The result is not impossible -- or even odd -- of course. Sentences perceived

as disparate (and that is not to say that such sentences are disparate) could result even though similar types of factors were claimed to be considered in cases generally. The perceived disparity in sentences could be a result of either different emphasis being placed on the same factors by all of the parties involved, or by different emphasis being attached to some factors versus other factors depending on either the specific case at hand or by the specific party making the recommendation. And all of the persons that we interviewed in Massachusetts acknowledged these possibilities.

One reason we conducted interviews in Massachusetts before the guidelines went into effect (aside from the fact that we wanted to learn about how the system in that state operated before guidelines) was to assess the initial reaction of judges, prosecutors, defense counsel and others involved in sentencing to the concept of sentencing guidelines as a disparity-controlling mechanism. Of course, all of the persons that we interviewed could only provide supposition about the reception and use of sentencing guidelines because, at that time, the guidelines model had not yet been released; even with this limitation in mind, our interviews did provide some intriguing responses about the general concept of guidelines and about their possible effects on the Massachusetts system of sentencing offenders.

The concept of sentencing guidelines was given mixed reviews by the judiciary and staff of the Massachusetts criminal justice system. That sentencing guidelines could serve as a mechanism to reduce sentence disparity was not a bone of contention to those interviewed;[95] however, exactly what constituted disparity, and how that disparity should be reduced by the guidelines, appeared to be the controversial issues -- and issues about which each of the occupational groups interviewed differed.

Judicial evaluation of the concept of sentencing guidelines was generally ambivalent; only a slight majority of the judges indicated unrestricted support of sentencing guidelines. Most of the judges thought that the sentencing guidelines research was worthwhile in that it could give them some idea of how their colleagues sentenced.[96] All of the judges believed that they knew what facts about the offender and the offense they took into consideration in determining sentence, and they all believed themselves to be self-consistent in their sentencing. Doubt existed among judges, however, as to whether they were deriving sentences in the same manner as their colleagues. To the extent disparity existed (though the judges generally believed it did not) it was probably the product of different sentencing techniques among judges. Guidelines would, they thought, give them specific information about their colleagues' sentences.

What should be done with the information about other judges' sentences proved to be the divisive issue for the judges. Some judges just wanted to learn about past sentences from the guidelines project and leave it at that, while others supported the concept of guidelines specifically as a tool that stated the court's previous (and probable future) sentencing policies. One judge noted that the guidelines could give confidence in the sentence in the average case -- i.e., a plea disposition where the judge has little or no time at all to learn about the character of the offender or the particulars of the offense.[97]

There were two major reasons for disapproval of sentencing guidelines among the judges interviewed. A number of judges, who may have had some interest in the guidelines project's study of past sentences, did not think that the results of that study should become sentencing standards in the form of guidelines. The judges believed that the sentence decision process is one totally unique to every case; to the extent that guidelines set standards for similar cases, the unique aspects of every case would be treated unfairly. (A number of these judges, however, apparently equated guidelines with a detailed form of mandatory sentencing.)

Secondly, a number of judges were adamantly opposed to any type of judicial sentencing guidelines at all. Aware of the continued legislative interest in sentencing reform, these judges believed that any structuring of the sentencing process should be done by the legislature, not by the judiciary.

Individual judges raised additional concerns about the guidelines structure and possible side-effects. If guidelines, for example, were to be so broad as to include all types of past sentences, what would be the purpose of having them at all? Yet, judicial discretion to vary sentence for the unique case should still be available to the sentencing judge without requiring a specially-justified "departure", so the guidelines range should give the judge some "running room". Also, while all of the judges said that it would not be difficult for them to state their reasons for imposing a particular sentence in "departure" cases, a reluctance to do so was evident because a number of judges believed that the practice would open up a new area for appeal of sentence. Judges noted that the factors included in the guidelines, as well as those used for deviating from the guidelines, might be challenged for appropriateness on appeal.

These same concerns, in essentially the same form, were echoed by prosecutors and defense counsel throughout the state. Again, a concern with the judicial authority to develop guidelines, with guidelines ranges and items included,

and with appellate litigation was evident. However, these concerns were now discussed from the respective vantage points of prosecution and defense -- both parties saw the need for control of judicial discretion and sentence disparity,[98] but each party saw sentencing guidelines as a potential tool of the opposition. If the guidelines ranges were fairly severe, for example, defense counsel believed that more trials (rather than guilty pleas) might result because a defendant would have "nothing to lose" by going to trial, since he would not get any worthwhile reduction in sentence as a consequence of his plea. If the guidelines ranges were too low, prosecutors felt that the public interest would not be served. Both prosecution and defense reiterated the belief that every case is unique -- but obviously for different reasons.

The fears of the prosecution and defense can really be summarized briefly: the prosecution was afraid that the sentences called for by the sentencing guidelines would be too low, while defense was concerned that the guidelines sentences might be too severe. Several defense counsel noted that a guidelines model could just provide the illusion of restraint of judicial discretion. After all, they noted, the model first provides a guidelines sentence, then a guidelines range, and then allows for exceptions. That system could end up, after exceptions are invoked, being just what they now had in operation.

Probation staff had only a tangential interest in the concept of sentencing guidelines -- the role of the probation department staff in the guidelines operation. Their concern was not misplaced, nor should it be derogated: probation staff, usually responsible for preparing actual cases for sentencing under most guidelines models in operation to date, have a tremendous (as yet mostly unstudied) power over criminal sentences.

Whether any sentencing guidelines model is eventually approved of by judges, prosecutors, defense counsel, and others, and how that guidelines model is used in practice in Massachusetts will clearly depend on how well two important issues are resolved. The first of these issues is whether the final form of the guidelines includes some realistic specification of sentence ranges and of offense and offender information. The second criterion, and one much more difficult to accomplish to everyone's satisfaction, is that of the proximity of the guidelines to what was perceived by all as the actual sentencing process. As we will see in more detail in the next chapter, in resolving both of these issues the Massachusetts sentencing guidelines project constructed guidelines that are largely the product of, and a response to, political and public influences.

Notes to Chapter 6

[1] The Massachusetts sentencing guidelines project stated that the guidelines to be developed would describe past sentencing practice, but qualified that claim by acknowledging that factors deemed unacceptable (for moral or ethical reasons, etc.) would not be incorporated in the guidelines and thus used as a partial basis for future sentence decisions. More will be said about this in the next chapter.

[2] Minnesota was the only other state in the business of developing statewide sentencing guidelines at that time, that did not make the claim that the guidelines would simply "describe" past practice (with or without a few minor changes to remove unacceptable factors); rather, the Minnesota Sentencing Commission clearly intended to prescribe sentencing policy, as we discuss in more detail in Chapter 9. However, generally speaking, the rhetoric of "descriptive, not prescriptive" has been central to the palatability of the sentencing guidelines concept, as we discussed earlier in Chapter 3 of this report.

[3] That one should have some notion of what past sentencing practice was, is important not only to those who may be evaluating a guidelines model; presumably it should be of importance to guidelines developers as well. In fact, guidelines developers need to have a general idea of sentencing policies prior to the start of the research not only to avoid a particularly atheoretical research strategy but also to ensure that information items presumed to be important to decisions can be obtained from retrospective data or some other means collected and their actual importance statistically tested.

[4] The Massachusetts sentencing guidelines were constructed after statistical analyses of cases sentenced in Superior Court during parts of 1977 and 1978. We will have more to say about the time frame and sample chosen later in Chapter 7, but for the moment it is important to note that since the guidelines were based on 1978 sentence decisions, our concern is with the Massachusetts legal sentencing structure as of 1978. We do not purport, here, to be describing the current legal structure of sentencing in Massachusetts. (We will, however, note a number of significant changes to that structure as we go along.)

[5] The methodological and statistical critiques of the Massachusetts statewide sentencing guidelines follow in Chapters 7 and 8 of this report.

[6] The items listed were not unusual for that time period; they were in fact modelled after the factors found to be influential in the determination of bail by the Vera Foundation's Manhattan Bail Project study. The Massachusetts Bail Reform Act items were:

> ...on the basis of available information, take into account the the nature and circumstances of the offense charged, the accused's family ties, financial resources, character and mental condition, the length of residence in the community, his record of convictions and appearances at court proceedings, or of any previous flight to avoid prosecution or any previous failure to appear at any court proceedings.

It is also clear that the "open-ended" feature of language discussed in Chapter 3 makes this list of factors non-exhaustive; despite the specific mention of these items, it is not stated that items not listed here that a judge might deem important for a particular case should not be considered.

[7] The March 4, 1968 study by the Governor's Committee on Law Enforcement and Administration of Criminal Justice recommended that the Criminal Law Revision Commission address six major objectives during its tenure.

(1) Development of general principles of responsibility and culpability.
(2) Statutory restatement of the many areas of law which are now covered only in common law terms.
(3) Coherent organization of the statutory provisions which create crimes.
(4) Elimination of overlapping, fragmented and prolix definitions of, and inexplicable gaps between, existing crimes.
(5) A rational grading and classification of crimes, with an attempt to develop criteria for sentencing.
(6) Mandatory post-release parole periods for serious offenders, to facilitate their post-release readjustment to life outside prison.

[8] The summary to the code provided in MCLRC, 1972:5 states:

> The levels of punishment authorized are:
> Class A felony: maximum-life, minimum-ten years or 1/3 of maximum
> Class B felony: maximum-20 years, minimum-five years or 1/3 of maximum
> Class C felony: maximum-10 years, no minimum
> Class D felony: maximum-5 years, no minimum
> Class A misdemeanor: maximum-2 1/2 years, no minimum

Class B misdemeanor: maximum-6 months, no minimum
Violations: fine and probation, no imprisonment authorized

[9] Release decisions for the other classes of offenses were left to the discretion of the parole board "because information about how an offender responds to correctional treatment simply does not exist at the time of sentence....It is the assumption of this chapter that in general it is far better to allow the judgement about release to be made at the time when there is the best information available, and that time is after sentencing when the man has already been imprisoned." (MCLRC, 1972:6). However, the Commission noted that "the public sense of security requires an assurance that a dangerous offender cannot within a short time be released" (p. 6) and so allowed judges to set minimum sentences for serious felonies though noting that "with a responsible paroling authority, the practical effect of such a minimum is likely to be negligible" (p. 6).

The Act also would have abolished good time; minimum and maximum sentences would have represented actual time to be served unless paroled. Mandatory parole terms -- eliminating the practice of "maxing out" would have been set for the more serious offenses as well (MCLRC, 1972:6-7).

[10] Imprisonment was presumed necessary for the "protection of the public" if the court found (1) an undue risk of recidivism or (2) a need for institutional correctional treatment or (3) non-incarceration would depreciate the offense's seriousness (MCLRC, 1972:78). Clearly these standards are drafted after the Model Penal Code (ALI:1962).

[11] If the court did not find any one of the 3 general criteria (listed in note 10 above) that would negate a probation sentence present then 12 additional items were to be considered before deciding that the offender should receive probation in lieu of imprisonment. The Commission stated that "the following factors, while not controlling, shall be accorded weight in determining whether to place the offender on probation: whether

(1) the offender's criminal conduct neither caused nor threatened serious harm;
(2) the offender did not contemplate that his criminal conduct would cause or threaten serious harm;
(3) the offender acted under a strong provocation;
(4) there were substantial grounds tending to excuse or justify the offender's criminal conduct, though failing to establish a defense;

(5) the victim of the offender's criminal conduct induced or facilitated its commission;
(6) the offender has compensated or will compensate the victim of his criminal conduct for the damage or injury which was sustained;
(7) the offender has no history of prior delinquency or criminal activity or has led a law-abiding life for a substantial period of time before the commission of the present crime;
(8) the offender's criminal conduct was the result of circumstances unlikely to recur;
(9) the history, character and attitudes of the offender indicate that he is unlikely to commit another crime;
(10) the offender is particularly likely to respond affirmatively to probationary treatment;
(11) the imprisonment of the offender would entail excessive hardship to himself or his dependents;
(12) the offender is elderly or in poor health;
(MCLRC, 1972:78-79, emphasis added). Compare Ferri (1921), discussed in Chapter 2 above.

[12] The defendant in this case, Angelo Jackson, was arrested on 5 July 1975, 3 months after the Bartley-Fox bill took effect. He was convicted of carrying a firearm without a license and sentenced to the mandatory minimum term of one year in jail. The Massachusetts Supreme Judicial Court review was decided on 15 March 1976.

[13] Massachusetts Constitution, Declaration of Rights, Article 30.

[14] The "efficacy" of the Bartley-Fox law in particular has been questioned and researched in several studies since the statute's enactment. Generally, the research has proved inconclusive as to whether the mandatory sentence prescribed by that law has been a deterrent to serious violent crime as the Massachusetts legislature intended. (See Beha, 1977:96-146; and Beha, 1977:289-333).

[15] Massachusetts actually has a number of statutes that require minimum terms of imprisonment, if imprisonment is imposed, for specific offenses or for second offenders; however, there are at least two major differences between those statutes and the Bartley-Fox law. Prosecutors have the ability to reduce the charge or not to charge an offender at all with one of those crimes that requires a mandatory minimum sentence of this sort. Also, judges may file charges for these offenses. Neither prosecutors, nor judges, may so adjust or file charges that involve violations of the Bartley-Fox Gun law. Finally, and most importantly, the mandatory minimum sentence for these other offenses must be

imposed only if the judge sentences an offender to incarceration; in lieu of incarceration, however, a judge may suspend sentence or sentence the offender to a term of probation or impose a fine. The Bartley-Fox mandatory minimum demands that the defendant must be incarcerated and for a specific one year term.

[16] Since the enactment of the Bartley-Fox gun law, numerous other pieces of legislation have proposed mandatory minimum sentences for a varied group of offenses -- most notably, drug and auto theft offenses.

[17] At about this time, the Boston Globe "Spotlight Team" published a special report on rising crime rates, especially violent crime rates, throughout Massachusetts.

[18] In addition to the federal monies allocated to the initial sentencing guidelines feasibility study and the dissemination projects we discussed in Chapter 3, a considerable sum of state and federal monies was expended through the State Planning Agencies in a number of states for the purpose of studying sentencing practice or developing sentencing guidelines.

[19] This proposal, titled "Program to Review and Improve Sentencing in Massachusetts", was submitted by the Crime and Justice Foundation to the Law Enforcement Assistance Administration's Regional Office on 29 April 1976. The applicant for the funds was Henry J. Mascarello, who later, as a staff person in the Superior Court, was instrumental in obtaining funds for the sentencing guidelines project in that state.

[20] Special Joint Committee on Uniform Sentencing and Revision of the Criminal Law Statutes, First Interim Report (1978) and Second Interim Report (1979). The presumptive sentences for drug offenses proposed by these reports were not enacted; however, the legislature did finally enact mandatory minimum sentences for drug dealers in 1980.

[21] Senate No. 1108 (submitted during the 1979 session of the legislature) would have taken away the power of prosecutors not to prosecute certain firearms offenses covered by the Bartley-Fox one year mandatory minimum sentence. A number of prosecutors who were interviewed at that time indicated their outrage over the legislation; it too was not enacted.

[22] Rather than detail the House and Senate legislation numbers for each proposed mandatory minimum sentence, it will be easier to report the statutes that were amended to provide

mandatory minimum sentences. (This convention will be followed for the next two footnotes as well.) The sentences established for several sexual assault offenses (ALM GL C.265, ss. 13B, 22, and 24) did not mandate a specific minimum term; rather, offenders were required to serve, as a mandatory minimum, two-thirds of the minimum term imposed by the judge in any particular case. The sections of the General Laws listed above were amended by the legislature in July 1980.

[23] Legislation passed in July 1980 required that offenders convicted of a number of auto theft offenses (ALM GL C.266, ss. 27A, 28, and 37) serve one year mandatory minimum sentences. Also, a special Governor's Task Force on Auto Theft report (1980:50-54) urged the enactment of mandatory sentences for unauthorized use of a motor vehicle and motor vehicle theft, as well as mandatory restitution to insurance companies in cases involving fraud.

[24] Legislation passed in July 1980 required that offenders convicted of maintaining a house of prostitution or of deriving support from a member of a house of prostitution (ALM GL C.272, ss. 6 and 7) serve a two year mandatory minimum sentence.

[25] Special Joint Committee on Uniform Sentencing and Revision of the Criminal Law Statutes, Third Interim Report, also known as House Bill Number 6304 (January 1981). The legislation has not yet been acted upon.

[26] We have noted that the first two reports of the Special Joint Committee were not enacted and that the third report is still in the Massachusetts legislature.

[27] Massachusetts House, No. 6595 (1977).

[28] ALM GL C.94, ss. 32-32H as amended in July 1980.

[29] Opinion of Justices to House of Representatives, 393 N.E. 2d 313, (July 6, 1979).

[30] The pieces of legislation in question were Senate, No. 777; Senate, No. 813; Senate, No. 814; and House, No. 507 (1979).

[31] The Court was also of the opinion that the legislation did not appear to be unconstitutional because of the vagueness of the term "street value" that was used to determine the worth of the drug in question.

[32] Chief Justice Hennessey, in a concurring opinion to the case of Comm. v. Appleby (402 N.E. 2d 1062, (1980)), saw reason to comment on the sentencing process:

> In sum, the sentencing result here is one which focuses on the compelling need for reasoned application of the broad sentencing discretion ordinarily available to our trial court. (emphasis added)

The comment on sentencing discretion by Mr. Justice Hennessey deserves note here, especially since the major issues of the case did not involve that topic.

[33] Although a number of bills were debated during the 1979 legislative session, all of them -- including the drug offense legislation -- were passed during the 1980 session.

[34] Massachusetts House, No. 6304, January 12, 1981.

[35] The bill proposed the addition of Chapter 266A to the Annotated Laws of Massachusetts.

[36] House, No. 6304 states in its introduction that:

> Mandatory sentencing, it should be mentioned, was considered and rejected by this Committee. While a mandatory approach would provide the requisite candid statement of sanctions, it will neither eliminate unjustifiable sentencing disparities nor provide substantive sentencing justice. Experience has long shown that mandatories result in unenforced laws and unpunished crimes. Particularly in these days of eroding public confidence in government, sentencing laws can not be shams of illusionary stringency but must be functional reliable vehicles for the measured and appropriate response to crime. (pages 4 and 5)

[37] Apparently rough drafts of this legislation had been informally circulated for comment. We were not able to obtain copies of this legislation and, since we do not know for sure what it contains, we will not comment about it further here.

[38] The maximum penalty for each offense is stated within the statutory definition of the offense in the Annotated Laws of Massachusetts. To find the statutory definitions and penalties provided for most common offenses (excluding drug offenses) see generally the chapters cited below:

ALM GL C.264: Crimes against Government
ALM GL C.265: Crimes against the Person
ALM GL C.266: Crimes against Property
ALM GL C.267: Forgery and Crimes against the Currency
ALM GL C.268: Crimes against Public Justice
ALM GL C.269: Crimes against Public Peace

ALM GL C.270: Crimes against Public Health
ALM GL C.271: Crimes against Public Policy
ALM GL C.272: Crimes against Chastity, Morality, Decency and Order

[39] The specific offenses over which the District Court has concurrent jurisdiction with the Superior Court are detailed in Chapter 218, section 26 of the Massachusetts General Laws. At the time of the sentencing guidelines development, the offenses of concurrent jurisdiction, other than those that carried a state prison sentence of 5 years or less, were: breaking and entering in the nighttime, breaking or attempting to break a safe with intent to commit larceny, breaking and entering in the daytime with intent to commit a felony, possession of burglarious instruments, theft or unauthorized use of a motor vehicle, receiving a stolen motor vehicle, escape or attempt to escape from a penal institution, and forgery and the utterance of forged notes. (See Speer, 1979:2) A 1980 amendment to the General Laws also gave the District Court jurisdiction over the disposition of assault and battery offenses.

[40] This overview of the legal structure of the sentencing process in Massachusetts is intended to give the reader some familiarity with how cases are processed and disposed of; it does not intend to provide an analysis of all of the intricacies of the then (or now) existing law pertaining to criminal sentences. For a more detailed discussion of that law, see Shapiro, 1976:87-102.

[41] An offender could be charged, of course, with two offenses of differing seriousness -- a serious felony and a minor misdemeanor violation, for example. In such a case, both charges would be disposed of in the court (either Superior or District) which has jurisdiction over the more serious felony offense.

[42] Chapter 274, section 1, of the Annotated Laws of Massachusetts states: "A crime punishable by death or imprisonment in the state prison is a felony. All other crimes are misdemeanors."

[43] ALM GL C.218, s.30 gives the District Court power to decline to exercise its jurisdiction.

[44] Comm. v. Rice, 216 Mass. 480 (1914) and Corey v. Comm., 364 Mass. 137 (1973).

[45] Corey v. Comm., 364 Mass. 142 (1973). In addition, the District Court may not decline jurisdiction after it has held a hearing on the evidence (Comm. v. Clemmons, 370 Mass.

288 (1976)). Speer (1979:7-8) discusses a number of other procedural issues relevant to the bind-over decision as well.

[46] ALM GL C.218, s.27.

[47] ALM GL C.279, s.23 and s.18.

[48] The sentencing discretion of the Superior Court is limited, however, if the case is being tried in Superior Court on appeal from the District Court. In this instance, the Superior Court may only impose a sentence within the 2 1/2 year time boundary available to the District Court. Though this restriction does not mean that a sentence imposed by Superior Court cannot be more severe than the original sentence imposed by the District Court, one writer has noted that in practice "when a jail sentence is imposed, the case is usually appealed and frequently results in milder disposition from a Superior Court judge." (Flaschner, 1973:125). It should be noted however that the author, the Hon. Mr. Justice Flaschner, does not provide any supporting evidence for this belief.

[49] Sentencing disparity may, indeed, result from the Massachusetts dual court jurisdiction over certain types of offenses, but the disparities are not easily noticeable and seem to have gotten lost "between the cracks" in that state's criminal justice system. To our knowledge, no study has ever been done on the differential effects of court level of offense disposition on the length and type of sentence for those offenses subject to concurrent jurisdiction in the District and Superior Courts. Also, the Massachusetts sentencing guidelines project did not address this point at all in their analyses of sentencing practices.

[50] It should be reiterated here that the cases we refer to in our hypothetical example are similar in all respects: offense, circumstances, offender background, etc. It may be the case that sentences in the District Courts are lighter for concurrent jurisdiction offenses than they are in the Superior Court, but those differences may be justified by differences in the circumstances or the offense on the offender's background and criminal history. However, since no studies have been done to date which could verify or discredit some of these questions, our hypothetical example could also be the case.

[51] Our interviews with prosecutors throughout Massachusetts revealed that both of these strategies are fairly common. In one county, certain high-volume offenses, as well as habitual offenders, are screened by district attorneys at the police station and district court level as

part of a major-offense project. Cases identified by this project are brought directly to the grand jury for indictment. See the final section of this chapter (and notes 73 forward below) for a complete description of our interview sample and questionnaire procedures.

[52] Comm. v. Moran, 353 Mass. 166, (1967).

[53] We will discuss the trend in Massachusetts toward the enactment of mandatory minimum sentences -- a trend that has not abated -- in the next section. However, in 1978 the only mandatory minimum sentences were required for violations of weapons offense statutes. See ALM GL C.269, s.10.

Sentences of imprisonment may not be suspended for several offenses (assault and battery to collect a loan, armed burglary, and gun possession without a license) and this limitation on the judge's discretion does in practice result in a sort of mandatory minimum sentence for these offenses. See ALM GL C.265, s.13C, and s.14; C.271, s.10; and C.269, s.10. Also, see the last section of this chapter and notes 83 forward below.

[54] ALM GL C.127. Various sections of this chapter outline the mandatory minimum periods of imprisonment in correctional facilities. Section 133 of ALM GL C.127 provides that offenders not convicted of certain specific crimes or of crimes committed while on probation or parole shall serve "one third of such minimum sentence, but in any event not less than 1 year." This stipulation was later changed during the 1980 legislative session as will be recounted in the next section.

[55] ALM GL C.278, s.28A.

[56] ALM GL C.278, s.28A.

[57] ALM GL C.278, s.28A. In fact, as Shapiro, 1976:98 notes:

> Insofar as these limitations prohibit women from appealing sentences of between 2 1/2 and 5 years, while allowing men to appeal similar sentences of that length, there appears to be a violation of the Equal Protection Clause of the Fourteenth Amendment.

[58] ALM GL C.278, s.28B. The staff of the Evaluation of Statewide Sentencing Guidelines project was fortunate in that the judges of the Appellate Division allowed us to attend one of their sessions and observe the decision-making process of that body. We are most grateful to them for extending the invitation.

[59] Walsh v. Commonwealth, 260 N.E.2d 911 (1970).

[60] ALM GL C.278, s.28A, s.28B, s.28C.

[61] Ziesel and Diamond (1977) noted several cases in which the Appellate Division did increase the original sentence on appeal. Hrones has noted as well that:

> ...Very few sentence appeals succeed unless the trial judge was way out of line on his sentence or there are very special factors concerning the defendant that were not brought to the attention of the sentencing judge.
>
> ...Thus, if an experienced criminal lawyer considers the sentence as not unusual under all the circumstances, one should strongly advise his client not to risk his sentence being increased where his chances of getting it reduced are very slim. (p. 6)

[62] The Parole Board has set this initial parole eligibility time requirement for House of Correction sentences under its jurisdiction. See, Massachusetts Parole Board and Advisory Board of Pardons, Decision-making Guidelines and Procedures for Parole Granting, Parole Revocation and Pardons (1978:3).

Of course, inmates sentenced to Massachusetts correctional facilities are also allowed "good time" deductions to their sentences by statute; however, judges did not indicate to our interviewers that they made any attempt to calculate the effect of good-time deductions on the sentence imposed since such deductions would be variable depending on the conduct of the offender while incarcerated. For a description of how deductions for good conduct are calculated, the reader is advised to see ALM GL C.127, s.129.

[63] ALM GL C.127, s.133.

[64] ALM GL C.127, s.133. In 1978, the Massachusetts Parole Board has paroling jurisdiction over all offenders confined in the state prison system, county jails and houses of correction except offenders sentenced by the District Court to terms of less than a year. However, during the 1980 legislative session, the minimum period of imprisonment, for determining who should be eligible for parole, was reduced from 1 year to 60 days. ALM GL C.127, ss.128, 130, 131, 148, 149, as amended, thus consolidated paroling authority over all offenders in the Parole Board. This amendment, however, does not affect the 1 year minimum sentence; rather, it merely broadens the jurisdiction of the paroling power of the Board to include authority over offenders serving sentences of less than 1 year.

[65] ALM GL C.279, s.23.

[66] Massachusetts Parole Board and Advisory Board of Pardons, Decision-Making Guidelines and Procedures for Parole Granting, Parole Revocation and Pardons, 1978:2.

[67] ALM GL C.127, s.133 details the offenses for which two-thirds of the sentence must be served before parole eligibility. This section of the chapter also contains a clause that allows the parole board to advance the eligibility for an offender convicted of one of the stated 'two-thirds offenses' "upon the written recommendation of the superintendent or director of the prison camp, and the commissioner of correction, and, with the consent and approval of a majority of the parole board." If such a request for advance eligibility is made and approved, the offender would still be required to serve one-third, or at least one year, of the sentence before parole. See ALM GL C.127, s.133. Offenders convicted of 'two-thirds offenses' committed while on parole may not, under the provisions of this section, have their parole eligibility date advanced.

[68] The term "persons offense" as used to determine parole eligibility should not be confused with the general classification by Massachusetts statute of "Crimes against the Person." While most of the Crimes against the Person are also "persons offenses" receiving two-thirds parole eligibility, a number of such crimes against the person are not considered "persons offenses" and are eligible for one-third parole eligibility. Examples include such offenses as Libel and Defraud/Willful Misapplication which are listed by statute as Crimes Against the Person, but which are not violent "persons offenses" for the purpose of computing parole eligibility at two-thirds of the minimum sentence.

[69] The following offenses, recognized as not only potentially life-threatening crimes, but also as offenses often intentionally committed to obtain insurance funds while jeopardizing life, were added to the list in 1979: Burning a dwelling (ALM GL C.266, s.1); Burning of Meeting House (ALM GL C.266, s.2) and Burning, Insured Property (ALM GL C.266, s.10).

[70] ALM GL C.265, s.2; and ALM GL C.127, s.133A.

[71] ALM GL C.127, s.133A.

[72] ALM GL C.127, s.133.

[73] Since Massachusetts judges are assigned to county courts on a rotating basis, they are usually familiar with

regional differences in types of crimes and offenders throughout the state. We, however, were not. Thus, in selecting four counties as sites for interviewing court personnel, we attempted to capture as much of the state's diversity on the following characteristics as we possibly could: crime rates, population size, population ethnicity, per capita income, regional location, urban/rural nature, county geographic size, and staff size. The last item considered was the number of cases that each county had contributed to the Massachusetts sentencing guidelines project's data base. On the basis of these characteristics, the four counties chosen as interview sites were: Hampden, Plymouth, Suffolk, and Worcester.

[74] We requested permission in each county to interview the entire staff of the district attorney's office, the Massachusetts Defenders Committee, and the Superior Court Probation Department. With one exception, that request was granted, and we were able to speak with the majority of personnel in each office; personnel who were not interviewed were either on vacation or were tied up in court matters that they could not put off.

[75] The interviews that we conducted with staff in each of our site counties could not have contained detailed questions because the guidelines had not yet been completed. We decided to limit the discussion to several topic areas (which are listed below) and to urge the respondents to contribute any additional thoughts about the sentencing or paroling process.

[76] We interviewed all of the Superior Court judges that were available in each of these counties, as well as a number of judges from adjacent counties. In interviewing the judiciary, our real aim was to speak with as many of the judges as possible -- rather than to limit our sample to just those judges then located in our county sites -- because the judicial reception of sentencing guidelines would, by and large, determine the use and effect of the guidelines.

[77] No real attempt was made to interview District Court judges who were not assigned to Superior Court since that group would not be using the sentencing guidelines.

[78] We mentioned in passing above that there was one exception to the approval of our request to interview all members of a particular office's staff. That exception was the Worcester County District Attorney's office where we were only permitted to speak with the District Attorney himself, and one of the First Assistant District Attorneys. The reasons why we were not given access to the remaining staff

are still not clear to us. (During our second phase interviews in 1980, however, we were able to speak with several of the assistant district attorneys from Worcester County, as well as the District Attorney himself.)

[79] These topic areas were discussed with all of the judges and personnel interviewed; additional topic areas were included, however, for prosecution and defense counsel (and for probation staff when applicable) that pertained to sentencing recommendations and how such recommendations were derived. Probation staff was asked, as well, about supervision requirements, court case load, and pre-sentence investigations.

[80] The discussion, in this section, of factors that may influence judicial sentence decisions, is based, we would like to repeat, on data supplied to us by judges, prosecutors, defense counsel, probation staff, and others involved in sentencing in Massachusetts. We are restating, for the reader's benefit, only what these actors told us happened in the sentencing process in Massachusetts, and these assertions as to how sentencing actually occurred are not the result of data analysis; this, however, does not mean that data analysis conducted in such a manner as to capture these items of information thought to be relevant to sentencing would not support the judges' and others' contentions about how sentences were derived. It may well be the case that judges and others can perfectly well account for how their decisions are made, and a sentencing study that has some notion of what to look for may find empirical evidence to support these beliefs. (As we will discuss in the next chapter, however, the Massachusetts project -- partially constrained by the nature of the data that they were able to collect empirically -- was not able to simulate this description of judicial decision-making.)

[81] The word "caused" here is obviously not used in the social science research sense -- judges merely want to know what happened and what triggered it.

[82] Of course, police initiate the majority of complaints and a good number of the final charges issued; however, police in any given jurisdiction work closely with the prosecution and, thus, know what are considered "acceptable" charging practices by that office.

[83] See the preceeding section and note 53 above for a listing of which offenses those were in 1978. There are also requirements for certain additional offenses in the statutes that sentence man not be suspended for second or subsequent convictions of a broad range of offenses.

[84] With the exception of the statutory mandatory jail sentence for the possession of an unlicensed firearm, the other offenses are sufficiently severe that the statutes call for state prison incarceration if incarceration is the judges sentence.

[85] Massachusetts judges stressed the point to us during interviews that they were more than capable of telling exactly "what type of case the prosecutor had" from a preliminary review of the evidence. The importance of this criterion should not, in our opinion, be underestimated as a good number of the judiciary interviewed who espoused this belief were, at one time or another, either prosecutors or defense attorneys required to deal with this dimension of the evidentiary process on a daily basis.

[86] This is, of course, a summary of what judges told us during our interviews; whether or not we believe that they do not consider these system operation factors as the real reasons for mitigating a guilty plea disposition, is another matter.

[87] The order of judicial (and other actors') processing of certain types of information was an area that we had also proposed to study in our never-funded phase II of this project. (See Chapter 1 for a description of that proposal.) Given the lack of a comprehensive present state of knowledge on the topic of judicial information processing, we can only make assumptions, and we acknowledge that they are of doubtful validity, about how judges' decisions are really made.

[88] The distinction often made by judges, prosecutors and defense counsel, and probation officers in particular about prior succesfully completed probation terms when considering appropriate sentences (obstensibly including the possibility of another probation term) for current offenses appears to us to be a ludicrous one. After all, if an offender who has "successfully completed" a prior probation term is now being sentenced for a subsequent offense, obviously the previous probation term was not much of a success.

[89] This sort of procedure is considered a good tactic, as far as we could discern from our interviews, to be used on offenders who have been racking up records of petty offenses, have been receiving light dispositions in the past, and have an attitude characterizing that of a "smart alec" toward the power of the criminal sanction.

[90] Judges are not alone in reversing the order of sentences from that of charges and thinking that this

procedure is appropriate (although it does appear to us that the legislature could perhaps find considerable problems with the procedure -- since it after all does imply that the sentences proscribed by that body for certain offenses are so inappropriate that they should be circumvented); the point earlier was that district attorneys tailor the charge package for that reason as well.

[91] The same answers about the method used to set a sentence were mentioned too often and have stages so idiosyncratic to the Massachusetts system to have been mere rote responses to questions about how judges sentence. None of the judges gave us the popular "bifurcated decision-making process language", we should note.

[92] Actual time might not, statutorily, match to the institutional choice. This would be the result should a judge impose a state prison sentence on a very young offender (as we observed in one case) and later inform the Department of Corrections that the sentence should be served at Concord (an institution used for younger offenders and for less serious offenses). The reason that a judge would do something like this is because he or she may want the offender's sentence to be subject to the parole rules for the state prison, rather than those for Concord, where times served are usually much shorter.

[93] Except for the disposition of "filed awaiting apprehension", judges did not appear to have preference for any particular "filed" disposition. One aspect of "filed" charges and dispositions that all persons interviewed agreed to was that once a charge has been filed -- either after a plea of guilty or not -- the charge is never "unfiled" even though the prosecutor has the power to do so.

[94] In the discussion to follow, we will note in detail that judges believed that sentence dispositions were less influenced by recommendations after trial cases than plea cases; it is important to remember during the discussion that follows however that most cases are disposed of by plea.

[95] Of course, even when a respondent said that he or she would support sentencing guidelines in general, it was noted that ultimate support of the Massachusetts Superior Court guidelines would depend on exactly what factors were in the guidelines and how the guidelines operated.

[96] The distinction here is between the study of past sentences and the development of guidelines from past sentences. Judges did not mind the study of past sentences at all.

[97] Ironically, as we will see more clearly in the next chapter, it is precisely the average case -- or plea disposition case -- that the guidelines do not apply to.

[98] Unlike judges, both prosecutors and defense counsel believed that tremendous disparity existed in the Massachusetts sentencing system. It could be the case that sentences that prosecution would cite as examples of disparity would be the defense's example of an appropriate exercise of judicial discretion, and vice versa.

Chapter 7: Developing and Implementing the Massachusetts Sentencing Guidelines

The Massachusetts Superior Court sentencing guidelines project began in June 1978 to study the sentencing practices of judges throughout Massachusetts with a view to the possible development of statewide sentencing guidelines. The study was funded by the National Institute of Justice through the state's planning agency -- the Massachusetts Committee on Criminal Justice. The Committee on Probation and Parole of the Superior Court -- a group of seven judges, headed by the Honorable John T. Ronan -- was delegated the responsibilities of overseeing the activities and direction of the project and of reporting the project's findings to the Court.[1]

The primary objective of the project (as stated in its funding proposal) was "to ascertain the extent to which there exists sentencing disparity in the Superior Court." Further, the proposal stated the intent of the project "to construct an empirically derived sentencing model that will provide a recommended, nonbinding sentence range for similarly situated offenders ... by using as a base past Superior Court sentencing averages". The project was to have three phases. Phase 1 would involve a statistical analysis of then current sentencing practices; Phase 2 would deal with the selection of a sentencing guidelines model by the project research staff and the Committee on Probation and Parole; and during Phase 3, the judges of the Superior Court would test the model throughout the state. The press reports released during the first few months of the project's operation stressed that appropriate sentencing guidelines should be based primarily on the statistical findings of the study, yet also noted that the Superior Court might decide that certain variables "should not be reflected in the guidelines".[2]

Project activities during the first five months mainly involved sample selection and development of data collection instruments. Training of data collectors and data collection itself were conducted during the following five months, ending in June 1979. Statistical analysis of the data then began, and this stage of the project ended in November 1979, when a preliminary statewide sentencing guidelines model was distributed to the judges of the Superior Court for use on a "test basis." The testing period of the preliminary guidelines model lasted from sometime in December 1979 until early in 1980.[3] After some revisions, the final guidelines model developed by the Massachusetts Superior Court sentencing guidelines project was presented to the Superior Court judges for further testing on 18 April 1980. At that session, the judges were also trained in the use of the guidelines for

particular case examples -- though, as we make clear below, this "training" was not all it could have been, and it did not compare with the familiarization and training done in Minnesota (discussed in Chapter 9 of this report).

The guidelines project then held several training sessions during May 1980 to familiarize the Superior Court probation officers with the day-to-day operation of the guidelines and with completion of various sections of the guidelines forms. Following the initial training of probation staff, however, the Massachusetts sentencing guidelines project was terminated due to lack of funds (at the end of May 1980). The task of monitoring the continued testing of the sentencing guidelines was then delegated to the Superior Court Probation Department -- in particular, to the Assistant Supervisor of the Superior Court Probation Services -- through the Superior Court Committee on Probation and Parole.

The testing of the final guidelines model developed by the sentencing guidelines project lasted from May 1980 to May 1981. At the semi-annual Superior Court judicial conference in May 1981, the justices voted to adopt that model without revision. It was decided that:

> Commencing June 1, 1981. After conviction by jury verdict or upon findings of guilty by a trial judge, a sentencing judge, SHALL consult the sentencing guidelines structure; shall complete the form; make whatever disposition as is deemed appropriate; and thereafter return the form to the probation department.

The legal structure of the Massachusetts sentencing process and the informal sentencing policies of judges, prosecutors, defense counsel and others in the state's criminal justice system were discussed in Chapter 6 of this report. Also in that chapter we attempted to shed light on the social and political forces that led to the judges' decision to develop statewide sentencing guidelines. This chapter will examine how those guidelines were actually constructed by the Massachusetts sentencing guidelines project. The structure of the adopted sentencing guidelines will be then looked at in some detail; and the final section of this chapter will analyze the initial reactions of judges, legislators and criminal justice personnel to the content and use of the Massachusetts statewide sentencing guidelines.

Development of the Massachusetts Sentencing Guidelines

The Massachusetts sentencing guidelines were constructed after analyses of data about 1,440 defendants convicted in the Massachusetts Superior Court between 1 November 1977 and 31

October 1978. Since the final sentencing guidelines are not solely the result of data analysis but, rather, include policy modifications made by the Superior Court Committee on Probation and Parole, our commentaries (which follow) on the methods of data definition, collection, and analysis might be considered moot issues by some.[4] The general methods used by the Massachusetts project, however, in large part replicate the procedures followed by a number of other guidelines projects to date; one purpose, therefore, of the commentaries will be to illustrate, with a concrete example, some of the basic problems inherent in this traditional guidelines development methodology -- especially since it has been followed by most of the recent projects that have been concerned with developing sentencing guidelines.

That the research findings of the Massachusetts sentencing guidelines project were modified by policy considerations, does not preclude the possibility that those guidelines were also flawed by methodological problems stemming from that project's empirical research and preliminary statistical analysis of sentencing decisions. It may be the case, for example, that imperfect statistical representations of sentencing policy might be made even less precise by policy modifications. There are (due to methodological error) a number of ways in which this could occur and, a second purpose of this methodological critique will be to point out some of those possibilities, especially as seen in the Massachusetts sentencing guidelines.

Defining and Collecting the Data

We have already mentioned that the Massachusetts sentencing guidelines project collected information about 1,440 sentenced defendants. Exactly what information was obtained? How was it decided that that information would be what was relevant to sentence decisions? How were the data collected? How was the sample of defendants chosen? What procedures were followed to ensure that the final sample of cases on which data were collected was a valid sample, yielding valid data? Unfortunately, these basic questions of research design usually must be decided at the onset of almost any study -- at a time when researchers new to the study of a particular topic do not really know the idiosyncracies of the area under study, or how to compensate for them. In fact, one of the most important criticisms that this evaluation can make of the Massachusetts sentencing guidelines is that the guidelines project did not collect the right information -- in other words, the researchers did not collect the information most relevant to past judicial sentencing decisions in Massachusetts.[5] Though we will indicate the specific problems in a moment, it should be said (in fairness to the

Massachusetts sentencing guidelines project) that the problems with the sentencing guidelines which derive from the data are fairly common problems encountered in many research projects that depend heavily on retrospectively collected data.

The Massachusetts Superior Court Probation Department, in conjunction with the county court clerk's office, maintains a daily record of every case disposed of in the Superior Court. One reason that the Probation Department maintained such an extensive listing of cases when the Massachusetts guidelines research began (for, obviously, not all defendants receive probation as a disposition) was because the department was responsible for the preparation of a "pre-trial" investigation report for each Superior Court case. Among other things, the report contained (for the judges' use in setting bail, etc.) background information about each defendant and the offender's prior record. (Subsequently, as we shall see, the courts decided to use pre-sentence reports instead of pre-trial ones.)[6]

The Massachusetts sentencing guidelines project decided to use the records of the Superior Court Probation Department as the data base for the selection of a sample of cases because of these two attributes: the records represented the entire population of cases heard in Superior Court, and the records contained the most consistent background and prior offense information available in one place.[7] The project then decided that a random[8] sample of approximately one-third of the yearly population of dispositions[9] would be adequate for the purpose of examining sentencing practice. A decision was made during the planning stage of the research to exclude three counties because of small caseloads -- Dukes, Franklin, and Nantucket; later on during the data collection phase of the project, Barnstable county would also be excluded on the basis of its caseload and the project's time and funding constraints.[10]

The development of a coding manual, which detailed the exact information to be collected about each defendant in the sample, was the next task of the Massachusetts sentencing guidelines project. The project tested a preliminary coding manual on a statewide sample of 50 cases[11] with the stated aim of "finding out what information is available in probation case folders" since the records of the probation department had been chosen as the major source of information for the data base. The final coding manual, including revisions thought necessary after an analysis of the 50 pre-test cases, was approved by the Committee on Probation and Parole on 2 February 1979 and the collection of data by approximately 30 trained[12] data collectors began shortly thereafter.

We mentioned earlier that a major criticism of the Massachusetts sentencing project's guidelines evolves from the fact that some information especially pertinent to judges' sentencing decisions was not collected by that project. Observations of actual sentencing hearings by staff of this evaluation project, as well as our numerous interviews with judges, prosecutors, and defense counsel about the sentencing process, have confirmed the generally held belief that some of the most important determinants of sentence dispositions are the sentence recommendations of prosecutors and defense counsel -- and the extent to which those two sets of recommendations agree. While the Massachusetts sentencing project did include items relating to prosecutors' sentence recommendations in their data coding manuals,[13] much of this information was not uniformly collected.[14] Finally, the sentence recommendations of defense counsel were not collected at all.[15] There are several reasons why this information was not collected for every case. The major reasons that the prosecutor's recommendations were often missing for sample cases are 1) that this sort of information was not present in probation department files and 2) this information often was not recorded in the district attorney's records.[16] More importantly, staff of the Massachusetts sentencing guidelines project did not (at least during the earliest design stages of the project) observe sentencing dispositions themselves, or interview judges or other court personnel;[17] thus, they were unaware, until it was too late, of the actual importance of both prosecutors' and defense counsel's recommendations. Whatever the reasons, the end result was that what is arguably some of the most important information predictive of sentence dispositions was excluded from the research that led to the sentencing guidelines model.

While the structure of the data collection instrument was being finalized, several sentencing guidelines project staff collected information on the most serious offense, the court of sentence disposition, and the judge at sentence disposition, for all of the 1,500 cases initially selected for inclusion in the sample. The purpose of this separate exercise was to insure that the sample was representative of sentence dispositions, and judges imposing these dispositions, throughout the state on these three characteristics.

One possible reason that the project wished to check for the representativeness of judges in the Massachusetts sample involves the method for assigning judges to counties and to particular court sessions in Massachusetts, since the judges in that state typically rotate between counties on a "circuit" basis. However, during our interviews with judges and others in the state, we were told that the criteria used to make these assignment decisions are somewhat informal. Apparently,

although there is a definite attempt to try to assign each judge to either his or her home county or a county close to it, circuit assignments are further influenced by the seasons and the weather, access to highways to get from home to the circuit court during months of bad weather, and seniority. Further, the last judge assigned to each circuit is usually the judge who does the most traveling within the circuit.

One judge in each area's court serves as the "assignment judge" and, as such, presides over the "first session" -- which hears motions, takes pleas of guilty, sets trial dates, etc. The remaining sessions (the numbers of which vary by the size of the county) are trial sessions. Assignment judges are chosen using the following informal criteria: whether a judge wishes to sit as the assignment judge, and the legal background of the judge. We were also told informally, during our interviews, that the choice of assignment judge also depends on the ability of the judge to "get guilty pleas without giving away the courthouse."

The Massachusetts project, aside from detailing the nine most frequently appearing offenses,[18] did not report any further findings from their review of the preliminary data. As the number of cases quoted differs, however, from the final number of cases actually used for the development of the Massachusetts sentencing guidelines, a discussion of the selection of that sample, and the reasons for the deletion of cases from the sample, appears to be in order.

A clear policy decision was made by the staff of the Massachusetts sentencing guidelines project to exclude certain cases from the original sample where the majority of information to be collected from the primary data source would be missing. A memorandum sent to all data collectors during February, 1979[19] indicated that the major reason for disqualifying a case from the sample would be that the pre-trial investigation report was either missing or had not been updated. A later memorandum on the same topic specified that the pre-trial investigation report must be less than one year old for the case to be included in the sample.[20] Although lack of a fairly recent pre-trial investigation report was cited as the major reason for excluding a case from the sample, a number of other reasons would also be acceptable for disqualification; however, these additional reasons for omission were never specified by any of the project memoranda or analysis reports.[21] In any event, each case that was disqualified (for whatever reason) was to be replaced with another randomly selected case from the same county.[22]

Table 7.1 details the outcome of the case selection process for each county in the Massachusetts sentencing

guidelines project's sample. The information contained in this table was taken from an appendix to a project memorandum[23] and is only partially explanable. The first row of Table 7.1 indicates the number of cases in each county originally designated for inclusion in the Massachusetts disposition sample. The next three rows indicate the numbers of cases that were dropped from the original sample. These cases were excluded because: the district attorney's case file was missing; the probation file (containing the pre-trial investigation report) was missing, or the pre-trial report that was present was more than one year old; or for an additional unspecified reason. In the fourth row of the table, we have provided the percentage of cases that -- based on the figures presented in the preceding three rows -- had to be dropped from the original sample. The next two remaining rows of the table contain numbers of cases included in the final sample after two resampling procedures, and the last row contains the final case sample size for each county.[24]

Insert Table 7.1 here

We consider the rows of the table that specify the number of resampled cases to be incomplete. Unfortunately, the Massachusetts project only listed statistics for those resampled cases that made it into the final case sample; the number of cases that had to be dropped (for any of the reasons listed earlier) from the two resamples was not also provided. Thus, it is not possible to calculate the overall percentage of cases that had to be excluded from the Massachusetts sample, since we do not know the overall number of cases sampled during the resampling procedures. We can only discuss the percentage of cases that had to be dropped from the original Massachusetts sample, before any case resampling was done.

Quite a bit can be said about the Massachusetts sample, however, even when the discussion is limited to a review of the number of cases eliminated from the original sample, rather than from the final sample. For starters, almost one-fourth (about 24 percent) of the original sample cases had to be dropped from the final sample. Of those 348 cases, the primary reason for exclusion was that the probation case file was missing, or because the pre-trial investigation report present in a probation case file was out of date.

It is not surprising that these particular reasons are the most common since the Massachusetts project had, after all, stated that missing probation files or incomplete reports would disqualify a case. But while we might expect that

Table 7.1: Massachusetts Sentencing Guidelines Sample Selection*

	Berkshire	Bristol	Essex	Hampden	Hampshire	Middlesex	Norfolk	Plymouth	Suffolk	Worcester	Total
Original Sample Size	30	86	125	190	22	303	117	81	346	181	1481
Cases Dropped-D.A. file Missing	0	2	5	1	0	8	6	0	6	3	31 (2.1%)
Cases Dropped-Probation file missing or outdated	0	5	15	13	0	73	2	3	95	5	211 (14.2%)
Cases Dropped-Misc. Reasons	0	0	14	16	1	44	0	10	16	5	106 (7.2%)
Percent Cases Dropped from original sample	0	8.1	27.2	15.8	4.5	41.2	6.8	16.0	33.8	7.2	23.5
1st Resample N	0	7	22	27	0	68	8	13	87	9	241
2nd Resample N	0	2	9	0	0	21	0	0	34	1	67
Final Sample Size	30	88	122	187	21	267**	117	81	350**	178	1441

*This table contains figures obtained from an appendix to a Massachusetts sentencing guidelines project memorandum dated January 29, 1980. Mathematical errors in the original version have been corrected here.

**The final Massachusetts sample of cases received by the Evaluation of Statewide Sentencing Guidelines project contained 1,440 cases. The discrepancies appear in Middlesex County (N=268) and Suffolk County (N=348)

missing information -- of one sort or another -- would be the major justification for excluding any case, the extent of information missing from the Massachusetts sentencing guidelines primary data source is somewhat alarming. Though it is not the only reason for alarm,[25] the general lack of information in Massachusetts probation records, might suggest that perhaps those records were not the best choice as a primary data source.[26] While 60 percent of the cases were excluded from the sample for lack of probation record information, only 9 percent were excluded because information was missing from district attorney files. Although we cannot say whether more cases would have been eliminated for lack of information in the district attorney files had those files been used as the primary, rather than secondary source,[27] it seems to us that -- given the apparent inadequacies of probation records, the district attorney files, as well as alternative data sources, should have been explored.[28]

In any case, it is one thing to have a sample which contains some cases on which some information is missing; it is quite another to have a sample from which such cases are completely excluded. The former situation produces some tough analytical problems; the latter may introduce sample bias which ex hypothesi cannot be estimated at all.

Clearly, from even a cursory review of Table 7.1, there is wide variation in the county rates for disqualified cases from the original sample and the variation does not appear -- with two exceptions -- to be systematic in any readily identifiable sense of that term. The exceptions, where one can identify some apparently systematic trend in county rates for excluded cases, are the two counties with the smallest, and the two counties with the largest, original sample sizes. Since the original sample was drawn randomly on a statewide basis, each county's original sample size should be proportionate to that county's Superior Court caseload. We might hypothesize that, in counties with small court caseloads, records might be kept in a more consistent fashion and thus that fewer cases would have to be excluded from the final sample for reasons of missing case files, information, etc. On the other hand, in those counties where the court's caseload is very large, a much higher proportion of cases might have to be excluded from the sample because of incomplete records. The Massachusetts project staff in fact noted in a 29 January 1980 memorandum on the quality of probation department records that "the larger counties with the longest backlogs and largest staffs have the most serious problems with pre-trial reports", thus indicating that the staff believed that missing or not-updated information was a function of county and court caseload size. Of course, if this hypothesis were to be correct, one would expect that the

percentage of cases dropped from the original sample would be a linear (or close to a linear) function of the county's original sample size; this does not appear to be true from the data in Table 7.1. There does not appear to be any other discernible reason for the extreme variations in county case exclusion rates -- with the exception of the two smallest and the two largest counties as we noted above.

One clear result of the varying county case exclusion rates was that the number of cases originally selected for each county differed somewhat from the county N's for the final sample. (Apparently, after the second resampling procedure, for reasons of time, etc., the Massachusetts project did not sample yet a third time to make the final county case N's equal to the original county case N's.) Once again, the differences are not proportionate to the original case size in any of the counties. Aside from three counties where the final sample N was the same as the original N, the differences between the original and final N's were small in magnitude with the exception of Middlesex County. Middlesex County was under-represented in the final sample by 36 cases, compared to that county's original sample size. However, the final N in that county is still quite substantial. Table 7.2, which shows the weight of each county's cases in both the original sample and in the final sample, confirms our belief that the relative representation of counties in the Massachusetts sentencing guidelines final sample was probably not so different from that of the original sample for it to matter much. However, the unclearly described (and in some respects curious) sampling procedures which were used leave us uncertain what kind of a sample they wound up with. It is one thing to have enough cases per county; it is quite another to have an unbiased sample containing enough cases per county.

Insert Table 7.2 here

The actual process of collecting information about each of the defendants in the final sample began in February 1979 and continued until June of that year. Data collection took longer than was initially expected and the staff of the project attributed the delay to difficulties in matching probation case files to the corresponding files in the district attorneys' and county clerks' offices.[29]

After all of the necessary information was collected and keypunched for each of the finally selected sample cases, the project staff reviewed the data to determine which items of information should be excluded from further analyses. As was the case with the decision to delete a selected case from the

Table 7.2: County Case Representation in Original and Final Massachusetts Sentencing Guidelines Sample

County	Original Sample N	% of Total	Final Sample N	% of Total
Berkshire	30	2.0	30	2.1
Bristol	86	5.8	88	6.1
Essex	125	8.4	122	8.5
Hampden	190	12.8	187	13.0
Hampshire	22	1.5	21	1.5
Middlesex	303	[illegible]	268	18.6
Norfolk	117	7.9	117	8.1
Plymouth	81	5.5	81	5.6
Suffolk	346	23.4	348	24.2
Worcester	181	12.2	178	12.4
Total	1481	100	1440	100

original sample, the primary reason that certain items of information were excluded from further analyses was due to the extent of missing information. Project staff decided not to conduct further data analysis on any item where the rate of missing information was higher than 17 to 20 percent.[30] Several additional reasons were stated by the project staff for not analyzing some items of information collected. Among those reasons were, that the information was not originally collected for the purpose of developing sentencing guidelines and that the information captured by one variable was "duplicitious" of information contained in another variable also present in the data.[31]

What effect did these three rules have on the amount of information used in the sentencing guidelines research conducted by the Massachusetts project? The first rule governing the exclusion of variables from the sentencing guidelines research -- that variables not be analyzed if more than 17 percent of the cases evidenced missing information as a response -- had, by far, the greatest effect on the subsequent research. A substantial number of variables (about 28) in the Massachusetts sample were not analyzed at all because of this rule.[32] For some of these variables, the information requested was missing for most of the cases, while the reliability of the information present in the remaining cases was thought by project staff to be generally unreliable.[33]

Although the project noted the two additional reasons for variable exclusion as being that the information was not gathered for guidelines development purposes and that the information duplicated that obtained in other variables, these reasons were not often mentioned as reasons for the exclusion of particular variables from analysis. In fact, only two variables (whether the judge at conviction was assigned to the Superior or District Court, and whether the defendant had applied for appellate review subsequent to the disposition) were excluded as not pertinent to the development of guidelines, and only one variable (weapon use) was considered "duplicitious" since it would be analyzed later in relation to degree of injury.[34] Rather, a different set of reasons emerged when the reasons for excluding specific variables were examined. Variables relating to a specific type of offense were often excluded (about 8 in number) because the sample size for specific offenses was rather small;[35] and an additional 7 or 8 variables were excluded because, for those responses obtained, the cases evidenced overwhelming similarity -- that is, they displayed very little variance.[36]

The reasons given by the Massachusetts sentencing guidelines project for not analyzing the various information items listed above may be questioned. One might ask, for example, how the project determined that bail information in particular was generally unreliable when, in fact, the accuracy of information about all of the variables had not been verified? While questions such as this are important in their own right, they are trivial compared to the more important issue here of how the project decided to handle missing data in their analyses. It is a generally understood and reluctantly accepted fact known throughout the field that criminal justice data are usually either dirty (filled with error and bias) or missing or both.[37] Yet these problems did not just come about yesterday and a number of approaches have been developed to deal with them.

The first approach, and probably most common, to dealing with a problem of extensive missing data has been to throw the data out. This is the solution that was chosen by the Massachusetts sentencing guidelines project.[38] The rationale behind this solution is that, since most of the time the information was missing for the judge, as well as for the researchers, it must not have been of utmost importance. It is possible that this rationale is correct. However, there may be another way to look at evidence of extensive missing data -- from the reverse of the coin. The reverse view would not say that the data are missing for a majority of cases but, rather, present for a minority of cases. From such a perspective it becomes rational to assume that, when information normally missing, is present, it is present precisely because of its relevance to the particular case and its handling.

To summarize this commentary on the methods of defining and collecting the data to be used in the Massachusetts sentencing guidelines study, we emphasize the following points. First, it is apparent that the primary data source selected by the project was inadequate to sufficiently inform the staff of the sentencing guidelines project about what actually occurs in the sentencing process, despite rather sophisticated analyses (which we will address in a moment). The most important reason why that data source -- probation department records -- was inadequate for the task at hand was that much of the information truly relevant to sentencing was never recorded there. Secondly, the procedures used to select the sample of cases -- notably the first and second resampling procedures -- were not adhered to in any uniform fashion (for example, in one county all of the cases excluded from the original sample might be replaced, while, in a second county, only half of the deleted original cases might have been replaced) and the possibility of bias resulting from these

resampling procedures was not investigated. Finally, the definition of what information was to be collected apparently was governed not by the sentencing practices of the court, but rather, by an assessment of what was available in the major data source, and -- perhaps more so -- by what items had been collected by sentencing guidelines projects in the past. With the exception of a few items of information idiosyncratic to the Massachusetts sentencing structure -- such as items related to disposition and to appellate review -- much of the information required by the Massachusetts codebook could have been taken first hand from one of the then popular manuals on how to construct guidelines.

Probable by-products of this method of determining which items of information (that might or might not be pertinent to sentencing) to collect are that the collected information will contain a high proportion of missing data (after all why write down information that isn't really relevant?) or that the information that is collected will have little or no bearing on sentence disposition. Unfortunately, both of these possibilities appear to be true in the case of the Massachusetts data. A high proportion of missing data was evident, and those items of data that were available, by and large, were not particularly relevant to the sentencing process in Massachusetts. Of course, this method of analyzing variables that evidence a high proportion of missing data is not free from error -- indeed, it may well introduce additional distortion of various sorts into data analyses depending on how the technique is used. But since a large number of variables which had a great deal of missing data would have been thrown out of the analyses using the Massachusetts 17-20 percent missing data cutoff, this method at least utilizes what there is of the information (or lack of it) in a somewhat more constructive fashion[39].

Making Sentencing Policy in Guidelines Research

One aspect of the task of constructing sentencing guidelines about which the Massachusetts sentencing guidelines project was apparently quite clear was that the sentencing guidelines research, construction, and final format would all be heavily affected by decisions made about sentencing policy.[40] This fact was acknowledged at the onset of the project; as we noted earlier, the initial press release declared that while the guidelines would be based primarily on the results of the sentencing research conducted, certain factors (shown to be of importance in the research) might not be allowed into the final sentencing guidelines model.[41]

The project invited Jack Kress, one of the early proponents of sentencing guidelines,[42] to address the judges

of the Committee on Probation and Parole about substantive policy issues that the Committee would have to deal with during the guidelines development process. Professor Kress did attend one session of the Committee on Probation and Parole for this purpose and also prepared a paper, titled Potential Guideline Variables,[43] which contained a summary of policy issues that had faced other guidelines developers and suggested strategies of how to resolve potential controversial issues. The policy issues that Kress suggested the Committee resolve before developing guidelines focussed on whether a number of items should be included in the sentencing guidelines and, if so, how those factors should be categorized and weighed. In particular, Kress suggested that the Committee pay close attention to the following items:

1. prior record: adult or juvenile, prior arrests or convictions, serious offenses or all offenses, decay of prior record, patterns in the prior record;

2. socioeconomic factors: employment, other social stability indicators;

3. method of conviction: trial vs. plea;

4. defendant cooperation;

5. consideration of the "real offense" rather than the statute of conviction;

6. other offense variables: injury, weapon use, drug use.

The judges of the Committee on Probation and Parole did, in fact, address a number of the policy issues listed above both before and after the session discussion with Professor Kress. In particular, the Committee made decisions about how to categorize the offender's prior record, whether to prepare separate guidelines for plea and trial cases, whether to include injury, and weapon use in the final guidelines, and how to categorize the seriousness of the current offense.

The Committee made a number of other policy decisions as well. First, the judges decided that the sentencing guidelines should be based on research conducted only on cases disposed of by Superior Court judges.[44] The judges on the Committee apparently believed that the sentences of District Court judges temporarily assigned to the Superior Court might in some way be different from the sentences of Superior Court judges and, thus, they did not want the guidelines to retain such differences.[45] Secondly, the Committee decided that the guidelines should eliminate traffic and "non-serious"

offenses from the defendant's prior record and that only charges that had resulted in "valid convictions" should be used in the computation of a prior record score.[46] The judges decided that the guidelines should not reflect any regional differences, or differences in sentences that might be a result of differential resource allocation in any county or area of the state.[47] The judges also decided to convert the sentences imposed for their sample of offenders into estimates of the amount of time that an offender would serve, and to then use those estimates as the dependant variable in future analyses.[48]

In Chapter 6 of this report we noted that Massachusetts sentences are usually imposed for a "package" of charges rather than for each single charge. As we also noted, the concept of a "package" of charges and a corresponding "package" of convictions is in part a result of the statutory limitations in force that do not allow for certain types of sentences to be given to certain offenses.[49] The Massachusetts sentencing guidelines project and the Committee on Probation and Parole recognized the importance of the "package" of offenses charged, and one of their most important policy decisions that affected the content and use of the sentencing guidelines developed was to use the entire package of charged offenses that resulted in valid convictions as a measure of the seriousness of the current offense. This decision to use all valid current conviction offenses to categorize the seriousness of the current offense broke with then traditional guidelines development methodology -- that traditional methodology usually allowed only for the use of the most serious offense in determining the appropriate guidelines sentence.

> We initially viewed sentencing as other researchers have presented it. The judge considers the charge and decides whether or not to incarcerate. If the decision is incarceration, a term is set. We no longer hold to this view because many Superior Court judges have told us that it is wrong....
>
> From discussions with the judges and by scanning disposition sheets, we have concluded that sentences are determined from the package of charges and events as well as the personal history of the criminal and other factors. (Sentencing Guidelines Project Staff to the Committee on Probation and Parole, 12 April 1979, pages 1-2)

The concept of "packaging" charges and offenses, deserves, in our opinion, the serious attention that was given to it by the Massachusetts project and that has not previously

been accorded to it in the development of sentencing guidelines. Most sentencing guidelines to date (and we really cannot, off-hand, think of any exceptions) have specified sentences for the "most serious" charge and conviction for sample cases. At the same time, however, the researchers often collected information about the combined sentences for all of the charges for the same sample case. One clearly obvious misfortune encountered in collecting charge and sentence information in this manner is, of course, that the resultant guidelines sentences may be either over- or under-stated for any particular offense. Consider, for example, two cases of armed robbery: in one case, the offender committed an armed robbery and a separate physical assault on, say, a store owner, and for both offenses received a combined total sentence of 10 years; in the second case, the robber may have committed only an armed robbery and received a sentence of 5 years for that one offense. The average armed robbery sentence for these two cases, considering only the most serious offense (as most guidelines do) would be 7.5 years. Thus, the first armed robber's sentence under the guidelines would have been under-stated, while the second offender's sentence would have been higher than it should have. By looking at all of the valid convictions for any case, the Massachusetts project was able to circumvent this problem. Instead of proposing a single guidelines sentence for a particular offense, the sentence would be adjusted to all of the conviction offenses.

The question remained, however, as to exactly how to combine different offenses to derive the guidelines sentence. The problem is one of the addition of apples and oranges genre. The obvious solution was to find one element that all offenses had in common. The project examined several alternative methods for classifying offenses according to some common denominator and rejected several of them. A crime-specific categorization of offenses was rejected because it would require too large a sample size and would not accommodate the notion of the offense package -- at least when several offenses of conviction were not of the same type.[50] The Sellin-Wolfgang method of isolating important elements of crimes was tested[51] and eventually rejected by the project.[52]

Judges were also asked to rate the seriousness of offenses, without consideration of the elements of offenses per se: initially, the judges on the Committee on Probation and Parole and the staff of the sentencing guidelines project completed this exercise,[53] and the resultant categorization of offenses into levels of seriousness was used in the first sentencing guidelines model proposed by the staff to the full Court. Following the introduction of the first sentencing

guidelines model, all of the Superior Court judges were asked to rate the seriousness of offenses as well to test the uniformity of the ranking procedure across judges. Though this exercise was completed, judges apparently indicated some dissatisfaction with this method of assigning seriousness scores to offenses -- the judges apparently felt that the assignment of numerical scores to offenses was bound to be somewhat arbitrary[54] -- and the method was subsequently abandoned.

The final sentencing guidelines model tested and then formally adopted by the Superior Court judges contained a different method of assigning seriousness scores to offenses. In this final method -- and it is one that we think is the most appropriate given the legal structure of sentencing in Massachusetts -- offenses are assigned a seriousness score based on the statutory maximum penalty that may be imposed for the offense. While this method does incorporate some statutory idiosyncracies in maximum penalties into the guidelines sentences, on the whole, we do not believe the idiosyncracies to be large enough to do any damage. Serious offenses usually carry severe statutory maximums, while less serious offenses have less severe maximum penalties. Added together, they produce a fairly accurate composite of the seriousness of all of the charges against a defendant.

The most important policy decision made by the Massachusetts sentencing guidelines project and by the Committee on Probation and Parole was to exclude a number of factors found to be predictive of sentence length[55] in the data analyses from the sentencing guidelines. The reason that these several factors were thus excluded was, quite simply, because the Committee felt that the factors were inappropriate considerations in sentencing decisions. In all eight factors -- among thse analyzed -- were found to be related to sentence length decisions: seriousness of charges, instances of gun use, instances of knife use, total extent of injury to victims over all charges, total seriousness of all prior offenses, total amount of "real" time to which the offender was committed prior to the present trial, whether the offender was free before trial, and the amount of bail.[56] The report noted that:

> Analysis indicated that these eight factors related strongly and independently to sentence length. However, only five of these factors were included in the guidelines model presented in the report because of decisions made jointly by the Committee and the staff. Prior incarceration time, free before trial, and amount of bail were the factors that were eliminated. (Massachusetts Sentencing Guidelines, November 1979:5)

The specific reasons offered by the project and the Committee in the report presenting the first version of the Massachusetts sentencing guidelines are worth noting for several reasons: the stated reasons reflect the practical considerations that must be faced when considering how guidelines will operate in a crowded court on a daily basis, as well as the perhaps more important ethical considerations inherent in guidelines development methodology.

> Prior incarceration time was not included for two reasons. Although its inclusion would lead to a more accurate guidelines model, the improvement would be slight because prior time served included aspects of seriousness of prior offenses, a factor already in the list of eight. Nevertheless, prior time served was a factor pondered by most judges in composing a sentence, and it would have been included in the guidelines model had the Committee and the staff not felt that prior "real" time would be too difficult to compute manually.
>
> The other factors omitted -- free before trial and amount of bail -- were also powerful ones. Every thousand dollars of bail added about one additional week of actual time to be served. Similarly, being free before trial resulted in an average of twenty months less time for an offender who was identical to another in all relevant aspects of the record except that the latter was not free before trial. A defendant not free could either have been unable to post bail or have been incarcerated for another offense. Research findings regarding pre-trial freedom may reflect the opportunity free defendants had to show rehabilitation and maturity. Those regarding high bail may reflect the past dangerousness of the offender or additional gravity of the current charge(s) that did not show up in weapon use, injury, or the offense seriousness score. (Massachusetts Sentencing Guidelines, November 1979, page 5)

By the time that the second version of the Massachusetts sentencing guidelines were introduced to the judiciary in May 1980, the list of relevant factors predictive of sentence outcomes had been reduced to seven. The two items of "instances of gun use," and "instances of knife use" in the first version of the guidelines were combined into one measure labelled "instances of weapon use". The Committee on Probation and Parole again clearly stated that the factors present in the guidelines were influenced by policy considerations:

> Statistical analysis found that seven factors were most important in sentencing defendants convicted after a

> trial: seriousness of current offenses, weapon use, injury to victims, amount of bail required and pre-trial confinement (and seriousness of prior offenses and prior "real" time served). The Committee on Probation and Parole excluded bail and pre-trial confinement (and prior "real" time served) from the guidelines because these factors were considered inappropriate but retained the other (four) factors. (Massachusetts Sentencing Guidelines, May 1980, page 1-2.) [57]

One further indication that the judges of the Superior Court intended that the sentencing guidelines reflect a statement of sentencing policy was that the judges considered the possibility of racial disparity in their sentences during the early stages of the project's research. As will be recalled from Chapter 6 of this report, charges had been leveled against the Superior Court judges of racial sentence disparity. [58] The judges responded to these charges by stating that if such racial disparity did exist, it should be researched and corrected. And the correction of racial disparity in sentencing would reflect a sentencing policy statement by the court -- a statement that consideration of race in the imposition of sentences is clearly inappropriate.

The Massachusetts sentencing guidelines project was assigned the task of investigating possible racial biases in sentences imposed in the state as one part of their research. The project prepared two reports on this issue. The first report, A Study of Racial Disparity in Massachusetts Superior Court Department, found that:

> The results indicate that although the race of the defendant by itself does not directly influence the length of sentences, race indirectly influences sentence lengths because of the way judges weigh and combine the six factors. White defendants receive an advantage of significantly shorter sentences after trials relative to black defendants, and black defendants receive an advantage after pleas of slightly shorter sentences relative to white defendants. (Marx, 1980a:3, emphasis added)

The second report, The Question of Racial Disparity in Massachusetts Superior Court Sentences, was released one month after the first, after additional research suggested by the judges had been completed. The research findings stated in the second report do not differ much from those of the earlier report in content, though the language used is considerably more moderate:

> The results demonstrate that the race of a defendant by itself does not directly influence sentence length. However, the variables that do account for sentence length are weighted differently for black and white defendants. As a result, black defendants receive an advantage of shorter sentences relative to white defendants in the great majority of cases that are disposed of by a plea of guilty. White defendants, on the other hand, have the advantage after trials of substantially shorter sentences compared to black defendants. When pleas and trials are considered together, there is no disparity in sentence length between black and white defendants. (Marx, 1980b:2)

Four final policy decisions were made about the Massachusetts sentencing guidelines. These decisions were different from the earlier decisions made, for the entire Superior Court judiciary was involved in one way or another with the resolution of these issues. The Court, in order of occurrence, decided that:

> 1. Use of the sentencing guidelines would be voluntary among Superior Court judges, not mandatory;
>
> 2. Judges would be allowed a 50 percent range around the guidelines sentences in trial cases,[59] and approximately an 80 percent range around the guidelines sentences in plea cases;[60]
>
> 3. Separate guidelines sentences would be provided for trial and plea cases; and
>
> 4. Consultation of the guidelines would be, as of May 1980 and restated in June 1981, mandatory for trial cases and optional in plea cases. When the guidelines were consulted in plea cases, the trial guidelines sentences would be the only guidelines sentences consulted.

The first of these four policy decisions -- that the use of the guidelines would be voluntary among Superior Court judges -- was initially a response to judges' reluctance to adopt formally and mandate a sentencing model that had not yet been tested to the satisfaction of all members of the Court. Though the judges on the Committee on Probation and Parole were familiar with the guidelines operation, the remaining forty or so judges were not entirely convinced that the guidelines would really produce workable sentences in specific cases. Thus, the initial use of the guidelines was voluntary during this preliminary "testing" period, which began in November 1979.

Secondly, it was decided by the judges that the guidelines would allow for a range of guidelines sentences. The entire Court agreed to the ranges proposed during the November 1979 judicial conference at which the guidelines were introduced; but the staff of the guidelines project and the Committee on Probation and Parole were actually responsible for the decisions as to how much of a range would be provided. The rationale used for deciding upon a fifty percent range in trial cases and an eighty percent range in plea cases was stated by the Committee and staff in the November 1979 sentencing guidelines report:

> The figure of 50% (for trial cases) was chosen because the Committee believed greater latitude to be inconsistent with the purpose of guidelines, which is to incline more extreme jurists toward the average of the entire Court. On the other hand, ranges set on a lesser percentage than 50% would encompass so few of the actual decisions that judges might feel unduly restricted by the guidelines. (page 7)

> Furthermore, the Committee set these ranges (for plea cases) wide enough so that at least 80% of sentences are expected to fall within the guidelines. Most of the sentences that will fall outside the guidelines are expected to be just slightly less than or slightly more than the guidelines sentences. The Committee placed these "high" plea sentences outside the guidelines to assure most defendants a lesser sentence in return for their pleas. (page 9)

The above quotes do point to the fact that the spread of sentences in Massachusetts was probably quite large. A range of eighty percent was needed for the plea guidelines to capture the same percentage of the original sentences as guidelines sentences; similarly, a wide range of fifty percent was needed to encompass some unknown quantity of trial cases (the above quote does not specify an exact amount). The provision of a range that was less wide for trial cases, the report noted, would have resulted in guidelines that would have been based on only a few of the trial cases; this statement does seem to suggest that the dispersion of sentences in trial cases was also quite substantial.

The third policy decision we noted above was that separate sentencing guidelines would be constructed for plea cases as distinct from trial cases. This fact has already been acknowledged by the preceding discussion involving the different percentage ranges that the first version of the Massachusetts sentencing guidelines used in plea cases as opposed to trial cases. The point we wish to emphasize here,

though similar, is that the use of a different range for one type of case versus another necessarily results in two different schemes of guidelines sentences. The end result of providing a wider range around the guidelines sentences for plea cases, as opposed to trial cases, is that the range is necessary to accomodate the greater leniency often shown to cases disposed of by a plea of guilty. While the legitimacy of the practice may be questioned, the existence, in the earliest version of the Massachusetts guidelines, of two distinct sets of guidelines sentences -- one for trial, and the other for plea cases -- was a straightforward acknowledgement of the widespread practice of a sentence reduction after a plea of guilty.

The judges of the Superior Court, however, were not particularly comfortable with this aspect of the sentencing guidelines; no doubt, because they believed that public acknowledgement of reduced sentences for pleas of guilty would only result in an increased clamor about sentencing disparity -- the very problem that they were attempting to examine and, if necessary, correct. As a result, when the guidelines were presented to the full Court for adoption, in November 1979, the Court voted to "test" the trial guidelines and to allow judges to consult the plea guidelines in plea cases when they thought appropriate. By the following May (1980), when the second version of the sentencing guidelines were presented to the full Superior Court for adoption, a separate guidelines format for plea cases was entirely missing from the package. Instead, the guidelines contained only guidelines sentences for trial cases. And judges, though mandated to "test" the second version of the guidelines on all trial cases, were free to consult those same trial guidelines for guidance in plea cases as they thought appropriate. The guidelines were formally adopted for use (as opposed to continued "testing" by the Court) in June 1981, again with the stipulation that the guidelines were not mandatory for plea cases. This change over time in what types of cases the guidelines were applicable to, was the final policy decision made by the judges of the Superior Court.

Structure of the Massachusetts Sentencing Guidelines

The Massachusetts sentencing guidelines project developed two sentencing guidelines models for use by Superior Court Justices. Both models were initially referred to, by the judges and the project, as test versions rather than as sentencing standards formally accepted by the Superior Court.

The initial version of the Massachusetts sentencing guidelines was introduced to the judges in November of 1979 (Massachusetts Sentencing Guidelines, November 1979). After

subsequent testing and revision of several procedural as well as substantive aspects of that first model, the second -- and thus far final -- version of the guidelines was adopted by the Superior Court judges in May 1980 (Massachusetts Sentencing Guidelines, May 1980). Again, that model was implemented on a "test basis"; however, all members of the judiciary were urged to consult the guidelines for individual cases sentenced and to provide written feedback to the Committee on Probation and Parole about deviations from the guidelines sentence. Finally, as we noted above, this second version of the sentencing guidelines was formally adopted by the judges in June 1981.

The first and second versions are quite similar in the general kinds of items included even though the actual number of items considered is one more in the first model than in the second. For the purpose of describing the Massachusetts sentencing guidelines, the discussion mainly will focus on the latter model, although differences between the two models that are especially important to the adequacy of the guidelines will be examined in some detail.

The second and final version of sentencing guidelines developed by the Massachusetts Superior Court sentencing guidelines project calls for the consideration by judges of four different types of information to determine the appropriate sentence decision. The types of information considered, as well as the sentence that may be derived from consideration of these items, are unique to the Massachusetts system in many respects and need further elaboration here.

That the Massachusetts sentencing guidelines are to be used for only a select group of cases is the first unique characteristic of this statewide guidelines system. The guidelines are "applicable in convictions after trial but may at the option of the judge be used as a reference in pleas." While only trial cases sentenced by Superior Court justices were used as the data base for the guidelines construction, the final version of the guidelines does not prohibit use of the model by District Court judges assigned to the Superior Court.

The four items of information included in the most current Massachusetts sentencing guidelines are: current offense seriousness, prior offense seriousness, the sum score of instances of weapon use for each current conviction, and the sum score of degree of injury to the victim for each current conviction. The only difference between this and the earlier versions of the Massachusetts sentencing guidelines was that while the current guidelines measure use of any type of weapon, the first version differentiated gun use from the use of a sharp or dangerous weapon.

There are exclusions and special scoring techniques for each of these four items. In addition, policy decisions made by the Massachusetts sentencing guidelines project's steering committee -- the Committee on Probation and Parole of the Superior Court -- have modified the content of the guidelines to exclude information items considered inappropriate to a guidelines system and to broaden the scope of other included items beyond the boundaries of the data collected.

The Seriousness of the Current Offense

One computational basis of the Massachusetts guidelines sentence is the statutory seriousness of the current conviction offense. Each offense resulting in a valid conviction (about which more will be said in a moment) is assigned a seriousness score using a 5 point scale that ranges in value from zero to four. This scale categorizes offenses according to the maximum sentence that can be imposed for the offense by statute as shown in the chart below. Only felony offenses are assigned a score within the guidelines; although misdemeanor offenses may carry up to a 2 1/2 year jail sentence (the maximum allowable sentence to a House of Corrections in Massachusetts), they are not included in the guidelines model.

Seriousness Score	Statutory Penalty Range
0	Maximum penalty is less than 5 years
1	Maximum penalty is exactly 5 years
2	Maximum penalty is more than 5 and up to and including 10 years
3	Maximum penalty is more than 10 and up to and including 20 years
4	Maximum penalty is more than 20 years

Once the current conviction offense(s) have been assigned a seriousness score that score is then multiplied by the number of counts listed as part of the conviction indictment. This results in a score value representative of both the seriousness of the offense and the number of those offenses of which the offender is convicted.

The seriousness score values are then totalled for all current conviction offenses that fall within the guidelines jurisdiction. The guidelines direct judges as follows:

> Assign a score of 0 to offenses which are dismissed, not prosecuted, result in a finding of not guilty, continued without a finding, filed without a finding of guilt or filed after a finding of guilt, as well as to all misdemeanors.

If a charge results in any one of the dispositions listed above, then, that charge (offense) is not included in the guidelines. The guidelines further direct that any "duplicitous" charges should be assigned a zero score and not counted in the computation of the guidelines sentence. The one remaining exception to the scoring procedure outlined above involves offenses where the defendant attempted but did not actually commit the crime or where the defendant acted as an accessory to the commission of the offense. The guidelines again direct:

> If the offense is an "attempt," score it the same as the offense itself unless the attempt is separately enumerated. "Accessory before the fact" receives the same score as the offense itself. "Accessory after the fact" receives a score of 2 since the maximum penalty is 7 years.

According to these directions, "attempted offenses" listed in addition to the completed offense are also assigned a seriousness score (thus doubling the seriousness score for that charge) and all "accessory after the fact" charges are to receive a seriousness score value of 2. As will be seen during our re-analysis of the guidelines sentencing data we have not followed this latter direction in the computation of seriousness scores for some offenses. Only those offenses that would have merited a seriousness score of 2 or higher -- had the defendant had been the principal perpetrator of the offense -- were assigned a seriousness score of 2 if the offender was finally convicted as an accessory after the fact. The directions given by the guidelines for these cases are misleading as they would have resulted in inflated seriousness scores for accessory after the fact cases in comparison with the score that would be received by the major perpetrator of the same offense.

The major problem encountered when computing the seriousness of the current offense is that the seriousness score ranges, provided by the Massachusetts guidelines to tell judges and probation officers how to classify each offense's seriousness, do not in all cases correspond with the statutorily defined maximum sentence. The statutes are not phrased in the same language as are the guidelines scoring instructions, so one is left uncertain, at times, as to how to score the seriousness of some offenses. The most common example of a scoring problem caused by lack of agreement between the statutory penalty range and the guidelines range involves offenses which by statute may receive a sentence of not more than 5 years. Such an offense can thus receive any sentence less than or equal to 5 years. The question raised here would be whether the appropriate seriousness score for

such an offense would be 0 (a maximum penalty of less than 5 years) or 1 (a maximum penalty of exactly 5 years). Clearly neither of these options reflects the language of the statutory penalty.

Despite the difficulties that are encountered when attempts are made to classify offenses according to statutory maximum penalties, the method used to determine offense seriousness is much clearer in the second and final guidelines version than in the earlier model developed. The seriousness of any particular offense was determined in the earlier version by consulting a table of offenses rank-ordered by six levels of severity. The scale of seriousness used was empirically developed by the guidelines project after a number of Superior Court justices responded to a questionnaire that asked them to assign seriousness scores to offenses. The Offense Seriousness Chart that provided the basis for the determination of each act's seriousness was not inclusive of all offenses -- approximately 10 offenses were listed for each scale value. Offenses could be assigned a scale value of 1, 2, 3, 4, or 6 based on this chart; no offenses at all were listed as deserving a seriousness score value of 5. Offenses not specifically listed in the chart were assumed to have a score of zero.

Dangerous Weapon Use

The second item included in the Massachusetts sentencing guidelines is the use of a dangerous weapon. The defendant's weapon use score is computed by assigning a point for each instance of use of a dangerous weapon during the commission of the crime, regardless of the number of charges that may result from the crime. As the guidelines explain:

> If an offender uses one weapon and this results in charges of Assault and Battery with a Dangerous Weapon (gun), Carrying a Dangerous Weapon (gun), and Assault with Intent to Murder, then enter a "1" in this column.

If more than one kind of weapon is used during the commission of the crime, a multiple score would be assigned. For example, if an offender used both a gun and a knife during a single criminal event, a score of 2 would be given to that offender.

Unlike the seriousness of the current offense (and, as will be noted a little later, the seriousness of prior offenses), the scoring of the weapon use and degree of injury items involves some discretionary scoring judgements. What exactly constitutes use? Does showing a gun constitute use, or must the offender point the gun or shoot it for the action

to be considered use of a weapon? And in some special situations, how should instances of use be counted? For example, if a defendant waved a gun at a grocery store clerk, that action would be counted as one instance of weapon use. But had 4 or 5 shoppers been present when the offender waved the gun, would the score for weapon use then be a score of 4 or 5? The answers to questions such as these are not provided by the guidelines; judges must decide within the context of each case the defendant's weapon use score. The guidelines state that "after conviction, the trial judge will inform the probation officer of the score(s) to be recorded" and it is left to the discretion of the judiciary how best to determine that score.

It was earlier mentioned that the only difference between the first and second versions of the sentencing guidelines involved items relating to weapon use. The current guidelines ask for one weapon use score to be computed for each defendant, regardless of the type of instrument used as a weapon. The first version of the guidelines differentiated between use of a gun and use of a sharp or other clearly dangerous weapon such as a knife or other sharp weapon, an explosive or an "infernal machine". (Baseball bats, shod feet, hammers etc. were stated as exclusions to the first guidelines; presumably these items are also excluded from consideration in the current guidelines.)

While essentially the same information is captured whether weapon use is calculated for combined or separate instances of gun use and sharp, dangerous weapon use, a substantial difference in the guideline sentence would occur if the first guidelines model were used rather than the current model. The sentence differential would result because the weights assigned to the two categories of weapon use in the first model were both much higher in value than the weight assigned to the combined weapon use score in the current guidelines. More will be said about the weights assigned to these and other items in the "computing the guidelines sentence" section that follows.

Degree of Injury

The degree of injury to the victim is the third item of consideration in the Massachusetts sentencing guidelines. Similar to the weapon use assessment, the determination of the amount of injury suffered by the victim is left to the discretion of the judge with general descriptions of degrees of injury supplied as standards to aid in that determination. At the Spring 1980 Massachusetts Judicial Conference where the current guidelines were explained to the Superior Court judiciary, the judge presenting the guidelines noted that

determinations of degree of injury are made routinely by judges and not exclusively within the jurisdiction of the criminal court.

The final scale provided by the guidelines to assess injury ranges in value from 0 to 5. Each victim to whom injury was caused -- or each instance of injury -- is rated according to this six point scale. While degree of injury to the victim was an item of information collected in the 1440 case sample by the guidelines project staff for analysis, inclusion of the item in the form that it appears here is not solely a result of its statistical ability to explain variation in the judges' sentencing decisions but is a result of policy decisions as well. The information collected by the guidelines project assessed injury on a 5 point scale, ranging in value from 0 to 4 -- no injury to the victim to death of the victim as shown below. In addition, psychological injury to the victim was never included in any of the data collected.

Injury Generally:
0 = No injury
1 = Emergency treatment only/minor injury
2 = Hospitalization/no permanent injury
3 = Permanent significant injury
4 = Death

In contrast to the degrees of injury specified by the Massachusetts data collection coding instructions, victim injury is assessed by the current sentencing guidelines on a six point scale which also ranges from no injury to the victim to death of the victim, but which has included an additional category. As best can be discerned, that additional category (as comparison of the rating scheme below with the coding manual version above will show) appears to be injury level 3 -- serious physical/psychological injuries that are not permanent. In addition, the guidelines injury scale includes an assessment to be made by the sentencing judge of psychological trauma or injury to the victim.

Injury Scale:
0 = No injury
1 = Minor physical injuries requiring emergency treatment/psychological trauma of limited duration
2 = Significant, but moderate rather than critical injuries/psychological injury that results in temporary or partial disability
3 = Serious physical/psychological injuries that are not permanent
4 = Major injuries, physical/psychological in nature which result in permanent residual disability
5 = Death

Obviously, the degree of injury specificity spelled out in the Massachusetts sentencing guidelines cannot be duplicated by the Massachusetts sentencing data. This point is one critical to an analysis of the Massachusetts guidelines for it would have been impossible for them to develop an injury scale of this nature in their guidelines had they truly developed guidelines descriptive of past sentencing practice.

However, the Massachusetts guidelines claim only in part to be descriptive of past sentencing practice. The inclusion of psychological injury is one specific instance of a policy decision made by the judges of the Superior Court to include an item in the guidelines that could not possibly have its importance in explaining past sentencing decisions verified statistically using the data available.

The Seriousness of Prior Offenses

The final item included in the Massachusetts sentencing guidelines is that of the seriousness of the defendant's prior criminal history. The procedure used to determine just how serious an offender's prior convictions are is the same as that used to compute the seriousness of the current offense. The seriousness of each adult felony offense in the prior record is assigned a seriousness score from the statutory maximum penalty score also used to classify the current offense seriousness. Only convictions that have not been disposed of by the filing of the case, or by a continuance without a finding or with a finding of probation, etc. may be included in the scoring.

Out-of-state convictions as an adult for felony offenses are also included in the computation of the seriousness of the prior record; the guidelines direct that "the Massachusetts offense seriousness score which is closest to the non-Massachusetts conviction" should be assigned to the offense. The guidelines further advise that if the out-of-state record is "not reliable", the information pertaining to those convictions should be verified by checking with the state where the defendant was convicted.

Once the prior conviction offenses have been assigned the appropriate seriousness scores, the scores are then multiplied by the number of counts for each offense. Information pertaining to the number of counts is usually listed together with the prior indictment offense on the offender's Massachusetts prior record sheet. The combined seriousness and count scores for each prior offense are then totalled to result in the score representative of the seriousness of all prior convictions.

The Massachusetts sentencing guidelines prior offense seriousness score is obtained by considering all valid convictions; included among those convictions deemed valid for computational purposes are offenses that have been committed a substantial amount of time prior to the current offense. A number of statewide sentencing guidelines models proposed to date have included a "decay factor" in the final guidelines version. Generally, the "decay factor" is an instruction specifying that offenses committed more than 7 or 10 years prior to the current offense should not be included in the prior offense score. The Massachusetts guidelines do not include such a decay factor; all prior offenses, regardless of age, are used to calculate the seriousness of the offender's prior criminal history.

Computing the Guidelines Sentence

Once scores have been derived for each of the items considered in the guidelines -- seriousness of current and prior convictions, injury, and weapon use -- the guideline sentence can be computed. The scores obtained for each of the above categories are multiplied by a weighting factor assigned to each category. The weighting factors are representative of months of imprisonment and, in a sense, the weight of each factor relative to the other factors included in the model specifies how heavily the item weighs in the guideline sentence. The single major difference between the two sentencing guidelines models developed by the Massachusetts project was that the weights assigned to each category changed substantially from the first to the final version. Since the weights are indicative of months imprisonment to be served even small changes can result in dramatically different guidelines sentences. The weighting factors assigned to each of the categories in the first and second guidelines models are described by the chart below.

Item	Weight Version 1	Version 2
Current Offense Seriousness	1.5	2.1
Dangerous Weapon Use	-	9.0
Gun Use	15.0	-
Sharp, Dangerous Weapon Use	10.0	-
Injury to Victim(s)	8.0	9.0
Prior Offense Seriousness	1.1	1.6

The most substantial weight decrease is seen for the weapon use items. The first guidelines model assigned 15 months of imprisonment for each instance of gun use and 10 months for each instance of sharp or dangerous weapon use. The current version requires that 9 months be assigned for

each instance of use of any type of weapon -- six months less than that earlier proposed for gun use and 1 month less than that earlier suggested for sharp or dangerous weapon use.

The weights for the remaining items each increased in value. Presumably the weights for current and prior offense seriousness have increased in the second model in part because the seriousness scale has decreased from a 6 to a 4 point scale. Also, the weights provided by a regression equation for each item would understandably be affected if the scoring or weight of any included item was changed.

After the original scores for each category have been multiplied by their corresponding weights, the new scores are then added across all of the categories. If the sum contains any decimals (or fractions of months) those decimals are rounded to the nearest whole number. (A decimal of .5 is rounded up to the next highest number of months.) The rounded sum derived is the initial guideline sentence, or the number of months that an offender could be expected to serve before eligible for parole.

Since the second version of the guidelines is the final model developed and used thus far, we will use the weights assigned in that model for illustrative purposes. Suppose, for example, that an offender is convicted of one offense of Assault and Battery with a Dangerous Weapon. Since the statutory maximum sentence for this offense is more than 5 up to 10 years, the conviction receives a current seriousness score of 2. During the commission of the offense, the defendant used a weapon to threaten and caused some psychological trauma to the victim. The defendant's prior offenses also were assaultive in nature -- Assault and Battery with a Dangerous Weapon and the lesser offense of Assault. According to this scenario the offender's scores on each of the guidelines items and the new scores produced after the original figures were multiplied by their associated weights would be as follows:

Item	Score	Version 2 Weight	Weighted Score
Current Offense Seriousness	2	2.1	4.2
Weapon Use	2	9.0	9.0
Injury to Victim	1	9.0	9.0
Prior Offense Seriousness	3	1.6	4.8
	Initial Guideline Sentence		27.0

When the procedure for computing the guideline sentence specified by the model has been completed, the initial guideline sentence score that should be given to this offender is 27 months of time incarcerated prior to parole eligibility.

This figure can also be thought of as an average sentence for an offender with these scores on only these four characteristics.

The Guidelines Range

The Massachusetts Superior Court judiciary did not want the sentencing guidelines to pronounce a single sentence of a set number of months for a specific offender; rather they insisted that the guidelines provide a range of months above and below the initial guideline sentence that would allow them the flexibility to adjust the average sentence if other items of information not specifically included in the model proved to have some special ameliorating or exacerbating effect in a particular case. If they sentenced an offender to a period of months within the guidelines range provided, the sentence would be deemed to be a guideline sentence.

The initial guideline sentence was thus further modified by the addition of a range of 50 percent above and below the average sentence to the model. For example, if an offender received an initial guideline sentence of 27 months as our hypothetical offender did the range within which a given sentence would be acceptable would be from 13 to 41 months time to be served prior to eligibility for parole.

An initial guideline sentence of zero has both an upper and lower range of zero and such a score value calls for a non-incarcerative sentence. The choice of a non-incarcerative disposition is also available as the lower range sentence for offenders who receive an initial guideline sentence between 1 and 5 months. Other than the initial guideline scores of zero to 5 months, the 50% range around the initial guideline score rule applies in the determination of the appropriate sentence. While a guideline score of 5 would allow for a lower range sentence of non-incarceration, an initial score of 6 would have a lower range guideline sentence of 3 months incarceration and an upper range sentence of 9 months incarceration prior to parole eligibility.

Converting the Guideline Sentence to An Actual Sentence

In order to analyze judicial sentences to incarceration and to develop sentence length guidelines, the Massachusetts project staff first had to convert stated sentence dispositions for the various county and state correctional institutions into reasonable estimates of the amount of time that an offender would actually be incarcerated. The sample of defendants was one of persons sentenced fairly recently; thus, in many instances offenders who received incarcerative sentences were still doing time and the actual length of their sentence was still unknown.

The guidelines project staff estimated the amount of time that offenders would be incarcerated by applying the formal and informal rules concerning parole eligibility (that we described in Chapter 6) to each sentence. Of course, these time estimations are not perfect -- although the rules concerning parole eligibility are followed more or less exactly in most cases, paroling authorities throughout the state are able to alter the parole eligibility when the special characteristics of the offender or the offense present a reasonable probability of future criminal behavior should the offender be released. Without information from the Massachusetts Parole Board regarding each of the offenders in the sample, however, cases where deviations from the normal paroling process would occur could not be identified by the guidelines project. The ultimate effect of the estimation of length of sentence, thus, would be to underestimate the amount of time to be served by offenders perceived as particularly prone to continued criminal behavior.

The initial guideline sentence, as well as the 50 percent range of months on either side of that sentence, expresses the amount of time that an offender should serve prior to parole eligibility. Thus, before a judge can pronounce an offender's sentence, the guideline amount of time must be converted back into an actual sentence by using the formal and informal parole eligibility rules in reverse.

The Massachusetts sentencing guidelines do not suggest the particular place or institution where an incarcerated offender should serve time. Of course, the amount of time specified by the guidelines does limit the choice of institution in many instances as we detailed earlier in Chapter 6. For example, the maximum pronounced sentence to a House of Correction (county jail) in Massachusetts cannot be more than 2 1/2 years. Thus, any single sentence to a House of Correction cannot expect an offender to serve more than 15 months as, under the parole eligibility rules, offenders sentenced to a House of Correction must serve one-half of the pronounced sentence. Similarly, statute requires that offenders sentenced to state prison must serve one-third of their minimum sentence but not less than 1 year (for offenders convicted of non-persons offenses). Thus, a guideline sentence of less than 12 months to be served could not be translated into a state prison sentence.

A number of suggested guideline sentences may, however, be translated into pronounced sentences to more than one specific correctional facility. In these instances, the sentencing judge must first decide where he wants an offender to serve the time incarcerated. Once a place of incarceration has been chosen, the guidelines sentence is converted into an

actual sentence to that place using the parole eligibility rules particular to that institution.

Several specific offenses may have slightly different parole eligibility rules, but on the whole the parole eligibility date of most sentences can be determined by using the general sentence conversion rules pre ented in Table 6.1 of the preceding chapter.

Of course, those rules are not absolute; the paroling authorities have considerable flexibility in determining the exact date of parole eligibility to say nothing of their ability to grant or revoke good time credits. Yet we were able to ascertain, from interview data with judges, prosecutors and others in the Massachusetts system, that the formal and informal rules outlined above do serve as good general guides to the length of time an inmate will be incarcerated, and thus lead to what are appropriate sentences.

Deviating from the Guidelines

The structure of the Massachusetts sentencing guidelines is quite straightforward: one determines the seriousness of the current offense(s), the degree of injury, the number of instances of weapon use, and the seriousness of the prior offense record; each of the scores derived from these computations is then multiplied by its corresponding weighting factor and the resultant weighted scores are summed to derive the guidelines sentence length. At this point a chart, which indicates the range of acceptable sentence lengths around the guidelines sentence, is consulted. The judge then picks a sentence length thought appropriate for the case at hand from the relevant range and, using the rules regarding parole eligibility, calculates what sentence he would have to impose in order to have the offender serve the required time in the institution the judge has chosen.

For example, suppose that the weighted sum of the guidelines factors in a particular case (offense seriousness, prior record, etc.) comes to 22 months; this is the "guideline" sentence in that case. The chart just mentioned will show that the range associated with this sentence is from 11 to 33 months (i.e. plus or minus 50 percent); any sentence resulting in a term in that range will not be considered a deviation. Suppose that the judge wishes to send this offender to Walpole State Prison, and that the offender is one to whom the one-third minimum parole eligibility rules apply: the judge may pass a sentence of between 33 and 66 months to Walpole, without "deviating" from the guidelines. Since judicially imposed sentences to odd numbers of months are unusual, it seems reasonable to assume that the judge would in

practice regard the allowable range as being three years (or 36 months) to five years (60 months); but this need not be the case. Would a sentence of six years, for example -- representing a rounding-up to the next year of the upper limit of the guidelines range -- be regarded as a "deviation"? What about a sentence of two and a half years (i.e. 30 months)?

Providing that the sentence chosen is one which should (given estimated parole eligibility) result in a term within the range allowed by the guidelines, the judge has sentenced within the guidelines. If the sentence does not result in an estimated length of term that is allowed by the guidelines, the chosen sentence is considered to be a deviation (i.e., what is elsewhere called a "departure") from the guidelines, and the judge is expected to provide some written reason for the deviation. There are no further rules to govern the deviation process in the Massachusetts sentencing guidelines -- of the kind found in Minnesota, for example (see Chapter 9).

Implementing the Massachusetts Sentencing Guidelines

Given the political climate in Massachusetts concerning the topic of sentencing in general, and sentencing disparity in particular, the judges comprising the Committee on Probation and Parole, and the staff of the Massachusetts sentencing guidelines project, were -- or so we infer, from what they told us -- very concerned about the introduction, reception and ultimate use of the sentencing guidelines they were developing. Not only was this concern focussed on persons outside of the Superior Court -- such as legislators, news reporters, and the public; it was also directed toward a number of judges on the Superior Court who (it was thought) either might not agree with the concept of sentencing guidelines in general or might not approve of the particular guidelines eventually developed by the project for the Court. From our research in Massacusetts, and from our observations of judicial sessions where the guidelines were introduced to the judges, we conclude that the concerns of the project staff and of the Committee were not misdirected.

In an effort to pave the way for a smooth and uncontroversial acceptance of the sentencing guidelines, project staff and the Committee on Probation and Parole undertook a plan of implementation while the guidelines themselves were still in an early stage of development. The "implementation plan" (and that is our phrase, not theirs) appears to have involved several stages: involving the judiciary in sentencing guidelines policy decisions and enlisting their support; announcement of the sentencing guidelines to legislative bodies concerned with sentencing;

and the training of probation department staff in the use of the sentencing guidelines in particular cases. The final phase of the implementation plan would involve introducing the sentencing guidelines to the system itself, and that phase would be accomplished by the actual use of the guidelines by judges in court.

The first phase of that plan obviously involved alerting the judges to the concept and development of the guidelines. The Committee on Probation and Parole did not want to take the position of being the sole body that would be unilaterally stating the Superior Court sentencing policy; thus, the input of the remaining judges of the Superior Court in policy decisions was seen as a required part of the guidelines development process. Secondly, the judges of the Committee on Probation and Parole, and the research staff of the sentencing guidelines project, realized that the ultimate use and acceptance of the sentencing guidelines model that was proposed would depend in part on judges' familiarity with it. To accomplish these ends -- of receiving input on policy decisions, and of familiarizing the judges with the guidelines concept and the model in development -- the sentencing guidelines project staff prepared and distributed to the judges a series of "newsletters" or "staff reports" prepared under the direction of the Committee on Probation and Parole.[61]

Three "staff reports" [62] were distributed to the judges on the Superior Court. The first staff report explained the origin of the Massachusetts sentencing guidelines project and also noted the names of judges appointed to the Committee on Probation and Parole who could be contacted if questions about the guidelines arose. This report also explained the general concept of sentencing guidelines -- stressing the point that much of the structure of the guidelines would inevitably be the result of policy choices made by the Committee on Probation and Parole -- and invited the judges to attend meetings of the Committee to "express views regarding guidelines" or to obtain "additional information about (policy) issues".[63]

The second report sent to the Superior Court judges from the staff did not discuss further the concept of sentencing guidelines; rather, the report detailed for the judges the general structure of drug legislation then pending in the legislature. (The content of that second report was apparently chosen because of the effect such legislation could have on the ultimate use of sentencing guidelines if it were to be passed, but also just because the judges wanted to know the facts about the proposed law.) So, it was not until the third (and final) report that the further progress of the

staff in the guidelines research was presented. Aside from a brief introduction that restated the purpose of the guidelines project, the report contained a summary of the preliminary results of a frequency analysis of the data. The preliminary results presented included an overview of defendants' social, economic, and demographic characteristics, arrest data, charge and dispostion data, and prior record data. To our knowledge, this is the only report that lists these basic descriptive facts about the offender sample used for the guidelines research, and this report was only made available to the judges of the Superior Court -- not to the public.[64]

The staff reports sent to the judges of the Superior Court gave them a general notion of what guidelines were to be about and told them some brief facts about the defendants they had sentenced in the past. The reports did not really address any of the policy issues that the Committee, as we noted above, eventually was to decide. Nor did the reports specify what the developed guidelines were to include by way of factors. Thus, the first time that most of the judges on the Superior Court would learn about the sentencing guidelines that the project had developed was at the judicial conference where the guidelines were introduced.

Judicial Reception of Sentencing Guidelines

The presentation of the guidelines to the Superior Court took place on 2 November 1979 in Plymouth, Massachusetts. Staff of the Evaluation of Statewide Sentencing Guidelines Project were able to attend this session, at the invitation of the Committee on Probation and Parole.[65]

The introduction of the sentencing guidelines to the judiciary began with a presentation of the results of the sentencing study conducted by the sentencing guidelines project research staff. The project director, Dr. Michael Hutner, and the project's research director, Amy Craddock, first discussed the project's activities and the statistical results of the study and then provided an overview of how the guidelines would work. This first presentation of the Massachusetts guidelines, though it was not smoothly run,[66] did accomplish the intended purpose of introducing the guidelines to the judges, and eliciting thoughtful criticism of the proposed guidelines model. The major criticisms of the guidelines offered by the judges at this session were three, as we perceived them. Stated briefly, the three reasons for opposing the guidelines were:

1. The guidelines were being introduced at a time when political opposition from the legislature (and especially from the members of the legislature

responsible for the development of the pending presumptive drug law) was sure to result. The legislature, the judges believed, would respond to the announcement of judicially developed and mandated guidelines by enacting stiffer mandatory minimum legislation for an unforeseen additional number of offenses. A number of judges expressed the view, however, that whether the guidelines would bring about the enactment of further mandatory sentences was not really the issue; rather, these judges believed that the promulgation of sentencing guidelines by the judiciary was an encroachment upon the function of the legislature to stipulate crimes and penalties.[67]

2. A number of judges disagreed with the methods that were to be used to derive scores for two particular factors that were included in the sentencing guidelines. First, the judges noted that in deriving a score for the injury component in the guidelines, no account was taken of psychological injury to the victim; physical injury only was to be considered. Secondly, the method for determining the seriousness of current and prior offenses did not appear to the judges to have any sound basis. It will be recalled that for the first version of the sentencing guidelines, crimes had been ordered into a scale by the judges on the Committee according to their perceptions of the seriousness of the act. Many of the other judges of the Superior Court, however, did not feel that such a technique was a legally sound method for deriving seriousness scores.

3. A number of judges expressed some doubt as to whether the provision of a separate, and less severe, sentencing table for plea cases, as opposed to trial cases, would be either politically palatable or legally acceptable.

It is apparent that while it would be fairly easy for the sentencing guidelines project staff, together with any oversight committee, to resolve the two latter criticisms of the guidelines, the first issue is much more problematical. The Supeior Court was asked, nonetheless, to vote on whether or not to use the guidelines at that November 1979 session of the conference. The result -- a vote to "test" the guidelines until the following conference session of the court -- reflected a compromise on the part of the judges critical of the guidelines and those supportive of the guidelines. The judges believed that merely "testing" the guidelines would accomplish two ends: the legislature would not be offended since the guidelines had not been formally adopted without their advice, and the guidelines' content could be modified in

the interim period to reflect concerns of the judges about the scoring of specific factors included in the model.

That is, in fact, what subsequently happened. The second version of the Massachusetts sentencing guidelines was presented to the full Superior Court on 18 April 1980 in Framingham, Massachusetts.[68] The structure of this second and final version of the sentencing guidelines has been described in detail above, as have been the differences between the first and second versions, and we do not intend to repeat that discussion here. What we would emphasize, however, is that the modifications to the guidelines in the second version were clearly a response to the concerns of the judiciary as evidenced at the judicial conference where the first version was presented, and by the judges input about the guidelines during the testing period.

Some comments should first be made about the "presentation" itself[69], and the reaction of the judges of the Superior Court as a result of the style of the presentation. The presentation of the second version of the sentencing guidelines was done in three parts. First, Judge Ronan, Chairman of the Committee on Probation and Parole described the sentencing guidelines project and explained how the guidelines (and the specific factors in them) had been derived after an analysis of the past sentencing decisions of the court. Judge William G. Young, a justice of the Superior Court, then presented a illustration of the use of the guidelines in a particular case. Finally, Mr. Donald Moran, Assistant Supervisor of the Superior Court Probation Department then gave a short description of the role of the probation department in the preparation of the sentencing guidelines for judges' use on a daily basis, and in the monitoring and future revision of the guidelines. Throughout, judges from the "audience" repeatedly interrupted the presenters (especially Judge Ronan) with questions about particular aspects of the guidelines.

It is our observation here that, in light of the previous criticisms of the initial sentencing guidelines model proposed, both skill and organization in the presentation of these three topic areas were bound to be important in order to convince judges, especially those who had been skeptical in the past, of the adequacy of the second guidelines model. Unfortunately, the presentation of the Massachusetts sentencing guidelines lacked both of those elements. The presentation of the history and activities of the sentencing guidelines project did not fully summarize either the research findings, or the policy decisions, made by the Committee on Probation and Parole and the guidelines project staff. (In fact, it was noted during this section of the presentation

that the Committee did not make policy when constructing the sentencing guidelines though, as is evident from our earlier discussion of policy issues, this was blatantly not the case.) Neither was this material made available to the judges in written form. The end result of not informing the judges adequately about the research and policy issues involved in the construction of the sentencing guidelines was that the judges were put in the position of examining the model for flaws; in a session including over forty judges, a considerable amount of criticism over the content of the guidelines was generated.

The major flaw of the presentation, in our opinion, was that a large portion of the "public relations" work of teaching the judges about the sentencing guidelines could and should have been done in advance of this meeting; and it was not. (Compare the description we have just provided of the formal introduction process in Massachusetts with that of the presentation of the Minnesota guidelines discussed Chapter 9.) And since the judges had not had the opportunity to air their views (and criticisms) at a date earlier than that of the presentation, those views were expressed -- often forcibly -- when the guidelines were introduced.

In fairness to the Massachusetts sentencing guidelines project and the Committee, some changes were made in the second presentation which should have made it easier for the judges to accept the guidelines. First, a judge, rather than a member of the research team, summarized the research and the guidelines content; this change from the presentation of the first guidelines model appears to have been made so that the judges would realize that there was some judicial input in the guidelines development process, and that the guidelines were not the product of statistical research void of any judicial interpretation. Secondly, a judge (once again, rather than the research staff) explained the operation of the guidelines -- and he did so by using a particular case example. This procedure should have emphasized the fact that some judges had tried to use the guidelines for real life cases and had found that the guidelines did produce sentences that they thought were appropriate. The final positive aspect of the presentation was that the Committee had obviously anticipated that the judges might respond to the concept with the question of "How, given current court caseloads, were judges to complete all of the required forms for all of the cases that they sentenced?" The answer to this question was given to the judges at the conference by Mr. Donald Moran, who explained that the Superior Court Probation Department would take responsibility for much of the necessary paperwork involved in the guidelines' daily use. These three aspects of the guidelines presentation certainly were beneficial to some

degree; however, the benefit was not enough to offset the judges' criticisms of the guidelines' content.

We noted above that the judges criticized the initial Massachusetts guidelines because those guidelines might generate a legislative "backlash", included factors that were not scored to the satisfaction of the judges, and provided separate (less severe) guidelines for plea cases. The second version of the guidelines included the following changes:[70]

1. The seriousness of the current and prior offenses would be determined on the basis of the statutory maximum sentence allowed for each offense;

2. Psychological injury would be assigned one point on the injury scale included in the guidelines; and

3. The guidelines would be applicable to trial cases, with optional consultation by the judge in plea cases, and separate sentence length scales would not be allowed for plea cases.

These three changes in the content of the guidelines appeared to please the judges of the Superior Court, for those judges believed that the changes made the final model a good deal more objective and thus less likely to be a cause of challenge, either in court, or in the political or legislative arenas. Several new controversial issues were raised, however; and these issues were not fully addressed by either the Committee or the research staff at this session. Included among those new issues raised were: What effect would the sentencing guidelines have on the function of the Appellate Division of the Superior Court? Why did the guidelines not include prior misdemeanor offenses in the calculation of prior record seriousness (especially where those offenses had resulted in some time served in incarceration)? How could the guidelines account for administrative (correctional) decisions made after sentence was imposed, that would affect the length of sentence? Would the guidelines sentence concur with the district attorney's recommendation? Had the effects of race been eliminated from the sentencing guidelines model? And, once again, a major concern of the judges that had been raised at the initial guidelines session was evident: Would the guidelines result in a major legislative "backlash" of mandatory minimum sentences for new offenses?

The Committee offered only partial answers (at best) to most of these questions.[71] They noted that the guidelines were not intended to affect the Appellate Division's function at all. They replied that misdemeanor offenses were not included in the guidelines because the project staff had not

had adequate statistical information to do this. They noted that the guidelines could not account for the decisions of correctional authorities after the sentence was imposed, but neither could judges do so at the present time, so the use of the guidelines would be irrelevant to that point, though the guidelines would make for a more informed decision about sentence length generally. The guidelines also would not necessarily concur with the district attorney's recommendation -- the Committee did not have the information that would have enabled them to construct guidelines that would do so; however, the guidelines were to be optional for use in plea cases, and it is in those cases (it was said) that most district attorneys' recommendations are important. The studies done by the Committee and the project staff (which we discussed above) indicated that race was not an apparent factor in sentences in Massachusetts; thus, it was believed that the effect of race was not present in the sentencing guidelines.[72]

On the final issue raised -- that of the potential for legislative backlash caused by the guidelines -- the Committee noted that some steps had been taken with the legislature to forestall such an occurrence. Correspondence between the representatives of the legislative committees concerned with sentencing reform [73] and the judges of the Superior Court dates from 18 July 1979 (prior to the introduction of the first guidelines model) to 7 April 1980 (just prior to the introduction of the final guidelines model).[74] In a series of letters, the concerns of both of these bodies -- the legislature and the judiciary -- about the Massachusetts sentencing guidelines were spelled out in considerable detail. The general legislative concerns were three [75] in number:

> ....As the determination of a sentencing policy is the prerogative of the General Court and any attempt to substitute the collective judgement of the Court for the discretion that now rests with the individual sentencing judge under the indeterminate sentencing system can only be viewed as a fundamentally legislative act and not one, therefore, that may appropriately be undertaken by the judiciary.
>
> ....The same Constitutional flaw appears with regard to the Courts' classification of offenses by seriousness and determination of penalties by means of a point system. Again, it is the prerogative of the General Court to set penalties for crimes.
>
> ....For a judge to attempt to determine how long a prisoner will be incarcerated would appear to violate both Executive and Legislative prerogatives under the

> separation of powers doctrine. The General Court has established and authorized the Parole Board to regulate the release of prisoners within the parameters set by statute. The judiciary should not attempt to circumvent or out-guess the proper exercise of discretion on the part of the Parole Board.[76]

The Chairman of the Committee on Probation and Parole, Judge John T. Ronan, responded to the concerns of the legislature listed above by noting the following points:

> Guidelines are a reflection of current sentencing practices, not the judicial creation of a sentencing policy.
>
> The guidelines project has no constitutional flaw in that it does not replace the present discretion of the individual judge....Since the sentencing judge is free to impose any sentence that is legislatively permissible and required only to make explicit or articulate the norms and considerations utilized by him in the sentencing process....
>
> The guideline project has no constitutional flaw in that it does not intrude upon the prerogative of the General Courts to set penalties for crimes.
>
> The use of real time to be served does not violate any legislative or executive prerogative....Superior Court Judges do appear to sentence with an eye firmly fixed on real time (the so-called "flat time" concept).
>
> There are no departures from current sentencing practices which would be occasioned by utilization of a guideline.
>
> ....That those who have previously said that difference is not between the judges who sentence but the difference is between defendants and their characteristics were more correct than many of us believed. This finding means that our practices have been uniform and rational in our collective approach to sentencing. It means that disparity in result is in a large measure earned by the offender. [77]

It is apparent from the points noted above that, while there was considerable divergence in the two groups' views of the guidelines' content and use, attempts had been made to discuss the differences in a rational fashion, and to resolve some of these issues prior to the introduction of the second version of the Massachusetts sentencing guidelines. The

judges of the Superior Court were informed of the discussions that were in process between the legislative committee and the Committee on Probation and Parole, and the fact that such discussions were in process did appear at the second presentation of the guidelines to reassure a number of the judges that their actions regarding adoption of the sentencing guidelines would not be construed as an intended subversion of the legislative function. However, not all of the judges were of this belief; thus, some tension over this aspect of the guidelines was still evident.

The decision to continue testing of the guidelines was made at a session of the judges during the following day of the conference. In May 1981 the judges voted to formally adopt the sentencing guidelines (unchanged) for use. During the interim year, the testing of the model apparently convinced the judges skeptical of the possible legislative retaliatory potential of the guidelines that enactment of further mandatory minimum sentences would not necessarily be a function of the adoption of sentencing guidelines.

The Probation Department and Sentencing Guidelines

The Superior Court Probation Department was assigned the task of supervising the operation of the sentencing guidelines on a day-to-day basis. (At about this time, the research staff of the Massachusetts sentencing guidelines project disbanded due to a lack of funds -- thus, the Assistant Supervisor of the Superior Court Probation Department took over the guidelines monitoring function done earlier during the test period only by the project's research staff.) Several regional training sessions were held during May 1980 to familiarize local probation staff throughout the state with the content of the sentencing guidelines and to instruct the probation officers in completing certain sections of the guidelines for the judges. Probation staff were to complete the guidelines sections about the defendant's current and prior record, while the judges themselves would score the degree of injury to the victim and the number of times that a weapon was used.

The reaction of the majority of the probation officers at the training session that we attended[78] was one of positive criticism. The officers had been informed of the judges' intention to use the guidelines and had been urged to assist the judges willingly so as to "help re-establish our role as an aid to the court in sentencing" and to "combat our diminishing role in the sentencing process."[79] A number of changes in the probation department's activities relative to sentencing had taken place fairly recent to the introduction of sentencing guidelines[80] and these changes, for the most

part had reduced the paperwork necessary for the probation department's work in respect to sentencing. Thus, there was evident at this time a concern among probation officers that the reduction in their duties might in turn lead to a cutback in the staffing of the probation department. These concerns may have been an important factor contributing to the officers' positive assessment and criticism of the guidelines model.

At the training session where probation officers were first introduced to sentencing guidelines, the major criticisms of the model were limited to questions about how certain of the items would be scored. Officers noted that the guidelines forms were to be completed for all cases, although they would not apply to plea cases unless the judge decided to consult them voluntarily; yet for most plea cases, the probation officer would not be aware of the exact charges to be listed on the form until the plea was entered. Since, in many counties, the probation department did not have close working relationships with the district attorney's offie, probation officers might not be able to collect the information necessary about the current offense that would be required to complete the forms until the last minute -- producing a large backlog of work in court at the time of sentence.

Secondly, the officers noted that if all of the charges are used as a basis for determining the current offense score, and some of those charges are filed by the judges to save paperwork, the guidelines sentence would technically be higher than it would be if those charges were not to be counted at all. In other words, the sentence would be above the guidelines range, not because it was inappropriate, but because the judge decided to save the court clerk's office some time by filing some of the charges. [81]

In general, the concerns of the probation department at this point in time involved technical considerations that the officers believed that they would encounter in the daily use of the forms, rather than the philosophical issues that the guidelines' use might raise. Admittedly, concern over philosophical issues on the probation officers' part would have been pointless anyway; the judges had voted to use the guidelines, and since the probation department was a division of the court, the officers were responsible for assisting the judiciary -- and assistance would, because of the judges' decision, include the preparation of sentencing guidelines for the judges of the court.

Using Sentencing Guidelines: The Reactions of Judges, Prosecutors, Probation, and Defense

The vote to continue testing of the second version of the Massachusetts sentencing guidelines took place in May 1980. By June 1980, probation department staff had been trained in the use of the guidelines, and the judges were completing guidelines forms in court for particular cases. During June and July 1980, staff of the Evaluation of Statewide Sentencing Guidelines project again visited four Massachusetts counties and interviewed judges, probation officers, prosecution and defense counsel about the content of the final guidelines and the use of those guidelines in actual cases.[82] Although judges and probation staff had only been using the sentencing guidelines for about one month, and defense counsel and prosecutors had known about the guidelines for an even shorter period of time, the analyses of the guidelines that each of these parties related to us during those interviews were quite detailed, and should be reported briefly here.

Judicial criticisms of the sentencing guidelines can be generally classified into one of three categories -- philosophical opposition to the concept, mechanical problems involved in use, and criticisms of the content of the Massachusetts model. Far fewer judges than would have been expected, given the comments made at the judicial conference where the guidelines were introduced, were adamantly opposed to the guidelines by the time we interviewed them. Apparently, use of the model in actual cases had somewhat tempered some judges initial dislike of the guidelines. Instead, the focus now seemed to be with what specifically was wrong with either the general structure or the specific factors included in (or excluded from) the guidelines -- the "content" criticisms that we noted above.

Most of the judges that we were able to interview had been able to use the guidelines in a small number of cases prior to our discussions with them. Thus, the judges' criticisms of the structure and content of the guidelines were based at least in part on their experiences with the guidelines in specific cases. The criticisms of the structure of the guidelines most often noted by judges included the following:

1. The guidelines range of 50 percent is extremely wide; as a result almost every realistic sentence is within the guidelines.

2. The guidelines sentence scale makes it very difficult to have a case result in an "out" sentence.[83] This may have important ramifications for the guilty plea

disposition: If the guidelines are used in plea cases, fewer offenders will want to plead guilty because the guidelines will not usually recommend an "out" sentence; if the guidelines are not used in plea cases, more offenders may plead to avoid the probable "in" sentence that the guidelines would stipulate after trial.

3. A number of judges noted that even though the guidelines attempt to use an estimate of "real time" in determining the appropriate sentence, there is no guarantee that the estimate is correct. Since the Department of Corrections may adjust its rules regarding parole eligibility and release at any time, it might have been more profitable for the project (it was argued) to have developed parole decision-making guidelines for the parole board, than sentence guidelines that may or may not be accurate.

Before we go on to note the criticisms that the judges had to offer regarding the specific items included in the Massachusetts sentencing guidelines, we should note that the 50 percent range provided in the guidelines was mentioned by some judges as an advantage of the model, in contrast to those judges who saw the same range as a criticism of the model. A number of judges liked the 50 percent spread just because most of their sentences were within that range; while other judges noted that if all of their sentences were within the range anyway, the guidelines were not accomplishing much by way of reducing variability in sentences.[84]

The criticisms of the content of the sentencing guidelines generally involved the exclusion of certain items of information from the calculation of the guidelines sentence. Among the comments made by judges about the guidelines content were the following:

1. The classification of offenses according to statutory maximum sentences did not allow the judges to differentiate between levels of seriousness for the same crime type. For an example, an assault with a dangerous weapon where the weapon was a "shod foot" would receive the same seriousness score as an assault with a dangerous weapon that involved a knife or a gun.[85]

2. That misdemeanor offerses were not calculated into the prior record seriousness score was criticized by judges for two reasons: the seriousness score for prior offenses would not reflect the actual seriousness of the offender's prior record, and patterns of prior criminality would be ignored by the guidelines.

3. The guidelines did not formally take into account any mitigating factors, a number of which would be relevant considerations when sentencing. Among those mitigating factors mentioned by the judges were: age, employment, unstable home life, etc.

4. The calculation of the current offense score was thought likely to be extremely confused because of the common practice of "filing" charges without a formal disposition. (This point was noted above as a criticism offered by probation officers about the model; in fact, this criticism did turn out to be the major problem with the use of the Massachusetts guidelines.)[86]

5. A number of judges noted that the guidelines may include a racial discrimination factor since the model places substantial weight on the prior record of the offender and minority offenders usually have longer prior records.

6. A statement of specific reasons for departure from the guidelines might result in litigation as to the constitutionality of those reasons in both the Appellate Division of the Superior Court and in the Appeals Court.[87]

Despite what may appear to be numerous objections to the content of the sentencing guidelines (given the sample listed above), judges did not seem to be as worried about the mechanical problems that might be encountered in using the sentencing guidelines as they had earlier been. The judges we spoke with acknowledged that they consulted the guidelines; however, the consultation was done after they had already decided on what would be the appropriate sentence for a particular case, rather than using the guidelines as a tool to help them to derive the sentence in the first place. (That the majority of judges used the sentencing guidelines in this manner may pose a constraint to the ability of the sentencing guidelines to function as a decision-making tool, as was intended by the guidelines' developers.) Consulting the guidelines, in any event, did not appear to bother any of the judges; we suspect that this was the case in part because the judges themselves did not have to spend the time computing the guidelines scores for current and prior offenses.[88] And the judges did not seem to believe that the use of the guidelines would pose any mechanical problems for the probation officers responsible for completing the forms for them.

The judges most recently appointed to the Court appeared to be more supportive of the concept than their brethren who had been sitting on the Court for a number of years; in fact,

quite a few of the judges who had been appointed to the Court for quite some time suggested that the guidelines might be especially useful for more recent appointees (rather than for themselves). They argued that the guidelines would, by describing past sentencing practices, provide newer judges with a general overview of sentences that would be usually appropriate for certain types of cases. This sentiment was repeatedly expressed by almost all of the judges that we interviewed, even by those who themselves did not support the general use of the sentencing guidelines.

It will be recalled, from our earlier discussion of the introduction of the sentencing guidelines at the judicial conference, that the major criticism of the guidelines by the judges was that the guidelines interfered in some way with what was a legislative prerogative to determine the classification of crimes and penalties. A subsidary but related criticism of the guidelines is that, since the determination of crimes and penalties is a legislative function, only the legislature can curtail the discretion of a sentencing judge in any way. We had expected, prior to our detailed interviews with judges throughout the state after the introduction and use of guidelines that these two issues would continue to be the major issues concerning the judiciary. This proved not to be the case -- and speculation about why it was not provides a great deal of insight into the reception of the Massachusetts sentencing guidelines.

Two facts about the Massachusetts sentencing guidelines, often noted by the judges we interviewed, are key to understanding why the judges' initial dislike of the model (on the grounds that it infringed upon the legislature, and on judges' sentencing discretion) had become less adamant with time. The first is that the guidelines were, in essence, voluntary (since a judge did not have to adhere to the guidelines sentence), and not applicable to plea cases; the second is that the guidelines range was so wide as to make deviations from the model the exception rather than the norm in almost every case. These appeared to be the major reasons that judges did not seem as concerned (either positively or negatively) about the sentencing guidelines after they had used them, as they had before. In other words, the judges had discovered that the guidelines really did not affect what they were currently doing when sentencing. First, the applicability of the guidelines was limited to only a small number of trial cases on a mandatory basis and, secondly, the guidelines provided an extremely large range from which they were to choose a sentence for that small number of cases. As one judge put it: "The ranges are very broad...given a fifty percent range up and down, I'd be intrigued to see the case outside the guidelines; it will be very much of a rarity."

These aspects of the structure of the Massachusetts sentencing guidelines did quite a bit to lower the hostility of the judges toward the concept and the model's use.

The other actors in the criminal justice system who would be involved in the use of the sentencing guidelines on a daily basis did not seem to share the judiciary's general attitude toward the concept. Probation staff, initially receptive to the sentencing guidelines (as we noted above), came to think of the guidelines as an additional burden that had been imposed on them. Probation department staff, by and large, noted that although the guidelines might assist judges in determining equitable sentences, the procedures to be used in implementing the guidelines invariably caused tremendous management problems for the probation department staff.[88] The specific criticisms that were noted by the probation officers, and their supervisors, in the four counties where our staff conducted interviews included the following:

1. A number of computational issues were apparent in using the guidelines. Probation officers were often not able to obtain an offender's out-of-state charges for inclusion in the prior offenses, and often were not able to judge the appropriate seriousness classification for those offenses.

2. The question of how filed charges should, or should not, be scored in the current and prior offense seriousness items was never adequately resolved; thus, some probation officers included filed charges (and previous dispositions of that stripe) in the seriousness scores and others did not.

3. The completion of guidelines forms for all cases, including plea cases, was a waste of their time since the overwhelming majority of cases pleaded guilty and thus were not subject to the guidelines unless the judge decided to use them. (Also, as most cases were plea cases, it was often not possible to obtain the necessary information about the case from the district attorney's office so that the guidelines form could be completed.)[89]

4. The transfer of the offenders' prior records from the court records to the guidelines forms took a tremendous amount of time as well; a procedure should have been advised whereby all of the information about the defendants' prior records would not have to be rewritten on to another form, but merely scored for the guidelines purposes.

5. Judges often asked probation staff to score injury and weapon use, contrary to the procedure directed by the guidelines. As these assessments were qualitative in nature, a number of probation officers did not want to be accountable for that scoring.

6. The tremendous amount of time that the guidelines required to be completed took away from the amount of time that probation officers throughout the state had to devote to the actual supervision of their clients.

Probation staff noted that a number of the problems mentioned above could have been resolved if adequate preparation for, and training in, the use of the guidelines had been considered. However, they noted that the training session run by the Massachusetts sentencing guidelines project did not deal with a number of the issues that they subsequently found to be a problem -- such as the classification of out-of-state charges, and filed charges and dispositions. Further, as the problems were encountered in the probation department's daily use of the guidelines forms, there was no uniform procedure for deciding how these problems should be handled -- the Massachusetts sentencing project had disbanded, and the questions that arose tended to be dealt with on a local level, often resulting in local differences in the methods used for scoring the sentencing guidelines' factors. In fact, in our discussions across the state, we found considerable differences in the inclusion or exclusion of certain items from the scoring of various guidelines factors. For example, one probation department might include filed charges in the current offense seriousness computation, while another department would not. Of course, at the time of our interviews, the use of the guidelines by all of the personnel that we interviewed was fairly novel; a number of the questions that were raised about the use of the guidelines may have been resolved in a uniform manner since the guidelines' early introduction. Information on this point, however, is not available to us at the present time.[90]

One additional criticism, though not made by a majority of the probation staff, was noted by several; and this describes an important feature of the Massachusetts sentencing guidelines. The criticism is that the guidelines make it very difficult for all but the least serious offender, who does not have any prior record, to receive probation. Although the criticism has important implications for the statement of a sentencing policy for the state of Massachusetts, it affects the probation departments' perceptions of the guidelines as well. Ironically, though probation staff are required to expend a great deal of their time (that might otherwise be

devoted to supervision) in completing the guidelines for every case, almost any usual case will be a sentence of incarceration, rather than probation. The probation staff who did notice this feature of the guidelines were anything but pleased with it at the time of our interviews.

District attorneys throughout the state had known a little about the development of the sentencing guidelines from the inception of the project, in part because their records had served as one data source that the project had used to compile information about current offenses. After the collection of data, however, the district attorneys were not kept informed about the projects' analyses or about the sentencing guidelines models (both the first and the second) that were developed. Neither the project, nor the Committee on Probation and Parole, asked for input from prosecutors into the sentencing guidelines development process, or presented the completed versions of the guidelines to the district attorneys (or to their statewide association).[91]

Thus, most of the district attorneys (and their assistants) throughout the state learned about the sentencing guidelines when the judges started using them in court. A number of prosecutors did not even learn about the guidelines then; judges, they thought, might have used the guidelines in court but not informed them (or anyone else) that they were doing so. Moreover, a number of prosecutors only learned about the judges' use of the sentencing guidelines at the time of our interviews with them in 1980.[92]

By far the most outraged of all of the groups affected by the use of the sentencing guidelines were defense counsel throughout Massachusetts. The development and subsequent use of the sentencing guidelines had not been made known to them at all.[93]

Indeed, it was only because of our interviews with defense counsel during the preceding year that many (if any) of them ever knew that the Court was considering the development of guidelines.

We will discuss the reactions of both of these groups -- district attorneys and defense counsel -- to the sentencing guidelines together in this section. A number of the points raised by both district attorneys and defense counsel about the guidelines are similar; and a number contrast significantly as a result of their contrasting occupations. Both the similarities and the differences will be more apparent if the discussion focusses on both at the same time.

Many of the general concerns of district attorneys and defense counsel mirrored those of the judges. Criticisms of the sentencing guidelines fell generally into the same three categories as those described by the judges that we interviewed: philosophical opposition to the guidelines concept, criticisms of the mechanics of using the guidelines, and disagreement with the factors included in the sentencing guidelines model. Of these three reasons, disagreement over the factors included in the sentencing guidelines and over the manner in which those items were scored accounted for the bulk of the defense and district attorneys' criticisms.

If one had to summarize the thoughts of each of these groups about the sentencing guidelines in a single sentence, it would be that while sentencing guidelines generally may be an acceptable addition to the sentencing process, these specific guidelines produced sentences that were too low for the satisfaction of the prosecution and too high for the satisfaction of the defense. This summary, however, is much too general and does not indicate what it is that is wrong with the sentencing guidelines so that they should produce unacceptable results for both groups. We asked both defense counsel and district attorneys what was either omitted from the guidelines that should have been included, or vice versa, and why the guidelines resulted in the sentences that they did. Both groups responded by noting that the guidelines, as they now stand, are much too simplistic to capture all of the aspects of a case that are truly relevant to sentencing. Agreement was evident among the two groups about certain perceived flaws of the sentencing guidelines:

1. The guidelines should not use the statutory maximum penalty to score the seriousness of the offense. The statutory maximums often do not match the seriousness of the offense and, further, use of the statutory maximum for all offenses of the same type will not point out the differences in seriousness that are present for similar offenses.

2. The guidelines should include misdemeanor and juvenile offenses in the calculation of prior record seriousness. Although defense counsel generally liked the idea that such items were not included in the guidelines, they noted that non-inclusion was just not realistic. Even defense counsel, they noted, make some judgements about the offender on the basis of the pattern, and extent, of the prior record.

3. Some element of "decay" should have been included in the guidelines where an offender's only prior convictions were very old.

4. The guidelines should include some mitigating factors, including age, employment, relationship with victim, lack of injury (rather than presence of injury), etc.

5. The guidelines should have made provision for certain aggravating factors, such as alcohol and drug dependency, amount of drugs, provisions for white collar crimes, and potential for injury (rather than actual injury caused).

6. The guidelines should specify some reasons for possible deviations, so that judges will have less reluctance to specify particular reasons for either higher or lower sentences.

7. The weights specified by the guidelines for particular factors should be changed. Specifically, district attorneys believed that the weight assigned to prior offenses was too low; defense counsel believed that the weights assigned to injury and weapon use were too high.

8. Both groups noted that the guidelines made it very difficult for a person to obtain a non-incarcerative sentence, even if the offender was a first offender. While the groups may have differed in their estimate of whether or not this feature of the guidelines was desirable, they both acknowledged that it was not realistic; thus, the guidelines sentences should be changed to make probation an acceptable disposition for more of the casees.

We believe that what is most important about the specific observations offered by defense and district attorneys, including those listed above, is that they clearly point out the fact that sentences can be explained, if the method of examining them is not too simplistic. Sentencing is a complicated affair, and the many different sentences that result are evidently based on differences among offenders; while the differences among sentences may not always be seen by opposing parties as appropriate in a given specific case, the opposing parties are also clearly not willing to negate completely the consideration of those factors that will make for differences in sentences.

The defense counsel and prosecution staff that we interviewed also had a number of observations about the ultimate effect of the sentencing guidelines on the use of probation, the incidence of trials, the sentence recommendations of counsel, the populations of prisons, and the probability of appeal.

Counsel noted that the guidelines do not place much weight on the input of the probation officer into the sentencing process. Assuming that probation officers will tell the judge the guideline sentence, and that the guideline sentence will not usually be non-incarcerative under the present model, the independent recommendations of probation officers regarding suitability for probation will be lost. This was seen as unfortunate: often, it was noted, it is the probation officer who knows the most about the defendant and his past criminal conduct.

Defense and prosecution counsel had somewhat similar views to offer about the impact of the sentencing guidelines on the number of trials as well. Their relative lack of experience with the sentences produced by the guidelines in specific cases made most of their comments on this topic a matter of speculation, yet the speculation has its merit. Defense counsel noted that the guidelines might actually be a good innovation if they produced reasonable sentences for certain types of offenders. If the sentences produced by the guidelines were seen as not too harsh, then a benefit of the guidelines would be that "they would allow a defendant to opt for trial and still be able to avoid the worst excesses of trial sentences." On the other hand, if the guidelines did call for extremely severe sentences after trial, the model would be "a phenomenal plea-bargaining tool" whereby defendants would be under tremendous pressure to plead to avoid the high guidelines sentence after trial. Prosecution staff agreed with defense counsel's assessments on this matter.

These adversary parties also acknowledged that use of the sentencing guidelines could change the nature of the sentence recommendation process. It was believed that the guidelines sentence for a particular case could become the reference point in the future for sentence recommendations and for plea negotiations. Additionally, prosecutors noted that the guidelines sentence could become the upper limit to a sentence recommendation in a plea case -- thus, acknowledging that the defendant was getting a lower sentence after a plea, but not a lot lower.

Defense counsel expressed the fear, as well, that the use of the guidelines could change the nature of the sentencing hearing. Rather than focussing the sentencing hearing on the aspects of the particular case and defendant that should bear on the sentence, counsel believed that the focus in future hearings would be concerned with points assigned to various factors, and methods of scoring. In other words, the hearing would center on the guidelines sentence, rather than on the defendant. And in view of the fact that both parties

acknowledged that the guidelines could be manipulated to result in high sentences depending on how certain items were scored, the expectation that the focus of the sentencing hearing would change seems to us reasonable.

As the focus of the sentencing hearing would change from the defendant to the guidelines, so too would the focus of future sentence appeals to the Appellate Division of the Superior Court. Counsel expressed the additional fear that judges would be reluctant mostly to specify reasons of mitigation for going outside the guidelines sentence on the low side. Some mitigating factors, they noted, just don't look good on paper -- in large part because they can come back to haunt the judge later on. An example would be where a judge stated as a reason for a lower sentence that the "offender probably won't do it again" and the offender does do it again. Defense counsel believed that judges will not want to take the risk of mitigating the sentence in such cases in the future. On the other hand, aggravating a sentence because the offender "might do it again" would be much more palatable to the public; thus it was thought that judges will probably not be as reluctant to aggravate a sentence as to mitigate it.

Finally, both district attorneys and defense counsel observed that the manner in which the guidelines were ultimately used, and the types of sentences that the guidelines produce could dramatically affect the population (and thus the operation) of the Massachusetts prison system. If the basic structure of the guidelines was found to produce longer sentences, and if more trials resulted and longer sentences were imposed, then prison populations would rise as a result. Also, several district attorneys noted to us that the sentencing guidelines compute "real time" estimating that all offenders are actually released at their first parole eligibility; but this is not always the case. Should the number of those offenders estimated for release at first eligibility far outweigh those who are actually released at that time, again, prison populations will rise. (Of course, this projection, also assumes that some offenders are not released at the discretion of the Department of Corrections before their first parole eligibility, which we also know not to be the case.)

These views of defense counsel and district attorneys about sentencing guidelines can be said to be "case-oriented". Both groups of attorneys were concerned with how the guidelines would affect each specific case, and the strategies they use to handle those cases, with which they deal on a day-to-day basis. In contrast, we would label the concerns of the judiciary as "sentence-oriented". The judges, after determining that the guidelines would probably not drastically

affect their sentences, lost much of their initial concern over them. Finally, probation department views can be summarized as "workload oriented"; the concerns of probation staff overwhelmingly involved how that staff would produce the additional paperwork that the guidelines would require. (It is surprising that the concerns of probation could not also be labelled "case-oriented", as guidelines appear at first blush to affect the future quantity of probation sentences; but this was not found to be the case.) In the event, it will require a great deal of additional information on the long term use of guidelines in the state of Massachusetts to determine which of the concerns raised by the parties involved in the use of guidelines will prove to be the most accurate in this estimate of the effects of sentencing guidelines.

Notes to Chapter 7

[1] Chief Justice Robert M. Bonin of the Superior Court submitted the project proposal on 6 September 1977 to the Massachusetts Committee on Criminal Justice. The Committee on Probation and Parole was assigned oversight of the project by Chief Justice James P. Lynch (who replaced Chief Justice Robert M. Bonin.) Members of that Committee were: Justices John T. Ronan, Chairman; John F. Moriarty; Robert S. Prince; Paul G. Garrity; Robert V. Mulkern; Charles R. Alberti; and James P. Donohue. The project was directed by Michael Hutner; and project staff included: Thomas J. Marx, Statistical Consultant; Amy Craddock, Research Supervisor; and Cathryn M. Chadwick, Administrative Assistant.

[2] The Massachusetts Sentencing Guidelines Project press release, dated 11 September 1978, mentioned that "the study may indicate that certain variables (possibly bail status or court delay) currently have an influence which the Superior Court may decide should not be reflected in the guidelines."

[3] It is unclear exactly how long the initial test period lasted because the testing of the guidelines was on a totally informal basis. Judges could consult the guidelines if and when they desired to do so, and interviews conducted by our staff with individual judges indicated that some judges "tested out" the guidelines on only a small number of cases while others actually consulted the guidelines in every case they sentenced over the several months of the test period. While feedback about the guidelines accuracy in particular cases was requested by the Massachusetts project, such feedback is believed by the members of our staff to have been very informal, consisting mostly of verbal exchanges between the judges and members of the Committee on Probation and Parole, and with project staff.

[4] Most researchers would not quarrel with a methodological critique of a statistical model that claimed to replicate a decision-making practice that was unaltered by policy decisions. However, a number of those same persons might say that, since the guidelines claim to accurately represent sentencing as a pure statistical model, methodological criticisms are a waste of time. We do not agree. The initial statistical model -- that represents sentencing before policy modifications, for example -- might be grossly inadequate because of methodological flaws. In addition, since policy altered statistical decision-making tools, like those developed in Massachusetts, are likely to be replicated in other jurisdictions, the limitations of such models should be fully investigated.

[5] This is not to say that the information that the Massachusetts project did collect was wrong in some way; although there are some usual problems with the data they collected (such as, question and/or response ambiguities), the major problem is that they did not collect the information that was most important to sentences at all.

[6] The pre-trial investigation report also provided judges with background information about the offender which they used in deciding upon sentence. Since, as in most jurisdictions, the majority of cases are disposed of by a guilty plea, and since Massachusetts judges rotate between the courts of several counties during the year, sentence was usually pronounced immediately after the plea was entered or the verdict was reached. Thus, the investigation report was prepared before trial (or the entering of a guilty plea) so that the judge would have the necessary information about the offender available for immediate sentence determination. The pre-trial investigation had a number of drawbacks, however, not the least of which was that the probation officer compiling the report could not ask the offender about the offense and could not elicit signs of remorse from the offender for having committed the offense because, at the time that the report was being prepared, the offender had not yet been adjudicated guilty of any crime. Pre-trial investigations were dispensed with during 1980 and pre-sentence investigations (of the traditional sort) were substituted in their place.

[7] The other contenders to be the data base for the selection of a case sample were the district attorney records and the records of the court clerk in each county. Presumably, the court clerk's records duplicated the probation office records with one exception -- the probation department records also contained some detailed personal information about the offender. The district attorney's files were considered by the Massachusetts project to be incomplete and, as we did not study those files ourselves, we must assume that this was the case. Even though these data sources were not used as the data base for sample selection, however, both were used to collect specific types of information that was not available in probation department records. The district attorney's files were used to obtain specific information about the current offense for each sample case, and information about the offenders' prior offenses was obtained from the court clerk's files.

[8] The project memorandum that discussed the sample selection procedures stated that the sample was "selected randomly from the list and procedures will be established to insure adequate representation by judge, court, location, and

offense." (Memorandum to the Committee on Probation and Parole dated 12 December 1978.) Presumably, the technique of random sampling by itself provided these assurances. At any rate, staff of this project were not able to ascertain, beyond the content of this and other memos, exactly how the random sample was selected. At one meeting with staff of the Massachusetts project, a member of that project mentioned that every third case had been selected -- and such a procedure, whether or not it would have significantly affected the data analysis results, is not a random sampling procedure. We must take the Massachusetts project at their word that the sample was truly random.

[9] It is also not known how the figure of "one-third of the yearly disposition population" was decided upon as the adequate number of cases. Presumably, the figure was chosen after a review of the sentencing guidelines development material provided to that project by Kress as we noted in Chapter 3.

[10] We doubt that the guidelines were adversely affected by the exclusion of all four of these counties. Although there is some seasonal variation in their disposition caseloads (for all of these counties are resort areas), in general their caseloads are too small to have made much of a difference in a statewide sentencing model.

[11] Massachusetts sentencing guidelines project memorandum dated 12 December 1978. In a project memorandum, titled Development of the Coding Manual, dated January 1980, Amy Craddock, the research director, noted that the choice of items included in the data collection manual was influenced largely by materials from the Michigan Felony Sentencing Project, the New Jersey Sentencing Guidelines Project, and the report Developing Sentencing Guidelines (Kress, et al., 1979). It was also noted in that memorandum that: "Unfortunately there was little input from the judiciary, other than a general list of areas considered important for investigation." (page 1).

[12] The 30 January 1979 progress report of the Massachusetts guidelines project stated that the training of data collectors was to begin on 7 February 1979. After an initial two days of class instruction, data collectors were to spend eight days coding cases in Middlesex County under the close supervision of the project's research director, Amy Craddock. After that, data collectors would be dispersed to counties throughout the state, and one supervisor would be appointed to perform code-checking and spot checks in each county. Code-checking would also be done by other data collectors routinely for some of the information items.

Whenever unforeseen coding problems developed, supervisors were told to check with the research director before any coding decision was made. General instructions about how case information should be coded were prepared on 8 February 1979. We also were informed that all of the training information, as well as all subsequent coding decisions made in unusual cases, were codified into the final data collection manual that was distributed to judges (together with the first guidelines model developed) in November 1979.

[13] The Massachusetts data collection instrument included eight items that related to the district attorney's sentence recommendation:

1) Is the District Attorney's recommendation the same as the sentence?
2) Recommended disposition (place and suspension)
3) Recommended costs
4) Recommended amount of Fine/Court Costs/Restitution
5) Recommended sentence length - minimum
6) Recommended sentence length - maximum
7) Recommended probation length
8) Relationship of recommended disposition to previous sentence or previous recommended disposition

Our project's interviews with district attorneys, as well as with defense counsel, indicated that these items by themselves will not capture the essence of a recommendation in a plea disposition case in any event. We learned from these interviews that recommended sentences, in cases where there has been a plea disposition, can take any of the following forms:

> 1) Prosecutor and defense agree as to the sentence recommendation openly in court and both parties state the same sentence recommendation;
>
> 2) Either party agrees not to offer a sentence recommendation at all -- thus letting the recommendation of the opposing party stand as the sentence recommendation;
>
> 3) Prosecutor and defense agree as to the place of incarceration included in the sentence recommendation -- but either or both parties agree not to make a specific recommendation as to the length of the sentence. (Of course, the place of incarceration itself will put <u>some</u> limits on the choice of the length of sentence as we discussed earlier in Chapter 6.)

As is evident from the forms of agreed recommendations that we list above, very little applicable information about sentence recommendations was collected by the Massachusetts sentencing guidelines project.

[14] Information about the district attorney's sentence recommendation was recorded for only 506 of 1,440 defendants -- or for about 35 percent of the sample. The percentage of recorded recommendations is even lower if one considers charges instead of cases; the district attorney's recommendation was recorded for only 1,001 of 4,646 charges -- or for about 20 percent of the total number of charges.

[15] The eight items that we have listed in note 13 above were the only items of information pertaining to sentence recommendations offered to the court that were collected by the guidelines project. No attempt at all was made to capture the defense counsel's sentence recommendation for any case. Of course, in all probability that information would not have been recorded in any of the data sources that the project had decided to use (such as, the probation file, the district attorney's file, or the court clerk's file). Considering the stated importance of this piece of information, however, we believe that some attempt should have been made to obtain it from another source, or an alternate source of collecting data (such as through observation of sentencing hearings) should have been examined at the onset of the research design. We discuss this last point in more detail in the text and notes 25 through 28 below.

[16] A number of prosecutors informed our evaluation project staff during interviews that, very often, the sentence recommendation offered by a prosecutor in court is not, as a matter of routine, recorded in the case file.

[17] We were told by staff of the Massachusetts sentencing guidelines project that observation of sentencing hearings was not urged by the judges of the Committee on Probation and Parole until a late stage of the data analyses. We wish to note that, to our knowledge, such permission was not denied by that committee either. In a summary of an interview held on 30 May 1979, it was noted that one of the judges on that committee actually suggested that all of the Massachusetts project staff attend court sessions.

[18] The nine most frequently appearing offenses in the preliminary Massachusetts data were: armed robbery, assault and battery with a dangerous weapon, larceny, possession or distribution of a class A substance, firearm violation, breaking and entering at night, unarmed robbery, rape, and assault and battery. Unfortunately, the memorandum that lists

these results is undated, so we cannot be sure of exactly when it was distributed to the members of the Committee on Probation and Parole. Presumably, it was after some data collection had already begun, however, since the data collection began during February 1979 shortly after the preliminary data were collected.

[19] Massachusetts sentencing guidelines project memorandum titled, Data Collectors Training, General Instructions and Questions, dated 8 February 1979, page 1.

[20] Memorandum to Sentencing Guidelines Project data collectors and supervisors from Amy Craddock, research supervisor of that project, dated 21 March 1979.

[21] The 8 February 1979 memorandum states:

> There will be other reasons for disqualification. These will be left to the supervisor's discretion. The reasons for disqualification should be clearly stated on the back of the field card.

Our project staff did not have the opportunity to discuss the other disqualification reasons that were decided upon by the Massachusetts project; however, some figures -- though they appear to be incomplete -- were released by that project and these are discussed further below.

[22] Data Collectors Training, General Instructions and Questions, 8 February 1979.

[23] The figures presented in Table 7.1 were taken from a memorandum to the Committee on Probation and Parole from Amy Craddock, project research supervisor, on the topic of probation department records, dated 29 January 1980.

[24] It should be reemphasized here, although it is noted on the table, that the sample size obtained by adding together the numbers of sample cases from this table do not agree in two instances with the county sample sizes found in the data sent to us by the Massachusetts sentencing guidelines project. Moreover, we have no idea why the numbers of cases found in our data for the counties of Suffolk and Middlesex differ from those shown in Table 7.1. Fortunately, the differences between the two sets of statistics are minor and probably do not really matter.

[25] The additional reasons for alarm are several. First, the lack of updated pre-trial investigation reports might suggest that those reports (and the information that they contain) are not the primary information source that

judges consult when sentencing. If they were important to judges sentencing practice, in other words, they would be there. Secondly, it might be the case that the information that the project wanted (i.e., the data collection instrument) was inadequate to capture the information important to sentences that was present in the probation files -- whether that information was outdated or not. It might also be the case that judges in Massachusetts sentenced on the basis of the outdated information since that appears to be what was available. (In a conversation with the director of the Massachusetts sentencing guidelines project, on the related issue of the validity or accuracy of the data, Dr. Hutner indicated that he believed that the accuracy of the information was a moot point -- accurate or not, it was the information that the judge had access to and used to make the sentence decision. We wonder why this same logic did not also extend at least to those cases where the judge apparently had to sentence on the basis of outdated pre-trial investigation report information.) Of course, since we don't know which, if any, of the above possibilities are true, we really have no way of telling whether the final Massachusetts sample, although initially random, was truly representative of case and sentence dispositions in the end.

[26] Remember that the probation records were chosen as the primary data source because they supposedly contained complete offender background information; apparently they did not fulfill this expectation.

[27] The district attorney records were used as the primary source of information about the offense, rather than for offender information. It is possible that those records might be severely lacking in offender information and, thus, if they had been used as the primary source for both types of data, the number of cases dropped because of missing information in the district attorney records might have been higher.

[28] Our interviews with judges, prosecutors, defense counsel and others throughout the state, as well as our observations of sentencing hearings, led us to believe that observation of actual sentencing hearings would be one of the best Massachusetts data sources -- although such observations would not suffice as the sole data source, especially since judges also have access to the offender's prior record at sentencing. At those hearings, probation file information is often updated verbally, defense and prosecution attorneys offer sentence recommendations to the judge verbally, and the judge often comments on what he or she considers to be pertinent information about the offender and the offense. To our knowledge, the information we just described is not

available from any source other than the actual sentencing hearing.

[29] Sentencing guidelines project progress report, dated 24 May 1979. While the exact date of the end of data collection is not known to us, this progress report also notes that the Massachusetts project expected data collection to be complete by mid-June, as compared to the 30 January 1979 progress report which projected completion of data collection for the end of March.

[30] Sentencing guidelines project memorandum, titled Variables Dropped from Data Analysis, dated September 1979. This memorandum states that "Variables were dropped because analysis of the data collected would not be meaningful." While variables that had a higher than 17 percent missing information rate were excluded from analysis, this memorandum also notes that "the response rate for most variables was over 90 percent." Indeed, a substantial number of variables do have valid responses at least ninety percent of the time, however, the category of "none" or "not stated" was often a valid category in many of these variables, and in many instances, coders were advised to use such categories if the information was not listed in the files. Thus, it is questionable exactly how many of the data for any variable (including those variables quoted as having a 90 percent response rate) were actually missing.

[31] The project memorandum, Variables Dropped from Data Analysis, mentioned above in note 30, also lists these two additional reasons for deleting variables from data analysis.

[32] The specific variables deleted involved: the defendant's weekly earnings, number of weeks employed during last 12 months, total assets and debts, type and amount of residence payments, neighborhood type, history of drug and alcohol use, type of drug used, rehabilitative efforts before trial, use of drugs or alcohol at the offense, cooperation with police, defendant's role in the offense committed, the age and ethnicity of the victim, the amount of goods recovered, the value of the drug distributed, the age of the receiver in drug distribution cases, the amount of court costs assessed, the date and amount of bail setting, changes in bail amounts, revocation of bail information, number of bail revocations, whether the police had to chase the defendant to apprehend, and whether the case was prosecuted by the major violator's section of the district attorney's office.

[33] Information pertaining particularly to bail (dates, amounts, changes, etc.) was noted by the project as being generally unreliable.

[34] The project did note, however, that while several of the variables mentioned in note 32 above were excluded from analysis because of missing information, there were other variables that could provide information similar to that which would have been provided by those deleted variables. Thus, a number of variables excluded from analysis were duplicitious of information contained in other variables, however, the duplicity was not the primary reason for their exclusion.

[35] Variables were excluded from analysis due to sample size for the offenses of fraud, gambling and sex related crimes.

[36] The project noted that the majority of cases evidenced similar responses for the variables pertaining to major violator status, distribution of drugs to juveniles, defendant's role in the offense, victim precipitation of the offense, and amount and type of court costs. Some of these variables were also excluded due to missing information, thus, there was perceived a dual reason for not conducting further analyses of these items.

[37] Over and over again studies of discretionary decisions made at various stages of the criminal justice process have found that two factors consistently influence the decision-maker's choice of outcome -- the seriousness of the offense and the offender's prior record. The repeated finding should not amaze researchers; of all the different information items examined, these two items are almost always present and are more or less accurate. In other words, despite oft-mentioned problems with classifications of offense seriousness and despite the fact that criminal history records might be partially inaccurate, these two items of information appear to provide the most available and least dirty data that are around.

[38] The Massachusetts project did give some thought to alternative methods of dealing with missing data; in fact, the statistical consultant to that project, Thomas Marx, prepared a memorandum on that subject for discussion. We do not know if the issues presented in that paper were discussed by the judges supervising the project or, rather, if the paper was intended for review by only the project staff.

[39] A relevant discussion of methods of dealing with missing data generally can be found in Cohen and Cohen (1975), Chapter 7.

[40] Not many of the sentencing guidelines projects at the time acknowledged that the sentencing guidelines developed after empirical study of sentence dispositions are still,

necessarily, a product of policy decisions involving the type of variables to collect information on, and how the analyses of variables is done statistically. Minnesota was one of the few projects, at that time, that recognized these influences; more will be said about that project's activities in Chapter 9.

[41] Massachusetts sentencing guidelines project press release dated 11 June 1978.

[42] See Chapter 3 of this report for a more complete discussion of the role of Jack Kress and his colleagues in the development of sentencing guidelines.

[43] The full title of the paper prepared by Jack Kress for the Committee on Probation and Parole is: Potential Guideline Variables. An Analytical Policy Paper Prepared for the Use of the Massachusetts Superior Court Committee on Probation and Parole at their Meeting on March 2, 1979 in Worcester, Massachusetts. The paper, which is described by Jack Kress on the last few pages as "a partial "think piece" to set the basis for discussion", contains a discussion of policy issues that Kress had encountered during his previous efforts to develop sentencing guidelines, and the piece was probably useful to judges who, at that stage in their project, did not know what to expect from the guidelines research at all. We object to two suggestions of Professor Kress's at that meeting of the Committee, however, as they seem to us to cloud two important issues with semantics rather than true solutions. The first objection we raise is to the endorsement by Kress of separate guidelines for plea and trial cases because separate guidelines would be "a more honest approach than that adopted by other jurisdictions." He noted that "the Court (would be) indicating explicitly that it recognizes plea bargaining....and that the Court would in this way be able to recapture sentencing authority." (Quotes taken from the Summary of Meeting of Committee on Probation and Parole, March 2, 1979, page 5.) The second objection is to a suggestion stated in the paper we cited above, that judges use their perception of the actual or "real" behavior involved, rather than the behavior of which the defendant was convicted, as their measure of the seriousness of the offense. While Professor Kress claimed that "The description of the judicial practice of sentencing on the basis of the judge's individual interpretation or conceptualization of the actual criminal behavior of the offender has been seen by many commentators as one of the valuable aspects of the national sentencing guidelines research I directed" (page 23), we disagree; indeed, we believe that serious legal, as well as ethical considerations, should prohibit such an approach (regardless of whether or not Professor Kress's claim that it occurs is true).

[44] Summary of Meeting of Committee on Probation and Parole, 2 February 1979, page 3.

[45] In a later memorandum, Minutes of Meeting of Committee on Probation and Parole, 1 June 1979, it was noted that cases heard by District Court judges were present in the data. The project staff promised the Committee that they would analyze the District Court judge cases for apparent differences compared to Superior Court judge cases; the results of such an analysis, if it ever was completed, were not released to us.

[46] Summary of Meeting of Committee on Probation and Parole, 2 February 1979, pages 3-4. The judges decided that charges that were either "continued without a finding" or which were "filed" were not to be considered valid dispositions. The exclusion of these two types of charge dispositions we believe exerted a substantial effect on the calculation of offenders' prior criminal history scores, as both dispositions are fairly common in Massachusetts. However, we do believe that the judges acted correctly in excluding these charges, for they were not legal convictions.

[47] Summary of Meeting of Committee on Probation and Parole, 2 March 1979, pages 9-10.

[48] This deicison was made after much debate sometime in June 1979. The major points of this issue are contained in a project memorandum by Thomas Marx, titled Use of Flat-Time in Data Analysis, dated 8 June 1979, although that was not the first time that the issue was raised. The project staff had earlier urged that some estimate of flat-time be used in the research in a memorandum to the Committee on 12 April 1979, titled Research Questions, page 2. The judges of the Committee on Probation and Parole discussed this memorandum at a meeting held on 1 June 1979, and during that meeting judges noted that while the analyses might use flat or real time as the dependent sentence outcome variable, they decided "for reasons of public relations especially, to use the formal sentence of the judge in the guidelines." Minutes of Meeting of Committee on Probation and Parole, 1 June 1979, page 10.

[49] See the section of Chapter 6 that provides an overview of sentencing in Massachusetts for a more detailed discussion of the offense "package" and its relation to the ultimate sentence disposition.

[50] In rejecting the crime-specific alternative to assessing the seriousness of the offense, the project noted: "At one end of the spectrum lies a system in which each offense is a mutually exclusive category. Such a system would

lead to crime-specific guidelines, that is a separate guideline sentence for on each offense charged. This option would require a very large sample, as well as a practical justification for thinking of each offense as an entity. It appears that neither is possible in Massachusetts." Sentencing guidelines project staff memorandum to the Committee on Probation and Parole, titled Research Questions, dated 12 April 1979, Page 4, emphasis added.

[51] In the same memorandum the project staff noted that "We have used the Sellin-Wolfgang idea to construct a set of scales applicable to our data. The potential for harm (i.e. risk) both physical and psychological and the actual harm done to the person are two scales. The damage to and loss of property is the third scale. The scales would thus describe the events that led to the offenses charged." Research Questions, page 5.

[52] It is not known why such a scale was never adopted; no further mention is made in project memoranda of how well the scale worked or why it was ultimately rejected. One reason that the scale may have been unworkable, presumably, could be because of the extent of missing data -- much of the information required by the three scales mentioned in note 51 above was probably missing in the original data. The project staff did not initially collect the data with any intent to conduct this sort of exercise and, thus, probably did not go to extreme lengths to ensure the presence in the original data of those items that were eventually used to construct the above scales.

[53] Massachusetts Sentencing Guidelines, November 1979. In the section of the November 1979 Sentencing Guidelines report titled Presentation and Testing of the Guidelines Model, (page 10) the project noted that "The Committee formally went through this exercise as a group after many months of pondering the relative gravity of different offenses."

[54] Massachusetts Sentencing Guidelines, November 1979, page 10. "At the conclusion of the testing period, each participating judge is invited to rate each offense on a one to six point scale or on a scale with as many points as he/she thinks are necessary to distinguish offenses." The testing was conducted, as we noted in the text; however, judges present at the semi-annual judicial conference where the guidelines were first presented in November 1979 indicated that they did not feel comfortable with this method of computing seriousness scores. (It is possible that one reason that judges did not feel wholly secure with this approach might have been that the Committee's scoring exercise had

failed to produce any offenses that rated a seriousness score of 5.) When the final version of the sentencing guidelines was explained to judges in May 1980, the fact that the method of computing offense seriousness had been altered greatly pleased a number of judges.

[55] Unfortunately, since the Massachusetts project staff did not release any of the exact analyses, we cannot be entirely sure on what basis these items (and not other items as well) were determined to be strongly and independently predictive of sentence lengths. The project did not give statistical evidence in its reports to the judges; rather, the following analogy was drawn: "The procedure for finding a set of the most influential factors from the seventy seven or so considered can be likened to a tennis tournament....Unlike the tennis tournament, because the number of factors eliminated after each match depended on statistical evidence from that analysis, not one but eight winners emerged." Massachusetts sentencing guidelines, November 1979, page 4.

[56] Massachusetts Sentencing Guidelines, November 1979, page 5.

[57] The phrases in parentheses were added to correct some obvious mistakes in the report about the Massachusetts Sentencing Guidelines, May 1980, page 1.

[58] We mentioned in Chapter 6 of this report that a special spotlight team from the Boston Globe newspaper was conducting an inquiry into the prevalence of racial disparity in criminal sentences at the time that the sentencing guidelines project was just getting under way.

[59] The November 1979 Massachusetts Sentencing Guidelines report states: "The minimum and maximum proposed by the Committee for trials is 50% of the recommended guideline sentence." (page 7)

[60] The range for plea cases is not as straightforward as is the range for trial cases. The same report states: "For the test period, any guideline sentence between 0 and 60 months has a range defined by a lower boundary of non-incarceration and an upper boundary of 80% of the guideline sentence. For guidelines sentences over 60 months, the range is defined by a lower boundary of 5% and an upper boundary of 85% of the guideline sentence." (page 9)

[61] The first staff report issued by the Massachusetts sentencing guidelines project to all of the Superior Court judges stated, by way of introduction:

> This is the first in a series of staff reports designed to keep you informed about the progress of the Sentencing Guidelines Project. Future reports will discuss a range of issues related to the formulation and implementation of the guidelines. Your questions and views on the guidelines, as well as any suggestions regarding the content of these reports, are solicited. (Sentencing Guidelines Project, Staff Report No. 1, June 1979, p.1)

[62] The three Massachusetts project staff reports that our project was able to obtain are: Sentencing Guidelines Project, Staff Report No. 1, June 1979; Sentencing Guidelines Project, Staff Report No. 2, July 1979; and Sentencing Guidelines Project, Staff Report No. 3, September 1979. We are fairly certain that these three are all of the staff reports issued, since the date of the last report was only one month prior to the formal introduction of the guidelines at the judges' semi-annual judicial conference.

[63] Sentencing Guidelines Project, Staff Report No. 1, page 2.

[64] Of course, the Committee was not under any legal obligation to publish public reports of the guidelines project's analyses; thus, that they did not release such a report should not be misconstrued. However, in most research efforts, a report of that kind to a broader audience is somewhat standard. The project staff did prepare four "analyses" reports just prior to the presentation of the guidelines to the judges. However, these "analyses" reports were prepared by the project for the Committee on Probation and Parole, not (to our knowledge) for all of the Superior Court judges or the public. Those four analyses reports are cited: Sentencing Guidelines Project, First Report on Analyses, 2 October 1979; Second Report on Analyses, 15 October 1979; Third Report on Analyses, 19 October 1979; and Fourth Report on Analyses, 22 October 1979.

[65] Noting that the session of the judicial conference that we were invited to attend was not a session open to the public (indeed, the entire conference is especially private for reasons relating to judicial security), we would like to thank Chief Justice Lynch and Judge John Ronan for extending the invitation to the staff of our project.

[66] The presentation of the initial Massachusetts sentencing guidelines was far from professionally organized and run. The major problem with the presentation was that the Committee had not held enough open meetings prior to the introductory session to iron out potertial disagreements among

the judges about the content of the guidelines. Also, the major portion of the presentation (which did not have benefit of visual aids or case examples) was conducted almost entirely by the research team, rather than by judges -- the judges did not seem to regard the entire guidelines development enterprise as their effort, therefore.

[67] That it is the legislature's responsibility to stipulate crimes and penalties is a recurring issue in Massachusetts; we discussed this point in more detail in Chapter 6 of this report.

[68] Our research staff were also able to attend the presentation of the second version of the Massachusetts guidelines to the judges in Framingham, Massachusetts. We would again like to thank Chief Justice Lynch and Judge Ronan for the invitation.

[69] We did not attempt to examine the first presentation in as much detail as we examiine the second here because some flaws are expected in the first presentation of anything. There should have been some improvement in the latter presentation, however.

[70] See the Massachusetts Sentencing Guidelines, May 1980, and the preceding section of this report, for a full description of the structure and content of the final version of the Massachusetts sentencing guidelines.

[71] Since members of our project staff were able to attend and take notes at this session, as well as at the first, we are able to provide a complete description of the events of that session here.

[72] We disagree with this point. The interested reader should consult the forthcoming article by Stecher and Sparks, Removing the Effects of Discrimination in Sentencing Guidelines, in Sentencing Disparity (1981).

[73] In our discussion, in Chapter 6 of this report, of the political climate at the time of the guidelines development, we provided a full description of the activities of the legislature during the same period. The reader is referred to that chapter for more information.

[74] The correspondence to which we had access included: a letter from Chief Justice James P. Lynch of the Superior Court to Senator Arthur J. Lewis, 18 July 1979; a letter from Chief Justice Edward F. Hennessey to Representative W. Paul White, 29 January 1980; a letter from Representative White to Chief Justice Hennessey, 5 February 1980; a letter from Chief

Justice Hennessey to Representative White, 8 February 1980; a letter from Representative White to Judge John T. Ronan, 20 March 1980; and a letter from Judge Ronan to Representative White, 7 April 1980. We, of course, cannot be sure that this selection of letters includes all of the correspondence between the legislature and members of the judiciary; however, the selection does provide an overview of the chief concerns of both of those groups relative to the concept and use of the Massachusetts sentencing guidelines.

[75] There were also numerous questions posed by the legislature about the implementation and procedure to be involved in the use of the sentencing guidelines by the judges. A sample of the 26 procedural questions posed by the legislature are the following:

> Will there be specific findings on the record of the facts used in the enhancements?
>
> Will the defense counsel and prosecution have access to the guideline calculations used by the judge?
>
> What will be the burden of proof for finding the enhancement factor?
>
> Will testimony be taken on the issue of the enhancement factors, with direct and cross examination?
>
> If there is not a separate evidentiary hearing, what provisions have been made for alteration in the rules of evidence, especially as to relevancy of testimony, so that evidence for and against the enhancing factors may be entered?
>
> If a plea of guilty is entered and no evidence is entered as to the absence or presence of certain enhancement factors, will there be a hearing on these elements?
>
> If such information is not given to the defendant, will the plea be based on enough information to make it fully voluntary, and will the defendant then have some right to a hearing on the issue of enhancement factors?
>
> What will be the avenue of review for guidelines sentences? What will be the standard on review of a guideline sentence?
>
> Will a sentence outside the guidelines and not accompanied by written reasons be therefore on its face invalid?

What reasons will be sufficient for a judge to go outside the guidelines? Will the reasons for sentencing outside the guidelines be reviewable? If not, why must the judge specify the reasons in writing?

As is obvious from the sampling above, the legislature believed that adoption and use of sentencing guidelines could have important ramifications to the operation of the courts, both on the trial level, and the appellate level. These questions were asked in the letter from Representative W. Paul White to Judge John T. Ronan, 20 March 1980, pages 3-4.

[76] Letter from Representative White to Judge Ronan, 20 March 1980, pages 1, 2, 5.

[77] Letter from Judge Ronan to Representative White, 7 April 1980, excerpts from all pages.

[78] The session of the probation officers' training that our project staff were able to attend took place on 14 May 1980 in Framingham, Massachusetts. At that meeting, the Commissioner of Probation, Joseph P. Foley, gave a short speech to the department staff urging them to provide the judges of the Superior Court with any assistance possible. Donald Moran, the Assistant Supervisor of Probation, then explained the role of the probation department in the judicial use of sentencing guidelines; Judge John T. Ronan provided the officers with an overview of the sentencing guidelines history, research, and development process; and finally Michael Hutner explained the use of the guidelines for several specific cases. Once again, we would like to thank the staff of the Massachusetts sentencing guidelines project, the members of the Committee on Probation and Parole, and the Commissioner of Probation and his staff for the open invitation to attend these sessions.

[79] Introductory comments of Donald Moran on 14 May 1980 in Framingham, Massachusetts during probation officer training session.

[80] Just prior to the formal introduction of the Massachusetts sentencing guidelines, the probation department discontinued preparing pre-trial investigation reports for each case and replaced those reports with a short form pre-trial intake sheet. In addition, judges could request pre-sentence investigations for cases where they needed information, in addition to that on the pre-trial intake form, for sentencing purposes. A considerable amount of paperwork had thus been dispensed with along with the long version of the pre-trial investigation for each case. This should have allowed probation officers to have additional time to complete

guidelines forms and may have been a factor in the probation officers' initial positive reception of the guidelines. Ironically, as we will discuss in more detail a little later, the probation officers eventually criticized the guidelines because of the amount of time it would take for probation to complete the required forms. (For a more complete description of the change from pre-trial investigations to pre-sentence investigations, see the early sections of this chapter and Chapter 6.)

[81] The question of how filed charges should be included in the computation of the current and prior offense seriousness scores proved to be one of the biggest problems encountered by the judges and probation staff in the use of the Massachusetts sentencing guidelines. (See Chapter 6 of this report for a discussion of the complicated process of "filing" charges that is used in Massachusetts.

[82] A description of the procedures that we used to select interview sites, and of the number of persons of each occupation that were interviewed during the summers of 1979 and 1980 is provided in Chapter 6 of this report.

[83] As we will discuss in more detail a little further on in this section, this point holds important implications for the use of probation as a disposition as well.

[84] It all depends on the point of view. We suspect that a fifty percent range was fine for those judges who really didn't want guidelines -- a range of that size would not constrain their sentences very much, so it would be a nice feature. On the other hand, if a judge wanted guidelines to provide a specific sentence, the fifty percent range produced such a wide choice of sentences that the guidelines sentence was not much of an aid.

[85] It could be argued, of course, that the difference in seriousness for these two offenses would be made more explicit in the scoring of the weapon use item. However, judges noted that even for offenses that do not involve weapons, per se, there are usually a whole host of factors that will differentiate the seriousness of one offense from another of the same general variety.

[86] Judge John T. Ronan, the Chairman of the Committee on Probation and Parole, noted in an interview with our project staff that, though the Committee had expected the scoring of "duplicitious offenses" to be the major problem with the guidelines, it did not turn out that way. Rather, the question of how to score filed charges was the major problem encountered by the judges using the guidelines.

[87] We should note that none of the judges thought that the articulation of reasons for sentences was a problem with the guidelines; all of the judges expressed the opinion that if a judge could not give reasons for his sentence, then that sentence was irrational. The aspect that worried the judges was that the articulation of reasons for deviating from the guidelines might eventually lead to some limit on the types of reasons thought acceptable for deviation -- thus, limiting their discretion.

[88] The probation department staff, after all, would be responsible for obtaining the prior record from the court clerk's office, the current offenses from the district attorney and the injury and weapon scores from the judges, before the forms would be complete. The staff claimed that all of this running around took up quite a bit of time. We suspect that this was also the case due to the size of the range provided by the guidelines. See note 84 and text above.

[89] Probation staff thought that an alternative method could be that the judge would indicate to them for which of the plea cases guidelines forms would be required. Such a procedure would save them a lot of useless paperwork.

[90] We have not conducted any further detailed interviews with Massachusetts judges, probation, prosecutorial, or defense staff since the summer of 1980. The procedures involved in resolving questions about the guidelines may have been streamlined since that time.

[91] On 22 February 1980, Judge John T. Ronan of the Committee on Probation and Parole responded to a letter from Philip A. Rollins, President of the Massachusetts District Attorneys Association, with the following comments:

> Your interest and that of your Association in the Sentencing Guidelines Project is understood, appreciated and acknowledged. More relevantly it is needed since the expertise and experience that the District Attorneys possess would be and should be an invaluable source and reference.
>
> Your letter, however, seems to be based on a false assumption that the Court has been fashioning a specific sentencing structure and intends to present the same as an accomplished fact. Please be assured that that is not the case.
>
> In the event of a favorable evaluation, the proposal will be then subjected to the solicitation of comment, input, observation and criticism of all interested

groups. Specifically, the District Attorneys Association will be invited to participate and bring to bear upon any proposal its collective expertise.

However, the input of the district attorneys throughout the state was never requested prior to the implementation of the sentencing guidelines in May 1980. We would like to comment that this oversight (as well as the oversight of defense counsel) was particularly unfortunate in two respects. First, district attorneys in Massachusetts are political persons -- they are elected and are often past legislators; thus, they have quite a bit of contact and influence with the legislature and could have been a powerful tool that the judges could have used to curtail the legislature's propensity toward enactment of mandatory sentences, had their support been enlisted in the development and implementation of sentencing guidelines. Secondly, had the district attorneys been involved in the development of the guidelines, they might have been more pleased with the structure and content of the guidelines. This latter point applies to defense counsel as well.

[92] Regardless of which of these ways the district attorneys found out about the Massachusetts guidelines, they were still quite mad about them. We should note that these are the same three ways that defense counsel learned about the sentencing guidelines. All in all, the method used to alert persons in the system to the sentencing guidelines were not ones designed to elicit support of the guidelines. Compare the procedures for guidelines construction and implementation used in Massachusetts with those used by the Minnesota sentencing guidelines project discussed in Chapter 9 of this report.

[93] District attorneys had an idea that guidelines were being developed because their records had been used for data, but the defense counsel did not provide any data to the project, and were never even contacted about the project's existence.

Chapter 8: The Massachusetts Sentencing Guidelines: A Statistical Analysis

A Secondary Analysis of the Massachusetts Sentencing Data

The Massachusetts sentencing guidelines project compiled background and offense information on a final total of 1440 defendants sentenced in the Massachusetts Superior Court between November 1, 1977 and October 30, 1978. We have already described the sampling methods used to choose cases for inclusion in that data base and we have discussed the research analyses and conclusions developed by the Massachusetts sentencing guidelines project. Before we can describe our secondary analyses of that sample's data, a brief description of the characteristics of the sentenced defendants is in order. Unlike several of the other statewide projects developing sentencing guidelines throughout the country (notably Minnesota and Michigan), a thorough report of the contents and analyses of the Massachusetts sentencing data was never allowed public release by the sponsoring Superior Court judiciary.[1]

In presenting this basic description of the sample we will be unusually brief -- and we do not pretend that we are presenting to the reader a step-by-step summary of how the Massachusetts project staff actually developed their sentencing guidelines. The materials released by that project did not contain enough technical information for us to replicate exactly the research of that project.[2] Given the information available to us about the analyses that the project completed, and given the original data collected by the project, the best that we will be able to do is to generate what we believe is a comparable re-analysis of their data.[3]

To that end, we begin this chapter with an overview of what the cases in the Massachusetts sentencing guidelines sample were like. In this section we will also investigate relations between some items of information that evidence similarities, such as the number of prior convictions and the number of incarcerations. These need to be examined since they may affect decisions about which items, of a group of highly inter-correlated items, should be used in regression analyses, in sentencing guidelines based on such analyses.

We next examine the fit of the Massachusetts sentencing guidelines to the cases from which they were constructed. Our aim in this section is to see how well the Massachusetts sentencing guidelines actually are able to predict the judges'

sentencing decisions. In doing this, we will examine not only the factors (and their associated weights) in the guidelines, but also the ranges provided for by the Massachusetts guidelines. We also examine cases that may be thought of as unusual in one or another respect, and the effect that exclusion of such cases will have on the model's ability to predict sentences.

The final section of this chapter will present our project's after-the-fact construction of sentencing guidelines for Massachusetts. Differences between the model that the Evaluation of Statewide Sentencing Guidelines project developed, and the sentencing guidelines proposed and being used in Massachusetts, will also be discussed in that section.

Demographic and Background Information

More than nine of every ten defendants on whom case information was collected were male; only 114, or eight percent were female. The majority of the defendants (69 percent, or 982 cases) were caucasian; 20 percent were black, and the remaining 11 percent of the defendants were of Hispanic origin.

Some differences in racial distributions were evident when the county where the offender was sentenced was analyzed. Almost all of the defendants sentenced in Berkshire county, for instance, were white: by contrast, 42 percent of the Suffolk county (Boston) defendants were black, and 26 percent of the Hampden county defendants were hispanic. These figures do not surprise us, as they reflect the relative ethnic population groupings across the state; yet they are important to bear in mind during later discussions pertaining to between-county sentence variation.

Less than a fourth of the defendants (317 cases) were either married and/or cohabitating at the time the offense was committed. One fourth of the single defendants (271 cases) were either separated or divorced at the time of the offense; the remainder had never married. Fewer of the female defendants were single at the time of the offense, yet a higher proportion (44 percent) than that indicated for the male sample (23 percent) were either divorced or separated.

Half of the 1,379 sentenced defendants were employed, in school, or in the military at the time the offense was committed. One-fifth of the 675 employed defendants were laborers (as classified by the Socioeconomic Index for Occupations in the Detailed Classification of the Bureau of the Census), 20 percent were operatives; service and crafts personnel split another 28 percent of the cases, while

managerial staff accounted for 11 percent and clerical/sales staff acounted for 8 percent of the sample. Only six percent (37 cases) of the defendants were employed as professional staff. The overwhelming majority (95 percent) of the unemployed defendants were described as having no occupation at all; the remaining 5 percent claimed to be laborers. It is interesting to note that none of the other occupational titles were claimed by any of the unemployed defendants; this might not be the case today, given the higher unemployment rate since 1981.

Half of the cases involved unemployed defendants, but only 22 percent claimed to have no primary source of income. About half of the defendants claimed their assets, earnings, retirement or disability pensions as the primary source of their income; 16 percent received income as a primary source from their family or relatives, while another 12 percent relied on public assistance (of a number of types) for income. Only twenty-eight cases (2 percent of the sample) were listed as having a primary source of income gained by illegal activities.

Specific information about the defendants' working histories was available for 1150 of the 1440 offenders. A third of these defendants were not employed at all during the twelve months prior to the offense; another 28 percent (or 318 offenders) were employed for the entire year prior to the commission of the offense. The overwhelming majority of employed defendants (about 70 percent) held only one job during the year prior to the current offense; however, a small majority had held several jobs, so that the mean number of jobs among those employed in the preceding year was 2.6.

The majority of defendants lived in areas classified by the Massachusetts sentencing guidelines project staff (by what criteria, we do not know) as "average residential" communities; 28 percent, however, were listed as living in "low income, depressed" areas. Quite a number of cases -- 57 percent -- had to be dropped from the analysis of this item due to missing information, but some between-county variation in neighborhood type for the remaining 612 cases was still evident. Almost half of the defendants in Hampden County lived in low income areas -- the largest percentage found listed in this category of all of the counties considered. Hampshire and Suffolk Counties had substantially higher proportions of defendants living in "mixed residential commercial" areas than did all counties combined, while Plymouth County had a substantially higher proportion than all counties combined of defendants who lived in "average residential" areas (81 percent, as compared to 51 percent of the cases across all counties). Some variation appears in the

distribution of upper middle class area residents by county, but the numbers in all counties for this category are rather small to permit valid comparisons. In addition, the two counties with the largest sample N's -- Suffolk and Middlesex -- were under-represented by almost two-thirds of their cases on this item because of missing information.

Three of every four defendants had no military history prior to or at the time of the offense; less than ten percent of the defendants were in school at the time of the offense. Six percent of all of the defendants in the sample completed all or only part of grade school; a third completed all or part of junior high school; a little less than half of the sample's cases had attended high school. But only a third of the total sample (or 502 offenders) had completed high school, and of these 216 had continued their education in college or graduate school. Interestingly, the sample contained 14 cases where the offender had completed either a master's or doctoral degree program.

The offenders were generally classified as being in good physical and mental health and did not usually have any physical handicaps. A hefty minority of defendants (235 cases) were listed as having had past mental health problems that had been treated; however, over three-fourths of the defendants did not apparently have any past or existing mental health problems.

The proportions of defendants listed as having a history of alcohol or drug use, or as having a current drug or alcohol problem for which treatment was recommended, generally approached a fourth of the total sample. Almost a fourth of the 1136 defendants for whom this information was available had a prior history of alcohol use, and 19 percent of 1374 defendants had a current alcohol problem for which treatment was recommended. Similarly, about a fourth of 1177 defendants had a prior history of heavy drug use, and a fourth of 1375 defendants had a current drug problem that was in need of treatment. (The N's here exclude missing data cases.) As might have been expected, there was substantial association between current drug abuse and a prior history of drug use ($r = +.79$), as well as between current alcohol abuse and a prior history of alcohol use ($r = +.92$). There was, however, virtually no association between drug and alcohol problems. Drug and alcohol use at the time of the offense were not uncommon (they figured in 9 and 18 percent, respectively, of the cases).

Juvenile Criminal History Information

Several items of information included in the Massachusetts sentencing guidelines pertained to the offender's prior criminal history specifically as a juvenile. Among the items recorded were: year of first juvenile adjudication (later recoded to be the offender's age at the first juvenile adjudication), the number of juvenile adjudications for all types of offenses, the number of juvenile adjudications for serious crimes, the number of juvenile supervision terms imposed, the number of juvenile supervision term revocations, and the number of commitments to the Division of Youth Services.

Three out of every four defendants had no juvenile adjudications for any type of crime. The average age at the time of the first juvenile adjudication for those 380 offenders with a juvenile criminal history was 14 1/2 years. The bulk of the juvenile offenders (73 percent) were between the ages of 14 and 17 at the time of the juvenile disposition. For those with juvenile records, the number of adjudications averaged 4 1/2. Nine out of every ten juvenile offenders were adjudicated for "serious" offenses; the average number of "serious crime" adjudications averaged 3.7 per juvenile offender.

Predictably, all of the juvenile record items were positively intercorrelated (r between +.53 and +.94 for most pairs); in such a situation, of course, only one such item (preferably the most reliable) should be used for analysis, unless some better-performing composite can be found. In most of our analyses we simply used "number of juvenile adjudications"; we do not know exactly what the Massachusetts researchers did.

One traditional aspect of juvenile criminal records seems to be substantiated once again by the Massachusetts data: male juvenile offenders total up longer juvenile criminal records than do females. Over 12 percent of the male offenders had 4 or more juvenile adjudications, while this was true of only 3 1/2 percent of the female offenders. The differences between male and female juvenile records further indicated that males were adjudicated for more serious offenses, were sentenced to more supervision terms, had more supervision term revocations and were more often committed to the custody of the Division of Youth Services than females.

Criminal Justice Process and Offense Information

A little over half of the defendants sampled had committed the offense for which they were convicted during the

year 1977. A third of the defendants committed their offense prior to 1977, including 18 percent where the offense was in 1976; an additional 18 percent of the defendants committed the offense during 1978. Similarly, over half of the defendants were arrested during the year 1977. A fourth of the defendants were arrested, however, before or during 1976. Thus, many of the cases in the sample had "aged" substantially during the time before trial.

A substantial number of the defendants (40 percent) were arrested at the scene of the offense; only 16 percent of the defendants were arrested after a chase by the police. Over ninety percent of the offenders were legally at large at the time of the offense, although almost a third were on probation or parole; six percent were free on bail, and another[4] percent were incarcerated at the time of the offense. Finally, 3 percent of the defendants were either on a work release or furlough pass from a penal institution, or were fugitives or escapees at the time of the offense.

A number of items of information were present in the Massachusetts sentencing data that pertained to the offenders' status at the time of arrest and disposition, the offenders' prior pre-trial release periods, and the number of outstanding charges of which the offender was accused. The specific items of information included: the length of time detained from arrest to disposition, whether the offender was at liberty at the time of case disposition, whether the offender had been released on his own recognizance, whether the offender's bail had been posted, the final amount of bail set, whether a change in the type or amount of bail had occurred, whether there were any pending charges against the offender, the number of prior bail defaults, the number of pre-trial release periods in which bail defaults occurred, and whether the offender's current bail had been revoked. As was found to be the case with the offenders' juvenile criminal history items, a number of these items were associated with each other in degrees that ranged from moderate to strong: and as we shall see, several were significantly associated with sentencing dispositions.

A little more than half of the offenders in the sample were not detained at all between arrest and disposition. Those offenders (N=549) who were detained were held anywhere from 1 to more than 996 days, averaging 102 days per offender. At the time of disposition, 846 offenders (62 percent of the sample) were free. A total of 602 of these offenders had been released on their own recognizance, while 244 offenders had been released after bail had been posted. Thirty offenders were in confinement for diagnostic tests at the time of disposition and 214 offenders were incarcerated. An

additional 231 offenders were incarcerated in local county jails when the case reached disposition; 70 percent of these offenders were in jail because their bail was not posted, 15 percent were in lock-up due to bail revocation, and a final 15 percent had had their bail petitions denied.

A change in bail status occurred in a fourth of the cases (349 offenders). The actual type of bail change was not recorded (except for revocation) and could have been any of several types: persons previously allowed bail or release on their own recognizance could have had that release revoked, or persons who had previously not been granted bail could have had a bail petition reviewed and bail granted. (The amount of bail may also have changed following the filing of new charges or information.) Six percent of the offenders (or 84 offenders) had their release on bail revoked; the number of bail revocations for those offenders who at any time had bail revoked averaged 1.2. About a third of the defendants had additional pending charges at the time of the offense, and although the number of pending charges was not recorded in the data set, and nor was the time at which the pending charges were discovered, these charges probably figured prominently in the setting of bail. The amount of bail set (for the 555 offenders where that information was available) averaged about $14,000.

Over a third of the defendants had had at least one (and 6 percent had between 11 and 59) prior bail defaults. Many of the offenders apparently defaulted more than once in each prior pre-trial release period. The total number of pre-trial release periods in which a bail default actually occurred ranged from none up to 16.

Six percent of the offenders were prosecuted by the Major Violator section of the prosecutor's office. Like similar projects in operation throughout the country, the Major Violator's division identifies offenses that are thought to characterize "career criminals" and specializes detection, investigation, and prosecution teams to handle these types of cases so as to maximize the likelihood of conviction. Although information was available on the degree of offender cooperation with the police in only 400 cases, a fifth of these cases were said to have cooperated "completely" with the police and prosecution. An additional third of these cases offered "some" cooperation to criminal justice authorities, and about half of the offenders apparently did not cooperate at all with the police or prosecution staff.

Plea agreements accounted for the overwhelming majority of the cases' dispositions. Almost 1200 cases were settled by a plea, while only 198 (or 14 percent) went to trial before a

jury, and only 23 cases were tried before a judge (without a jury). Defense counsel was provided at public expense for 918 offenders (or about two thirds of the cases). Of those cases requiring public defense assistance, 70 percent were handled by the Massachusetts Defenders Committee; the remaining 30 percent were assigned private counsel by the court. While all of the cases in the final sample were disposed of in the Massachusetts Superior Court, 16 percent of the cases were heard before a District Court judge who had been assigned to the Superior Court bench to aid in the disposition of cases.

Twelve percent of the cases appealed their sentence dispositions to the Appellate Division of the Superior Court. As mentioned earlier, this Division, which consists of three judges, exists to review sentences considered either excessively harsh or lenient; however, only the defendant, and not the prosecution, has a right of appeal.

Current and Prior Offense Information

The Massachusetts sentencing data were collected in such a way as to provide information on all charges lodged against the defendant, as well as on the disposition of each charge. Due to the nature of the several statistical programs that we decided to use during the re-analysis of the Massachusetts sentencing data, it was necessary to aggregate information across charges and dispositions and provide case-level (as opposed to charge-level) data. This was accomplished primarily by noting the presence or absence of various items of information included in the original data base and summing across charges where the presence of such information was indicated. This procedure generated information that, while reflecting accurately the presence or absence of crime characteristics for an entire case, did not indicate particular charges within a case that the information applied to.

A total of 4,646 charges were filed against the 1,440 defendants -- an average of 3.23 charges per defendant. Using the figure of 4,646 as our statistical base for a moment, the two largest groups of charges fell into the classifications of Crimes against Persons (36 percent, or 1,672 charges) and Crimes against Property (26 percent, or 1209 charges). Fifteen percent of all of the charges were for Controlled Substance Violations, while 7 percent were listed as Motor Vehicle violations by the Massachusetts sentencing guidelines project. Finally, Crimes against the Currency and Weapons offenses split 8 percent of all of the charges noted, while the remaining 8 percent of the charges fell into the Miscellaneous Offense or Unknown Offense categories. The classification of conviction offenses substantially mirrored the distribution of the offenses charged.

Ninety-five percent of the defendants were charged as the principal actor in all of the charges against them; only 5 percent of the offenders were described as having a minor or accessory role in the offense.

As will be recalled from our description in Chapter 7, the Massachusetts sentencing guidelines do not provide specific guidelines for certain types of offenses; rather, the guidelines incorporate the statutory seriousness of all conviction offenses as one basis for the computation of the guideline sentence.

We attempted to duplicate the total seriousness scores obtained for each offender for charges, convictions, and prior records, by classifying each offense according to the statutory seriousness score above and then multiplying that seriousness score by the number of counts recorded for that particular charge. While the attempt was made to exactly duplicate the seriousness score developed by the Massachusetts sentencing guidelines project, we cannot ensure that we created an exact replica, for several reasons. First, the numeric descriptors attached to offenses coded in the Massachusetts sentencing data often do not match (exactly) statutory classifications of offenses. It was often necessary to make a judgement, for example, that the offense coded as "Assault with Intent" was really "Assault with Intent to Rob" as opposed to "Assault with Intent to Rape". The two do not necessarily carry the same penalty. Additionally, very often a label attached to an offense in the Massachusetts sentencing data could represent several different offenses as listed in the Massachusetts General Laws. For example, the offense "Unnatural Acts" is coded as "6625" in the Massachusetts sentencing data. It was assumed for the purpose of developing a seriousness scale, that this descriptor applied to both the offenses "Unnatural Acts" and "Unnatural Acts with a Child under 16" as defined in the Massachusetts statutes (cf. Chapter 272, Sections 35 and 35A).

Lack of some specific information pertaining to second and subsequent offenses also caused some difficulty in constructing offense seriousness scores. The Massachusetts statutes often provide substantially heavier penalties for offenders who are convicted of an offense that they have also been convicted of in the past. However, in order for the heavier statutory penalty to be applicable, the offense listed on the indictment must specify that the offender is being charged with a second or subsequent offense. It is not enough to scan the prior record for previous convictions for this same offense. Unfortunately, the Massachusetts project staff did not collect information as to whether the indictment listed specifically that the offense was a second or

subsequent offense, thus neither they nor we were able to take this information into account when computing the seriousness of offenses.

A further problem encountered during the duplication of the Massachusetts seriousness scores for charges and convictions was more encompassing and worrisome, and has been briefly mentioned earlier. The statutes do not provide sentences in the same phrasing as that used to measure the sentence seriousness by the guidelines project. The statutes rarely provide a stated punishment of "exactly five years" for example. Rather, the statutes typically state that punishment "may not be more than 5 years". Thus, for our analysis, an offense score of 1 ("Maximum penalty is exactly 5 years") includes all offenses where the statute states that punishment may not be more than 5 years. It could, in other words, be less than or equal to 5 years.

There is one additional difference in our replication of the seriousness scores of charged, convicted, and prior offenses. The Massachusetts sentencing guidelines specifically include an allowance for the offender's degree of participation in the offense in only one instance. The guidelines state that "Accessory after the fact receives a score of 2 since the maximum penalty is 7 years". According to this statement, the only adjustment that would be made to the seriousness score assigned to an offense would be to change it to a 2 if the offender was an accessory after the fact. While it is not clear whether this instruction was followed exactly by the project in computing seriousness scores, it did not seem to us to be appropriate, in light of statutory references, to score accessories after the fact in this manner for all levels of seriousness. 4 In addition, the Massachusetts statutes do not mitigate the seriousness of offenses only where the offender was an accessory after the fact; rather, the offense's seriousness is also lowered if the offense was a conspiracy or an attempt as well.

To develop offense seriousness scores that would most closely correspond with the Massachusetts statutory provisions the following procedure for assigning seriousness scores was followed. First, each offense was assigned a score from 0 to 4 reflecting the statutory maximum penalty for a completed offense of that nature. Accessory before the fact offenses were assigned the same score that the completed offense would receive. If the offense charged was conspiracy, the seriousness score was adjusted according to the chart below.

Conspiracy Offenses

Maximum Penalty Completed Offense	Original Score	Maximum Penalty Conspiracy	Adjusted Score
life	4	20 years	3
20 years	3	10 years	2
10 years	2	5 years	1
5 years	1	Misdemeanor	0
below 5 years	0	Misdemeanor	0

If the offense was an attempt, the seriousness score would be adjusted according to a somewhat different scheme depicted below.

Attempt Offenses

Maximum Penalty Completed Offense	Original Score	Maximum Penalty Attempt	Adjusted Score
life	4	10 years	2
20 years	3	5 years	1
10 years	2	5 years	1
5 years	1	5 years	1
below 5 years	0	Misdemeanor	0

If the offender was an accessory after the fact, the seriousness score was assigned as shown below.

Accessory After the Fact

Maximum Penalty Completed Offense	Original Score	Maximum Penalty Acc. After Fact	Adjusted Score
life	4	7 years	2
20 years	3	7 years	2
10 years	2	7 years	2
5 years	1	Misdemeanor	0
below 5 years	0	Misdemeanor	0

Seriousness scores computed for the totals of charged offenses, for each of the 1440 offenders in the Massachusetts sentencing data, ranged in numeric value from 0 (where the total seriousness score was based entirely on either misdemeanors or minor felony offenses) to a high of 244 points, with an average score of 6.22 per offender. The summed seriousness scores computed for offenses of conviction for each offender were somewhat lower: they ranged in value from a low of zero to a high of 118 points, with an average score of 4.28 per offender.

The total seriousness score for prior offenses was also computed for each defendant. The sums of the seriousness scores for prior offenses ranged in value from 0 to 77 points,

with an average score of 4.76 per offender. It should be noted, however, that the zero score for prior offenses could represent one of two possibilities: either the offender did not have any prior record, or the offender did not have a serious prior record (for example, that he or she had only a history of misdemeanor or low severity felony offenses).

Another measure of the seriousness and types of crime which offenders had been convicted of was also computed by us from the Massachusetts data. Due to the parole eligibility rules mentioned earlier, offenders who have been convicted of offenses defined by law as "persons offenses" are released at two-thirds, rather than one-third, of their minimum state prison sentences. Each of the current charges that resulted in a valid conviction was classified as either a persons or non-persons offense [5] according to the Massachusetts statutory scheme. The number of current "persons offense" convictions was then summed for each defendant. About half of the defendants were convicted of 1 or more persons offenses. The bulk of these defendants (54 percent of the 716 cases) were convicted of only 1 current persons offense, though a further one-third of this group had either 2 or 3 current persons offenses.

The number of prior "persons offenses" was also computed for each offender. Over a third of the defendants had a history of at least one such offense. While 40 percent of this group had committed only 1 such offense, the scores ranged in value from 1 up to 22.

The association between the summed seriousness scores for current convictions and the total number of current persons offenses was, as might be expected, quite strong (+.67); the same was true for the summed seriousness scores of prior convictions and the numbers of prior persons offenses (r=+.76). There was only low correlation, however, between the number of current person offenses and the number of prior person offenses (+.19) -- that one has committed persons offenses in the past is not highly associated with the number of persons offenses of which one is currently convicted, or vice versa. Similarly, there were only weak associations between the seriousness scores for current charges or convictions, and the seriousness scores for prior offenses.

A little over 60 percent of the sample had never been incarcerated prior to the current offense. Of those 544 offenders who had had one or more prior terms of incarceration, the average number was 4.9. Fifteen percent of the sample (or 202 defendants) had been released one or more times on parole.

It was possible using the data present in the Massachusetts sentencing study to classify defendants as to whether or not they had previously been convicted of charges for several specific types of offenses. Dichotomous variables were created that indicated whether any one (or more) of the offenders' prior convictions involved a robbery charge, a sex offense charge, a burglary or breaking and entering charge, an assault charge, or a first degree murder charge. If an offender was convicted of an offense that could be classified in one of these crime specific groups, the offender was given a score of +1 for that item. Without making value judgements as to the "most serious" of the current conviction offenses, we could thus look in a rough-and-ready way at the patterns of the sample's criminal careers.

Almost a fourth of the defendants had been convicted of at least one charge of robbery, and one out of every five defendants had previously been convicted of at least one charge of burglary or breaking and entering. Over a third of the defendants had been convicted of one or more charges of assault; six percent of the defendants had been convicted of one or more sex offenses; and 1 percent of the sample (or 14 offenders) had been convicted of one or more charges of first degree murder.

Information on Injury to the Victim

The Massachusetts sentencing guidelines project collected information on the amount of injury to the victim for each of the specific current offenses charged. As we pointed out earlier, the Massachusetts guidelines themselves call for the injury scores assigned to each conviction to be summed and the total score is then used in computing the guidelines sentence. We attempted to duplicate this total injury score using the data available from the sentencing project. However, since the Massachusetts project coded the level of injury according to a scheme which allowed for five degrees of injury ascending in severity, and the guidelines developed by the Massachusetts project coded injury according to a different six degree scale of severity, we can not be sure that the reproduction represents the same scale as that used in constructing the guidelines.

The best we can do is to generate a total injury score that sums the individual injury scores (as coded) corresponding to each current charge. When this procedure was carried out for each case in the sample, it was found that the scores ranged in value from zero for three out of every four offenders to a high score of 27 injury points.

The victim was known to the defendant in a little over 17 percent of the cases. It should be noted that for the purpose of this analysis, police were considered to be strangers to the defendant, as information regarding specific police-defendant relationships was not available.

Predictably, victims of crime were known to the defendant more often in some types of cases than in others. Only 11 percent of the burglars and 12 percent of the robbers knew the victims of their offenses, while half of the convicted first-degree murderers and 58 percent of the sex offenders were familiar with the victim of the crime. Almost 40 percent of the offenders convicted of assault knew their victims.

Although in a total of 247 cases the victim was known to the offender, victim-precipitation of the offense was indicated in less than a quarter of those cases or four percent of the total. Again, some variation was found in the amount of victim precipitation for specific types of offenses: as would be expected, it was much more common in assaults and sex offenses than in property crimes (including robbery). Victim precipitation of the offense was not indicated for any of the first degree murder cases, however.

Several items of information collected by the Massachusetts project pertained specifically to the injuries suffered by victims of sex offenses. The six percent of the defendants listed as sex offenders were convicted of at least one offense of either rape, incest or sodomy (or combinations thereof). Penetration was completed in over three-fourths of the cases where convictions of sex offenses occurred and sodomy was committed in about 20 percent of the total sex offense conviction cases. Also, 10 percent of the 88 sex offense conviction cases involved "sex organ torture" and 1 percent indicated "permanent sex organ injury" to the victim. In short, as such cases go, these were a fairly serious bunch.

Weapon Use Information

Using the same summary strategy as that used in the duplication of the guidelines injury score, three summary variables pertaining to weapon use were also created by us. The first of these variables sums scores across all charges for each instance of use of any weapon during the commission of the offense. (The original scores assigned to charges have been recoded 0=no weapon use, or 1=weapon use.) The majority of the defendants (840 offenders or 58 percent of the sample) received a weapon use score of zero, indicating that a weapon had not been used in any of the offenses that they had been charged with. Another fifth of the offenders had used a weapon in at least one offense, while 11 percent had used a

weapon twice; the remaining weapon use scores ranged from 3 to 27. This extreme skewness is not, of course uncommon in criminal justice data. But in this case it illustrates an acute problem for regression-like guidelines (and the regression analyses on which they and other guidelines may be based). Whatever weight is estimated or assigned to this variable, it is highly unlikely that it would be given the same weight by judges for each of 27 instances as it would be for only one instance of weapon use.

At various times during the development of the Massachusetts sentencing guidelines, as has been mentioned earlier, certain kinds of weapons (such as a gun) were distinguished from others (such as a knife). Using available data pertaining to the offender's use of various types of weapons for all of the charges, two additional weapon use variables were created. The first of these creates a total score for use of a gun specifically during the commission of any of the offenses charged; the second item computes a total score for the use of any sharp or dangerous weapon summed across all of the offenses charged.

The figures for gun use and sharp or dangerous weapon use by themselves show that the overwhelming majority of the sample did not use either of these types of weapons. Only about one out of every ten offenders used a gun, and only a quarter used a sharp or dangerous weapon, in the commission of any offense charged. The scores for gun use ranged in value from zero to 14, while the scores for sharp or dangerous weapon use ranged from zero to 13.

Judges' Sentence Decisions and Prosecutors' Recommendations

The sentences given to defendants convicted in the Massachusetts Superior Court can be described by several items of information present in the available sentencing data. These items include: whether the offender was incarcerated, the length of a direct sentence to incarceration, the length of a direct sentence to probation, the length of a suspended sentence (and, if that sentence is accompanied by a sentence to probation, the length of the associated probation term), the length of a sentence to probation that is to follow some time in incarceration, and the length of a suspended sentence (if any) that is to follow some time to be spent incarcerated.

Prosecuting attorneys throughout the Commonwealth of Massachusetts offered recommendations to the judiciary as to what they thought would be the appropriate disposition for the offender in a number of cases. During our fieldwork in Massachusetts we observed that the recommendation of the prosecutor can carry substantial weight in influencing the

judge's final sentence decision: analysis of the data confirms this.

Over half of the offenders in the Massachusetts sample were incarcerated as a result of their conviction in Superior Court. This figure represents 774 offenders; the remaining defendants were not incarcerated, but received a suspended sentence, probation, fines, or some combination of these.

Prosecuting attorneys made sentence recommendations in 506 cases, or slightly more than one-third of the total. In about three-fourths of these 506 cases, the prosecutor's recommendation was for incarceration (either in jail or in prison). An incarcerative sentence was called for in three out of every four instances. When no recommendation was made by the prosecutor, the judges incarcerated about 50 percent of the time; however, when the prosecutor did offer a specific recommendation, judges tended to follow it. Thus in 70 percent of the cases where the prosecutor recommended incarceration, this was what the judge did; perhaps more importantly, in 84 percent of the cases where the prosecutor's recommendation was for a non-incarcerative sentence, that kind of sentence was given.

The average length of term to parole eligibility, for those incarcerated, was approximately 31 months. The average length of incarceration recommended by prosecutors was 34 1/2 months (based on a total of 295 cases). The average length of suspended sentences (received by 334 offenders) was approximately 11 months. Again, the recommendations of the district attorneys for suspended sentences (in the 205 cases where these were made) corresponded rather closely with the average length of those imposed: the average suspended sentence length recommendation was approximately 13 months.

District attorneys also made specific recommendations about the type of sentence, the place where the sentence should be served, and so forth, in the majority of those cases where any recommendation at all was made. In some instances, the district attorney's recommendation was that some time be served in jail or prison, followed by an additional term that would be suspended. (Often, the suspended term was also to be accompanied by probation.) The rationale behind a sentence of this kind (we were told by both prosecutors and judges) was that following a term of incarceration, the offender could be returned to jail or prison for a longer period of time if he or she did not satisfy the conditions of probation, or if new offenses were committed. This particular combination of sentences was recommended by prosecutors in only 11 cases in the sample and the average length of the suspended term recommended after the direct sentence to incarceration was 30

months. Judges actually made use of this type of disposition for 66 of the sample's cases, however; and the length of time suspended following direct incarceration was much higher on average -- 61 months -- than that suggested by prosecutors.

Probation was recommended as the only sentence in no more than 29, or six percent, of the 506 cases in which any recommendation was made; and the average term of probation suggested was 31.7 months. A direct probation sentence was the only sentence disposition given by judges to 234 offenders, however; the average length of these probation terms was 35 months.

Probation was also recommended as the second part of a suspended sentence for 60 of the 506 cases; the average probation term recommended for these cases was approximately 30 months. The judges used this disposition in 343 cases, however; and the average length of probation terms for these cases was 28 months. Again, the judges were -- as might be expected -- more lenient than prosecutors.

A term of probation -- on average 23 1/2 months -- that would follow some amount of time served in jail or prison was recommended by prosecutors in 40 cases. Judges actually gave 269 offenders this type of a sentence, however; and the average length of supervision for these cases was 31 months.

Another sanction often used by Massachusetts judges is the ordering of court and other costs to be paid by the offender. The Massachusetts sentencing data show, first, that costs had been charged to over a fourth of the sample in connection with a previous conviction; and, for the 389 cases where information was available, the amount charged in connection with current convictions averaged about $475. (It is not clear from the information supplied to us, however, whether this figure represents the average amount of costs that were assessed or the amount actually collected.)

As Table 8.1 shows, judges generally agreed with the prosecutors' recommendations regarding the specific institutions to which offenders should be sent. For example, over half of the prosecutors' recommendations of direct sentences to the House of Correction were accepted by judges. (The underlined percentages represent concordant sentences.) As this table also shows, 65 percent of the prosecutors' probation recommendations were followed by the judges.

Insert Table 8.1 here

Table 8.1: Sentence Decision by District Attorney Recommendation

Sentence	District Attorney Recommendation												Total	
	House of Correction				Concord				Walpole				Probation			
	Direct		Suspended		Direct		Suspended		Direct		Suspended					
	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N
House of Correction																
Direct	56	94	12	6	20	20	-	-	16	20	14	1	13	3	30	144
Suspended	30	50	51	25	6	6	47	7	2	2	-	-	13	3	19	93
Concord																
Direct	2	3	-	-	40	40	-	-	18	22	-	-	4	1	14	66
Suspended	3	5	-	-	9	9	40	6	2	2	-	-	-	-	4	22
Walpole																
Direct	0	1	6	3	3	3	-	-	60	72	14	1	4	1	17	83
Suspended	1	2	2	1	4	4	-	-	2	3	57	4	-	-	3	14
Probation	8	14	29	14	18	18	13	2	1	1	14	1	65	15	13	65
Total	100	169	100	49	100	100	100	15	100	124	100	7	100	23	100	487

We have already noted that defense counsel in Massachusetts also offer recommendations to the judge as to the appropriate sentence for their clients. Unfortunately, it was not possible for us to assess the extent of agreement between defense counsel's sentence recommendations and the judges decisions, for the Massachusetts sentencing guidelines project did not collect that information.

The Decision to Incarcerate

Just over half of the defendants in the Massachusetts sample were sentenced to serve some amount of time in jail or prison, as the primary disposition of one or more of the charges of which they were convicted. Yet, when the characteristics of the sample -- in particular, the offenders' current and prior relations with the criminal justice system -- are reviewed in light of the incarcerative disposition, the incarceration rate fluctuates dramatically.

To begin with, the judiciary favored sentences that did not involve incarceration (including fines, probation, and suspended sentences) for two thirds of the female offenders, but for less than half of the male offenders. Sanctions that did not involve any incarceration were also used more frequently for white offenders (half of the cases receiving such a disposition), than for either black or hispanic defendants (about a third each). Both of these sets of differences were statistically highly significant.

Defendants living in areas classified by the Massachusetts sentencing guidelines project staff as "low income" or "depressed" tended to be incarcerated more often than did other defendants: two of every three residents of these areas went to jail or prison. By contrast, only about one of every five "upper middle class" area defendants was incarcerated.

More importantly, differential incarceration rates were also evident when that decision was considered in light of the county in which the offender was sentenced. Suffolk County and Hampden County had higher incarceration rates than did all of the counties combined (about 60 percent each, as compared to a general average of 54 percent). It should be remembered, from our earlier discussions, that these two counties had especially large minority (black and hispanic respectively) populations; and presumably they also had higher crime rates, and perhaps different crime patterns, though it is difficult, given the sample size, to demonstrate this statistically. Table 8.2 shows less between-county variation than has been found in other states (see, for example, the discussion of Pennsylvania in Chapter 9 below). A naive reading of this

table, however, might suggest that it showed disparity in sentencing. This possibility cannot be ruled out; but it cannot be assumed to be correct either, unless other between-county differences (e.g. in racial composition and crime patterns) are carefully controlled.

Insert Table 8.2 here

Four out of every five defendants who went to trial received a sentence of incarceration (either to the House of Correction, or to Concord or Walpole); this was so, however, for only half of those offenders who pleaded guilty. These statistics become even more interesting when the type of trial is considered. In Massachusetts, a defendant may opt for a trial either before a judge or before a jury. When the defendant's trial was before a judge, slightly less than half of those convicted received a jail or prison sentence; when the trial was before a jury, however, slightly more than 80 percent of those convicted were incarcerated. In part, of course, this also reflects variation in the decision to go to trial among those charged with different offenses: trials were more likely among those eventually convicted of first degree murder and sex offenses (57 and 32 percent, respectively) than among burglars (of whom only nine percent went to trial). Many factors -- such as a long prior record, or consideration of the evidentiary "strength" of the case -- may influence defendants' decisions to plead or go to trial; and we cannot conclude, without more data than we have, that the fact of going to trial by itself influenced incarceration rates.

Incarceration rates also varied significantly, of course, depending upon the type of conviction offense. Approximately 60 percent of both the assault and burglary offenders were sent to jail or prison, as were almost three-fourths of those convicted of sex offenses and robbery. At the other extreme, only about half of those convicted of theft or fraud were incarcerated.

Again, predictably, the number of persons offenses in the current charges had a strong monotonic effect on the chance of incarceration. The likelihood of incarceration increased even more dramatically for offenders who had been convicted of persons offenses prior to the current offense. Only 46 percent of the offenders who did not have any prior persons offenses were incarcerated; with five or more such convictions, the chance of jail or prison was 88 percent.

Table 8.2 : Incarceration Decision by County

County	Not Incarcerated %	N	Incarcerated %	N	Total %	N
Berkshire	50.0	(15)	50.0	(15)	100	(30)
Bristol	55.7	(49)	44.3	(39)	100	(88)
Essex	58.2	(71)	41.8	(51)	100	(122)
Hampden	40.6	(76)	59.4	(111)	100	(187)
Hampshire	42.9	(9)	57.1	(12)	100	(21)
Middlesex	51.1	(137)	48.9	(131)	100	(268)
Norfolk	48.7	(57)	51.8	(60)	100	(117)
Plymouth	45.7	(37)	54.3	(44)	100	(81)
Suffolk	38.5	(134)	61.5	(214)	100	(348)
Worcester	45.5	(81)	54.5	(97)	100	(178)
Total	46.2	(666)	53.8	(774)	100	(1440)

Chi square = 24.08 is significant at the .0042 level using 9 degrees of freedom.

Prosecution of the offender (as a "career criminal") by the Major Violator Division of the district attorney's office greatly increased the likelihood of incarceration: of those prosecuted in this way, 84 percent were incarcerated, against 38 percent of those not dealt with in this fashion. On the other hand, whether the defendant made efforts toward "self-rehabilitation" before the case was disposed of also appeared to reduce the probability of a jail or prison sentence, although not strongly. Among defendants who had made no attempt at rehabilitation before coming to court, 57 percent were incarcerated; but only 44 percent of the offenders who had made some such effort were incarcerated. That said, we must add that it is not entirely clear what this variable in the Massachusetts data means; presumably such efforts (which occurred in guilty plea cases) played some role in plea negotiations, however.

Of much more significant impact on the incarceration decision was whether or not the offender was "at liberty" -- i.e. not incarcerated, either awaiting trial or serving another sentence -- at the time of the case's disposition. (Of course, whether or not an offender is incarcerated prior to trial must also affect that defendant's ability to make an effort toward "self-rehabilitation"; so the two items are not unrelated.) In comparison with 37 percent of the defendants at liberty, 81 percent of those defendants who were already incarcerated at the time of their trial or plea received a further sentence to jail or prison.

Just as the number of current and prior persons offenses affected the judges' incarceration decision, so too did the number of prior incarcerations. As the number of times that a defendant had previously been incarcerated increased, so, generally, did the proportion of defendants incarcerated for the current offense. The chance of a jail or prison sentence rose from 41 percent for those with no prior incarceration, to 89 percent for those with eight or more prior jail or prison terms; the increase is not quite monotonic, though this undoubtedly reflects small numbers in some categories.

Length of Jail and Prison Terms

The decision to incarcerate has been said by Wilkins to be the first in a "bifurcated" or two-stage decision process -- the second part of which is a decision about the length of incarceration for those sent to jail or prison. As we noted at some length in earlier chapters, this representation of the sentencing process -- if intended as a psychological account of how judges actually decide what to do to offenders -- is far from being a well-supported theory. In fact, there are a great many alternative decision-making models which may more

accurately describe judicial behavior in the sentencing process.

That sentence length, more than a binary choice between incarceration or not, may be the most important decision made by judges when sentencing is emphasized by the structure of the Massachusetts sentencing guidelines. That structure required that judges first compute the estimated amount of time that the offender should (or would) serve, given his particular scores on various items included in the sentencing guidelines. The question of whether the offender should be incarcerated at all is not the first decision that the judge is required to make, when using the Massachusetts sentencing guidelines. Rather, the offender's scores on the various items included in the guidelines are multiplied by weights that supposedly reflect the amount of time in jail or prison that should be given for each item; these weighted items are summed, to provide the total estimated time -- the "guideline sentence" -- to be imposed. This estimate is supposedly based on sentences given to similar offenders (in terms of the guidelines factors) in the past. It is not until after the estimate of time to be incarcerated is derived that judges are, according to the guidelines' structure, even given the opportunity to think about the "in-out" decision.

The "in-out" decision is thus clearly a secondary one, in terms of the structure of the Massachusetts sentencing guidelines. The calculation of a time estimate in the first instance appears to create a strong presumption toward incarceration. If, under the most recent version of the Massachusetts sentencing guidelines, the offender's attributes yield a guideline sentence of less than 6 months, the judge is allowed to impose an "out" sentence; but the only presumptive "out" sentences would seem to be those with a score of zero. For guidelines sentences of more than six months, the presumption is that some amount of time spent in jail or prison should result. And given the items included in the Massachusetts sentencing guidelines (current and prior offense seriousness, weapon use, and injury to victim) and their respective weights (2.1 months, 1.6 months, and 9 months each), virtually no one other than a first offender convicted of a trivial offense would presumptively receive an "out" sentence, unless the case is treated as a departure.

Thus, the element of sentence length, as incorporated in the Massachusetts sentencing guidelines, is of major importance. In this respect, the Massachusetts guidelines differ markedly from those developed in other jurisdictions to date. (We discuss some of these in Chapters 4 and 9.) Though we will examine predictors of sentence length in some detail in the analyses we will present a little later on in this

chapter, we note at this point two important factors that are associated with sentence length in the Massachusetts data. The first of these is method of conviction -- in other words, whether the offender was convicted after a guilty plea or after a trial; the second is the county in which the offender was convicted.

These factors are important for several reasons. We noted earlier in this chapter that the method of disposition (trial or plea) was strongly related to the "in-out" decision; it is thus reasonable to assume that it also affects the length of terms imposed. It should be recalled as well that the Massachusetts sentencing guidelines (in their final form) do not differentiate between cases disposed of by a plea of guilty and those disposed of after a trial, when it comes to determining the appropriate sentence that should be imposed under the guidelines. Thus, if any real differences do exist in the lengths of sentences for offenders disposed of by different methods, those differences will be obscured or, at best, mis-estimated in some direction, when the guidelines are used uniformly on all future cases. Furthermore, it is at least possible that judges consider entirely different things when sentencing tried offenders and those who plead -- or that different weights are attached to similar factors.

We next examine the lengths of sentences in the Massachusetts data in light of the county where defendants were sentenced. Are there any substantial differences in sentence lengths across counties which cannot be explained by crime rates or patterns, etc.? If such differences exist, they could hold serious implications for the sentencing policy that would be established in a set of statewide sentencing guidelines. Differences in sentence lengths caused by county of disposition would be obscured or, again, mis-estimated, when the one guidelines model was used uniformly on all future cases. We shall see that this problem is a very general one, in the construction of "empirically based" guidelines; we believe that the Massachusetts researchers did not cope with this problem, but it is fair to note that they are not alone.

Sentence Length and Method of Disposition

The sentence lengths imposed after pleas of guilty and after trials are compared in Table 8.3. The distribution of sentence lengths for trial cases appears, from this chart, to be roughly equivalent to the severity of sentences imposed after a guilty plea with one important exception -- the right-hand tail of the distribution of sentence lengths in trial cases indicates that some of these cases received sentences of a much greater severity than did the plea cases. Indeed, nine percent of the trial cases received a sentence

equal to or more than 12 years in estimated time to be served until parole eligibility, while only about two percent of the plea cases received sentences of this length.

Insert Table 8.3 here

We do not mean to suggest that these results indicate that offenders convicted after a trial receive more severe sentences than do offenders who plead guilty, only because they exercised their right to a trial. Although that may be one cause of the discrepancy (as the literature in the field, as well as some Supreme Court cases, acknowledge), we do not think that the problem is quite that simple. Rather, a number of other facts about the defendant, or the offense, which discriminate trial from plea cases generally may contribute to this evidence of sentence severity variation. As we pointed out in earlier sections of this chapter, there is some evidence to suggest that there are legitimate differences between trial and plea disposition offenders, and the offenses that they commit; and these differences may result in legitimate differences in sentence lengths, though this cannot be clearly demonstrated from the Massachusetts sample. The differences can probably be summarized generally by the statement that offenders who choose to go to trial are often charged with more serious crimes, and have aggravated offense descriptions and prior records, though there are no doubt other factors involved as well. It follows that these offenders, if convicted at trial, would be likely to get more severe sentences than other offenders not displaying such dismaying attributes.

The point that we wish to make here about the effect of method of disposition on length of defendants' sentences is that even a cursory analysis (such as that above) suggests that there is a need to look separately at trial and plea cases, in the first instance, when analyzing data prior to constructing sentencing guidelines, in Massachusetts or elsewhere. The Massachusetts project did, in fact, do this; the end result was that they initially proposed that different ranges for determining acceptable guidelines sentence lengths be used for each type of case. However, the work of the Massachusetts project proved to be in vain, as the judiciary in that state ultimately chose to use the sentence length ranges based primarily on the trial guidelines for plea cases as well. We could expect, from the distribution of sentence lengths in trial cases presented above, that the effect of that decision would be to overstate the suggested guidelines sentences where plea cases are concerned. (In a later section of this chapter we will examine this issue more closely, when

Table 8.3: Time Incarcerated Until Parole, by Type of Trial (Incarcerated Cases Only)

Time	Plea %	Plea N	Trial %	Trial N	Total %	Total N
Up to 3 months	19.7	117	16.6	29	19.0	146
Over 3 to 6 months	18.8	112	16.0	28	18.2	140
Over 6 to 9 months	4.0	24	2.8	5	3.8	29
Over 9 to 12 months	16.7	99	14.3	25	16.1	124
Over 12 to 18 months	12.8	76	10.3	18	12.2	94
Over 18 to 24 months	6.2	37	4.6	8	5.8	45
Over 24 to 30 months	4.7	28	4.6	8	4.7	36
Over 30 to 36 months	1.8	11	1.7	3	1.8	14
Over 36 to 48 months	3.9	23	5.1	9	4.2	32
Over 48 to 60 months	4.0	24	5.7	10	4.4	34
Over 60 to 72 months	1.3	8	4.0	7	2.0	15
Over 72 to 84 months	2.4	14	1.7	3	2.2	17
Over 84 to 96 months	.7	4	1.7	3	.9	7
Over 96 to 108 months	.2	1	.0	0	.1	1
Over 108 to 120 months	.3	2	1.7	3	.6	5
Over 120 to 132 months	.0	0	.0	0	.0	0
Over 132 to 144 months	.7	4	.0	0	.5	4
Over 144 months	.5	3	3.4	6	1.2	9
Life (15 yrs)	1.0	6	2.8	5	1.4	11
Life (not 15 yrs)	.2	1	2.8	5	.8	6
Total	100	594	100	175	100	769

we estimate the proximity of the guidelines sentence to the actual sentence imposed in each type of case.) We ought to add, at this point, that this blunder -- for that is what it is -- is probably not the fault of the Massachusetts guidelines research team. Staff memoranda (which they generously provided to us) show that they tried to make this point clear.

Sentence Length and County of Disposition

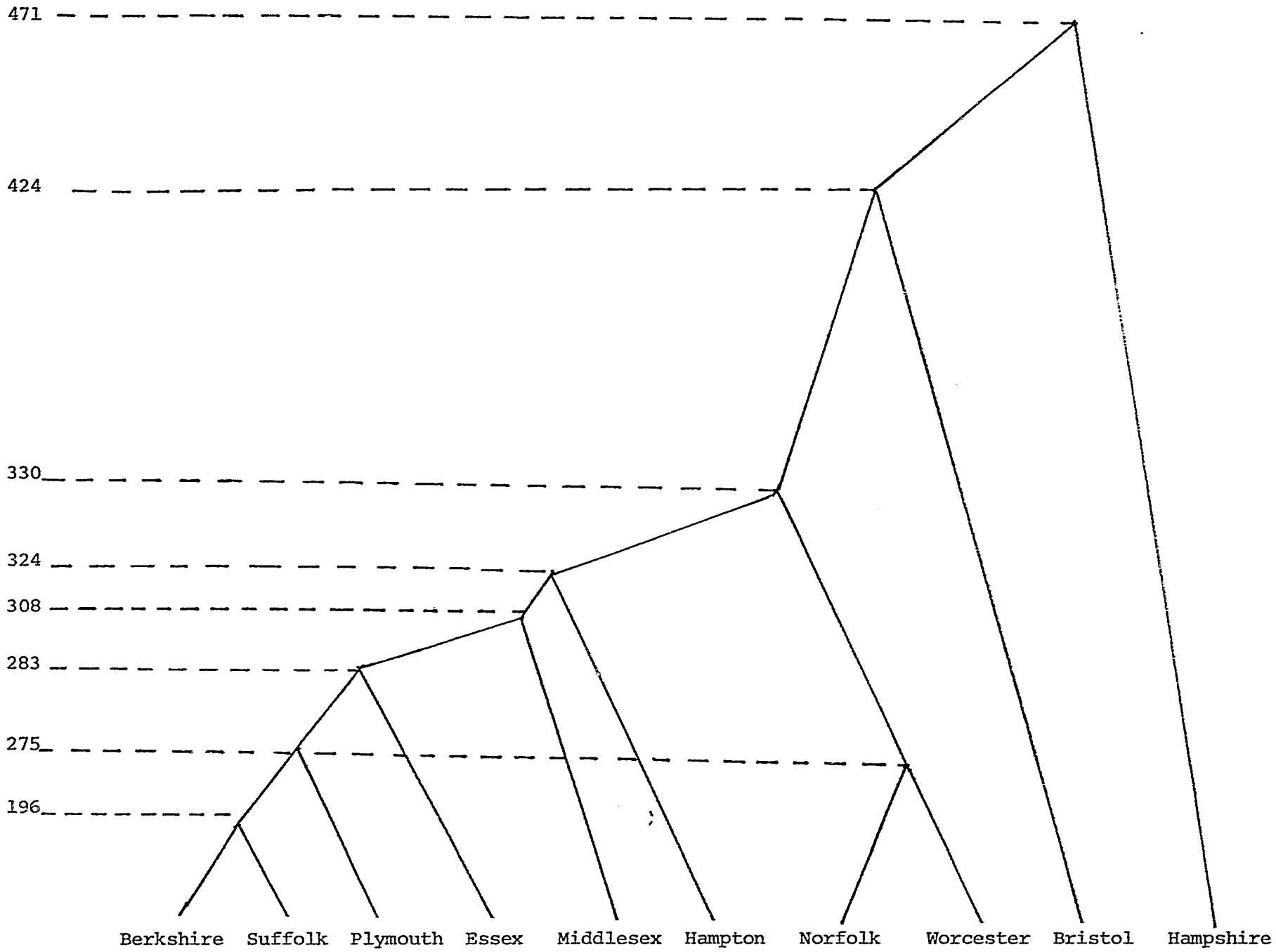
Figure 8.1 presents the results of a hierarchical clustering analysis of the Massachusetts counties in terms of their distributions of sentence lengths. The hierarchical clustering technique seeks to identify how similar the distributions of sentence lengths in each county are to the other counties' distributions, using as the measure of "similarity" the Kolmolgorov-Smirnov statistic. The scale values presented on the left side of the figure are thus scale values of similarity, equal to the cumulative relative frequency distribution for the county groups being compared. (A zero scale value would mean, for instance, that the county groups being compared displayed exactly similar sentence length distributions.)

Insert Figure 8.1 here

The graphics of Figure 8.1 make it immediately apparent that the Massachusetts counties are not all that similar to each other in terms of the lengths of sentences that are imposed in each. The two most similar counties, in the sense just explained, are those of Berkshire and Suffolk. This finding is a little surprising, since Berkshire is the smallest and most rural of all of the Massachusetts counties, while Suffolk (the home of Boston) is the largest and most urbanized of the counties. Both counties have fairly low incarceration rates. But this is presumably for different reasons: for instance, Berkshire does not have much crime, in comparison to Suffolk, but both areas are limited by the facilities available to them for incarcerating offenders. Thus, even though Suffolk County has a higher crime rate, and probably has more serious crime than other areas of the state, the lack of available prison facilities places a "cap" on the number of persons that can be incarcerated in that county.

Plymouth is the county most similar to Berkshire and Suffolk (taken together) in terms of its sentence length, or sentence severity, distribution. As can be seen by the dotted line which extends from the scale values over to the right of the table, Norfolk and Worcester counties are also similar to

Figure 8.1: Hierarchical Clustering Scheme Based On Distributions of Sentence Severities in Massachusetts Counties



each other, at about this level. Though Norfolk and Worcester counties are similar to each other at the same value as Berkshire, Suffolk, and Plymouth counties are, they are not similar to Berkshire, Suffolk, and Plymouth counties. In other words, the two groups of counties evidence quite dissimilar patterns of sentence severity from each other, though the distributions of sentences for each of the counties within each group are similar to the other counties in that group. Finally, Bristol and Hampshire counties display the most dissimilar patterns of sentence lengths from the other counties. They join the clustering scheme last, having scale values of the severities of their sentence distributions calculated at .424 and .471 respectively.

We should emphasize that this figure does not provide a statement about the relative severity of sentences in one county versus another county; rather, a statement is made about the distributions of sentences -- from lenient to severe -- in each county, as compared to other counties. In other words, where any similarity between counties does exist, it is a similarity in the sense that the entire range of sentences imposed in one county is similar to the range of sentences imposed in the second "similar" county.

The finding that the distributions of sentence lengths across counties in Massachusetts are not similar to each other can raise some important questions for those who intend to construct sentencing guidelines for statewide use. Should the pattern of sentence lengths found in Suffolk county, for example, be the pattern used in the future through stipulation in a guidelines model? The Massachusetts sentencing guidelines project, and the Massachusetts judiciary, managed to avoid direct confrontation of this issue -- by stipulating that one policy would be in effect, and that the guidelines would not take into account regional differences in crime or correctional resources. (We discussed this point in more detail in Chapter 7, and consider it in the case of Pennsylvania in Chapter 9). However, the incorporation of a particular pattern of sentence lengths into the guidelines may result in differences in the future allocations of persons to correctional institutions throughout the state -- a point which that project seems not to have considered. Had they thought of it, only the use of the guidelines in specific cases for some time to come will show the actual effects of the use of any particular pattern of sentence severities on Massacusetts' state and local correctional populations.

Replicating the Massachusetts Sentencing Guidelines

We now turn to the question of how well the sentencing guidelines developed in that state actually fit the data on

which they were supposed to be based. The purpose of this analysis is two-fold. We wish first to determine exactly how much of the variance in sentencing decisions in the construction data can be explained, using the factors described by the Massachusetts sentencing guidelines project as "predictors" of those decisions, and prescribed for future use; secondly, we are interested in documenting the proportions of the Massachusetts sample that received sentences within the plus-or-minus fifty percent range specified by the guidelines, and ranges of other magnitudes.

Four factors, as we have already noted, are included in the final version of the Massachusetts sentencing guidelines. These are: (1) the seriousness of current convictions and (2) prior convictions, (3) the degree of injury to the victim, and (4) the use of a weapon during the commission of the offense. Using the information provided to us in the Massachusetts sentencing data, we have first calculated each offender's score for each of these guidelines factors, according to the procedures used in the Massachusetts sentencing guidelines themselves. Once each offender had been assigned a score for each of the guidelines factors, those scores were then used as independent variables in regression analyses of (1) the "In-Out" or incarceration decision, and (2) the length of sentence decision.

The results of these analyses were a bit surprising, in two respects. First, the amount of variation in the sentence decisions that we found to be explained by the Massachusetts guidelines factors was much less than we had expected.[6] Secondly, the weights assigned to the variables by the regression equations did not, on first inspection, appear to match those weights assigned to the same factors in the Massachusetts guidelines.

Of course, we did not expect that the empirically-derived weights for those factors would exactly match the ones used in the guidelines; apart from questions of rounding, standard errors, etc., the Massachusetts research team (and those members of the judiciary who were responsible for overseeing their work) were quite explicit about the fact that they modified the project's empirical findings on grounds of policy (in a broad sense of that term). We shall show, later in this chapter, that this kind of relatively minor tampering with the research team's regression results is not of much consequence; it is certainly much less important than some other errors that (in our opinion) are embedded in the Massachusetts researchers' work.

Table 8.4 presents the results of our regression analyses. As can be seen from the sample sizes indicated for

each of the separate analyses summarized in this table, the entire 1,440-case sample was used for the analysis of the "in-out" decision. Initially, only incarcerated cases were included in the data base for the analysis of sentence length. But the estimates of regression weights derived from this analysis for each of the guidelines factors appeared to be a little strange (about which more will be said in a moment); so, ultimately, we decided to use the entire 1,440 case sample for the length of sentence decision analysis as well, as the Massachusetts researchers apparently had done.

Insert Table 8.4 here

As we have noted in earlier chapters, this approach is -- for a variety of reasons -- mistaken. That a set of sentencing guidelines provides for a direct calculation of "time incarcerated", with zero standing for an "out" sentence, does not show that the multivariate model used to describe previous sentencing behavior can simply set "out" sentences equal to zero, and estimate lengths of terms using all cases. The weights associated with the items included in such a model may have little or no relation to those which would accurately describe length-of-term decisions among the sub-set of offenders who, in the construction data, were actually incarcerated.

Regardless of the data base chosen, or, for that matter, the decision under examination, the Massachusetts sentencing guidelines do not perform very well in predicting the various sentence decisions. The four guidelines factors are collectively able to explain only about 14 percent of the variance in incarceration decisions, and roughly 12-13 percent of the variance in sentence length decisions, depending on which cases are included in the data base. It seems ironic that a 1 percent improvement in predictive ability (in the statisticians' sense of that term) for the sentence length decision can be achieved when non-incarcerated cases are included in that analyses -- since, as we have noted earlier, this is generally a mistake.[7] (Parenthetically, we repeat here a point made elsewhere in this report, which is that R square is not a very useful statistic in this context -- especially when the variable to be predicted is a dichotomous one (such as the "in-out" decision); the significance of regression slopes is far more important. As we shall show later in this chapter, part of the low R square values obtained by the Massachusetts researchers owes to their failure to exclude extreme "outliers" when estimating equations intended to describe the mine-run of cases.

Table 8.4: Massachusetts Sentencing Guidelines

Variable Weights and Predictive Ability

Variables	Incarceration Decision (N=1440)		Sentence Length (N=1440)		Sentence Length (N=794)	
	Estimate (B)	Significance	Estimate (B)	Significance	Estimate (B)	Significance
Seriousness-Current Convictions	.02	.0001	1.26	.0004	1.05	.0538
Injury to Victim(s) Score	.02	.0008	9.54	.0001	11.06	.0001
Weapon Use Score	.02	.0616	2.13	.1184	2.04	.3198
Seriousness-Prior Convictions	.02	.0001	1.34	.0001	1.36	.0001
Intercept	.36	.0001	-1.18	.6119	2.82	.5427
R Square	.148		.134		.124	
F Value	62.41		55.39		27.08	
Significance of F	.0001		.0001		.0001	

Further problems are suggested by our regression analyses of sentence length decisions, particularly by the unstandardized weights associated with each of the guidelines factors in those analyses. One would expect that a secondary analysis that incorporated exactly the same factors as were used to construct "empirically-based" guidelines would produce weights for those factors that were similar to, if not exactly the same as, those included in the guidelines; at least they ought to be in the same ball park. We noted earlier that we first analyzed the ability of the guidelines factors to predict sentence length, using only those cases where the offender had actually received a sentence of some time to be spent incarcerated. Such a decision still seems to us to be correct, even if the Massachusetts researchers thought otherwise; but the factor weights produced did not even approximately resemble those subsequently included in the Massachusetts guidelines. In particular, the weights associated with injury to the victim and with weapon use in the final guidelines model (nine months each) were quite different from the weights produced by our re-analysis of the sentence length decision. Using only incarcerated cases, the weights produced for those items by our regression analysis were 11 months and 2 months respectively. (The remaining two variables also evidenced weighting factors different from those called for by the guidelines, though the differences were not as startling and could possibly have been attributed to minor variations in our preparation of the data for re-analysis.)[8]

Faced with such extreme differences in variable weights, we decided to use the entire 1,440 case sample for the sentence length analysis, coding non-incarcerated cases as zero on the dependent time variable. This analysis not only produced a slightly higher R square value, but also resulted in guidelines factor weights that are more similar to those proposed by the final Massachusetts guidelines model. The weight for the item of injury to the victim, notably, was only half a month different from the guidelines score, while the current and prior offense seriousness weights departed from the guidelines weights by .9 and .3 months respectively. Unfortunately, though this procedure did increase the weight of the "weapon use" factor somewhat, the weight assigned by our regression equation was still seven months less than that called for by the Massachusetts sentencing guidelines.

One additional feature of the Massachusetts guidelines that is somewhat unique, given the method used to calculate sentence length, is that no account is taken of cases in the construction data that are obviously exceptional in either their seriousness or their triviality. Such exceptional cases, as has been often noted, can adversely affect the

predictive power of regression models, and more importantly can lead to a false picture of the weights which should be assigned to variables included in such models. It was surprising to us, therefore, that the Massachusetts sentencing guidelines project did not seem to have considered this problem of exceptional cases at all, during the development of their guidelines models. Clearly, however, this problem needs investigation. The next step in our analysis, therefore, involved examining how closely the (hypothetical) guidelines sentences matched the actual sentences imposed in the cases in the construction sample. By doing this we would be able to identify particular cases where the guidelines "predicted" sentence was clearly more severe, or more lenient, than the sentence that had, in fact, been imposed.

Of course, we expected in the first instance that some sorts of cases would have received sentences that would be markedly different from the sentence called for by the guidelines. In particular, we assumed that where an offender had received a life sentence, something about either the offense or the offender would be unusual in a fairly plain sense of the term. Thus, cases that had received a life sentence without parole (N=6) were excluded from all of these analyses. Even if the defendants in those cases had committed the most serious offense according to the guidelines (i.e. murder), and had similar scores for weapon use, injury to the victim(s), and prior record, the guidelines sentences derived in such cases would still be estimates of time to be served, which is not the same as "life without parole."[9] (Of course, it might have been possible to stipulate that any derived guideline sentence of over $\underline{X}$ months or years would be considered the equivalent of life without parole, but that procedure would almost certainly have introduced great error into our estimates.)

Figure 8.2 presents the residuals from our regression of sentence length on guidelines factors (i.e., the values found after subtracting the observed sentence from the "expected" sentence predicted by the guidelines, plotted against those predicted guideline sentences. To simplify the analysis somewhat, only those cases that actually received a sentence of incarceration were included in this analysis; a plot of this kind for non-incarcerated cases as well (bringing the total case size to 1,440) would have been even more cluttered than this one was.

The purpose of this exercise was twofold. First, we wanted to get a visual picture of how closely the Massachusetts guidelines actually "fit" the Massachusetts sample's sentence data; the number and location of residual values provide important information in this respect. A

second, though related, purpose of the analysis was to locate those cases in which the guidelines clearly did not "fit"; extreme positive or negative residual values readily provide this information in succinct visual terms.

Insert Figure 8.2 here

A number of things are immediately evident from even an eyeball analysis of Figure 8.2. The two most important facts that the figure illuminates about the Massachusetts sentencing guidelines are that (1) none of the cases received a predicted sentence of non-incarceration, and that (2) the sentences called for by the guidelines were generally up to about 75 months of time in jail or prison. Of course, it may be though that, since we used only those cases that had been incarcerated in our analysis, these results are not too surprising; those sample cases incarcerated prior to the guidelines might have been very serious in nature, thus requiring high sentences under the guidelines.

Further analysis of the figure, however, reveals that substantial differences exist between the predicted guidelines sentences and the actual sentences imposed on those cases that had received sentences of incarceration. The extent and nature of the differences between these can be seen clearly. Positive residual values, which indicate that the actual sentence received was more than the guidelines sentence, are shown to the right of the solid vertical center line; negative residual values, which lie to the left of that line, indicate that the sentence called for under the guidelines was more severe than the sentence the defendant had actually received. (For those cases in which the sentence actually imposed was exactly equal to that predicted, the residual is of course zero.) The points plotted within the figure, represented by the letters of the alphabet, show the relative concentration of cases in various areas of the plot. (Each letter stands for the number of cases which mark its place in the alphabet, i.e., A is shown for 1 case, B for two cases, and so on.)

As Figure 8.2 shows clearly, the majority of the residual values in the sample are negative -- indicating that the guidelines sentence "predicted" for the defendant was more severe than the sentence that the defendant had actually received. (An analysis of exact residual values indicates -- what this computer-generated plot observes -- that for 514 of the 768 cases used in this analysis, the guidelines sentence was more than the actual sentence, so the residual was negative.) At first glance, therefore, the Massachusetts sentencing guidelines appear to be "predicting" sentences that

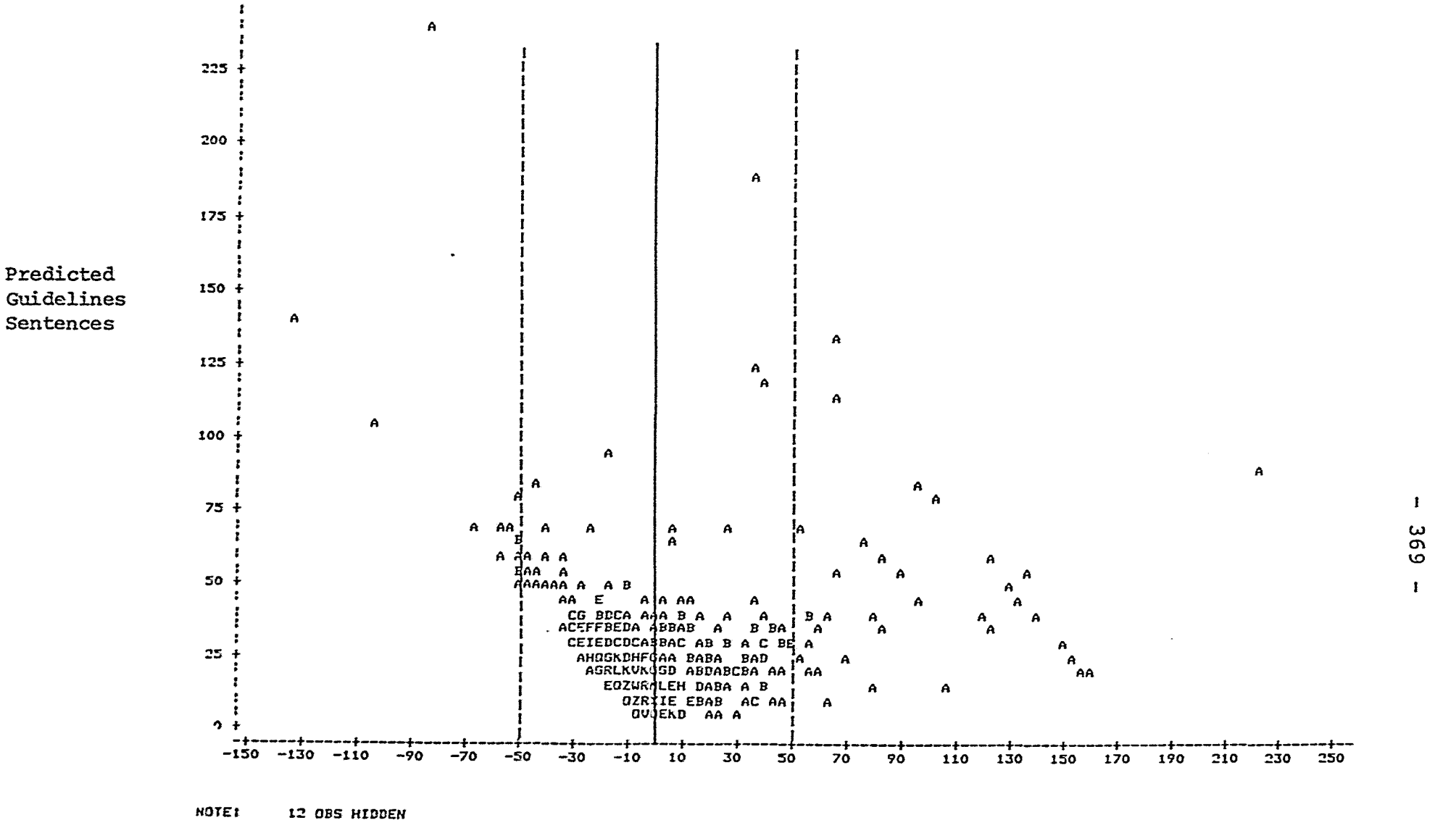


Figure 8.2 : Predicted Massachusetts Guidelines Sentences against Residual Values (Actual Sentences Imposed Minus Predicted Guidelines Sentences), Excluding Life Sentence Cases (N=768)

are very often more severe than those that have normally been imposed in the past. It should be noted, moreover, that since this analysis is based only on those cases in the construction data that received sentences of incarceration before the guidelines -- rather than the whole 1,440 cases sample -- it most probably understates the actual severity of the Massachusetts sentencing guidelines. One could reasonably expect, for instance, that a number of defendants who had actually received non-incarcerative sentences would receive predicted guidelines sentences that would call for incarceration.

To return to Figure 8.2, this plot also shows a number of cases where either the predicted sentence or the actual sentence was extremely different from the other, thus yielding extreme residual values. Most of the cases have residual sentence values of between plus and minus 50 months, with most of those residuals negative. The dotted vertical lines drawn on Figure 8.2 block these cases off; as a result, it can be clearly seen that 47 cases have residual sentence values of more than plus or minus 50 months.

It seems reasonable to assume, at least provisionally, that these cases with excessive residual values are unusual in some respect, and that this was the cause of the unusual. sentence imposed. (We will examine this assumption in more detail later.) On this assumption, we excluded these 47 extreme "outliers" from the data; we then repeated the earlier analysis to determine whether the exclusion of these cases resulted in a better "fit" of the Massachusetts sentencing guidelines to the actual sentences imposed.

Regression of sentence lengths on the guidelines factors, after excluding the cases that had life sentences and the 47 outliers just mentioned, not surprisingly improved markedly the Massachusetts sentencing guidelines factors' ability to explain variation in sentence lengths. Recall that the Massachusetts guidelines factors are able to account for approximately 12 percent of the variation in the 774 incarcerated offender's sentence lengths. After removing only 53 "exceptional" cases from the data base, those same guidelines factors were able to account for approximately 36 percent of the sentence length variation.

That we should be able to increase the proportion of variance accounted for by excluding such cases is not, of course, surprising; nor is it the point. Rather, the point is that if one is using a statistical procedure like regression analysis to describe a set of data like these, it is absurd to retain cases which are wildly deviant from the majority (as the 47 outliers in Figure 8.2 obviously are). Of course it is

better if a reason can be found for such extreme deviations (we consider this point further in a later section). If the object of the exercise is to find a statistical description of what happens in most of the cases, then the accuracy of that description is not likely to be improved by keeping in the analysis cases which are so plainly bizarre in comparison with the majority.

Thus, the improvement in predictive ability of the regression equation using the Massachusetts sentencing guidelines factors was accompanied, as would be expected, by changes in the regression weights associated with those factors. Unfortunately, the end result of the weight changes was not to bring them closer to the weights prescribed by the Massachusetts sentencing guideines, but to increase their differences from the guidelines weights. Table 8.5 provides a comparison of the factor weights produced by the two regression analyses. The new weights provided by our analyses not only indicate that the weights provided by the Massachusetts guidelines are an inaccurate reflection of sentencing "policy" as best that can be reflected in data analysis; they further confirm the fact that the Massachusetts project did not attempt to analyze unusual cases present in their sentencing data. Such cases may of course be evidence of "disparity", judicial insanity, or whatever. They may also reflect the fact that the researcher's model of sentencing practice is simply a bad model -- that it includes the wrong variables, or that the simple linear form of the model is inadequate. Either way, it is misleading in the extreme to include such cases when estimating a statistical model -- and then to claim that the data (in this case, the sentencing practice of the Massachusetts Superior Court) display a good deal of variability that cannot be "explained".

Insert Table 8.5 here

Finally, a second residual plot of the data was prepared, again excluding the 6 life sentences as well as the 47 outliers from the data base: see Figure 8.3. As is evident from this figure, the distribution of residual cases is still heavily weighted to the left, indicating that the guidelines sentences are for, the majority of cases, heavier than the sentence actually imposed; it is, however, slightly more evenly spread than Figure 8.2. The uneven distribution of residuals (a lot of relatively small negative ones, offset by fewer large positive ones) still obtains; and there are some new outliers, though these are not so enormous as the ones in Figures 8.1. Moreover, there is not much of a pattern in the residual plot; e.g. there is little apparent tendency for the

predicted sentence to "miss" the observed one by a greater amount as predicted sentences grow larger.

Insert Figure 8.3 here

Another aspect of their data apparently not explored by the Massachusetts sentencing guidelines project is the possibility of interactive effects between the guidelines factors. The most likely, though not the only, interaction that could be expected between any of those factors would be between the weapon use and victim injury scores. It seems plausible to believe that the use of a weapon, in conjunction with injury to the victim, might influence a judge's sentence decision, perhaps as a reason for aggravating the sentence. In a similar vein, if a weapon was present, but not used to cause injury, this fact might act as a mitigating factor in the judge's sentence decision. In any event, the possibility that a relationship, of whatever direction, might exist between the items is cause enough for testing to see if there is such a relationship.

The product of the offenders' scores for the items of injury to the victim and weapon use was thus added to the four Massachusetts sentencing guidelines factors, and included in a regression equation of sentence length, again using only incarcerated cases. The interaction term was able to improve the predictive ability of the guidelines factors only slightly. Whereas the regression coefficient (R square) for the earlier analysis of sentence length for incarcerated cases (excluding outlier and life sentence cases) was .36, the addition of the injury-weapon interaction term increased the value of the R square to .39. (The coefficient associated with the interaction term itself was statistically highly significant.) Though the improvement may seem slight, it nonetheless points out an additional area overlooked in the development of the Massachusetts sentencing guidelines; the improvement is after all caused by the addition of only one interaction term -- countless others might also have proved to be of value to the Massachusetts sentencing guidelines had they been examined. (We did in fact analyze the interaction effects possible between all of the guidelines factors. We do not put much stock in the analyses, though, since the inclusion of all such interactive terms, as well as the original factors, has no theoretical justification; for the same reason we have not tried to test possible interactions between the guidelines factors and other variables.)

We have thus far focussed our analyses of the Massachusetts sentencing guidelines either on the entire

Table 8.5: Regression of Massachusetts Sentencing Guidelines Factors on Sentence Length for Two Incarcerated Offender Samples

Massachusetts Guidelines Variables	All Incarcerated Cases (N=774)		Incarcerated Cases, Excluding Lifers and 47 Outliers (N=721)	
	Estimate	Significance	Estimate	Significance
Seriousness-Current Convictions	1.05	.0538	.97	.0001
Injury to Victim(s) Score	11.06	.0001	4.98	.0001
Weapon Use Score	2.04	.3198	1.19	.0083
Seriousness-Prior Convictions	1.36	.0001	.52	.0001
Intercept	2.82	.0001	4.27	.0001
R Square	.124		.364	
F Value	27.08		102.70	
Significance of F	.0001		.0001	

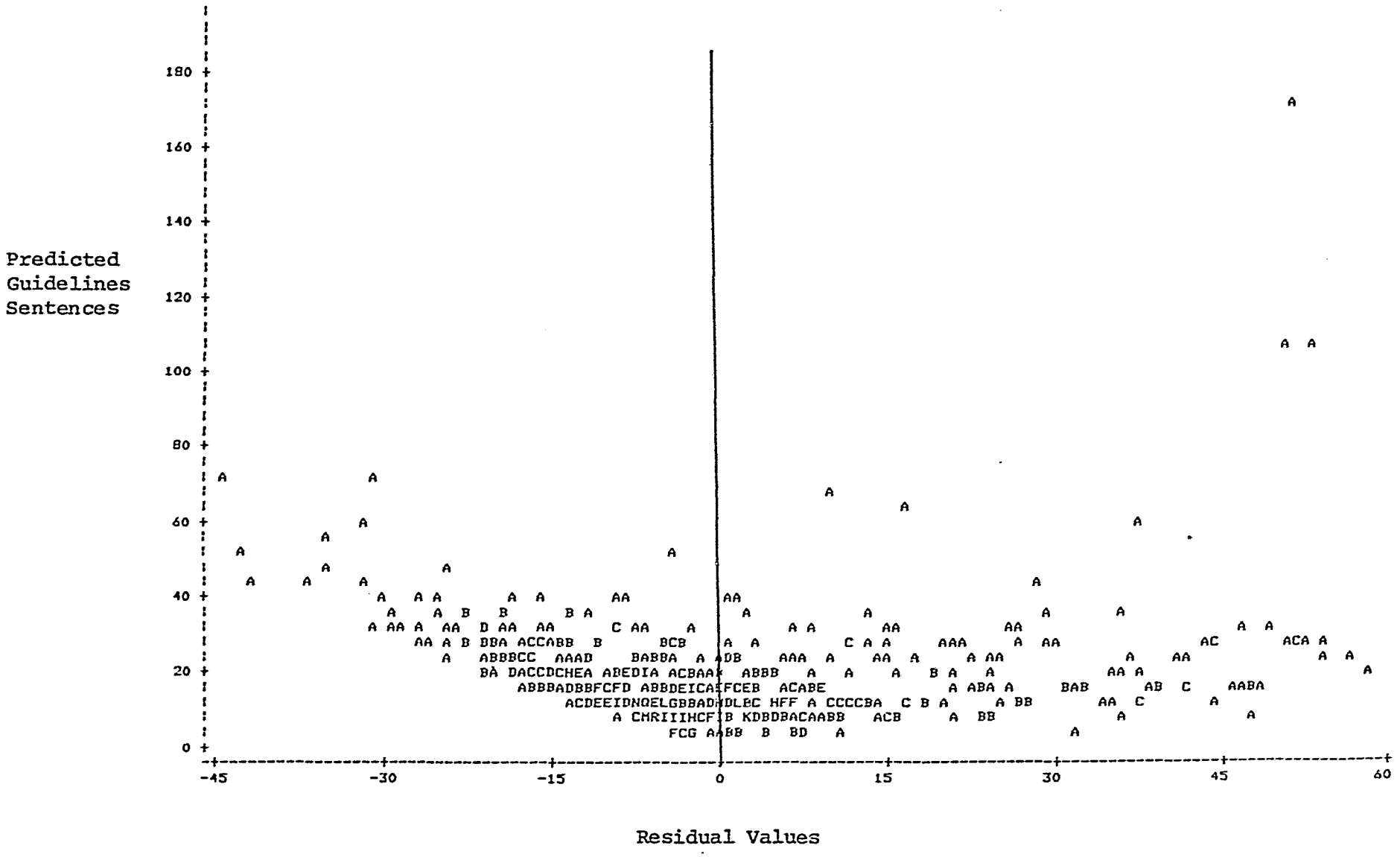


Figure 8.3 : Predicted Massachusetts Guidelines Sentences against Residual Values (Actual Sentences Imposed Minus Predicted Guidelines Sentences), Excluding 47 Outliers and Life Sentence Cases ($\bar{N}^{c}$=721)

sample of cases, or on the entire sample of incarcerated cases. It will be recalled, however, that the Massachusetts sentencing guidelines were initially developed to provide guidelines sentences for cases sentenced after trials, by Superior Court judges. These two early limitations on the application of the sentencing guidelines pose additional questions as to the ability of those guidelines to explain past sentencing decisions. We have already pointed out that the guidelines, promulgated by the Massachusetts sentencing guidelines project and the Superior Court judiciary, do not appear to account very well for the variation in judges' sentencing decisions in all cases -- those in which there was a guilty plea, as well as those tried. As we noted earlier, the guidelines are able to explain only about 15 percent of the variation in the decision as to whether to incarcerate, and can explain only 12 or 13 percent of the sentence length decisions (depending on whether all or only incarcerated cases are used as the data base). Considering the intended application of the guidelines, however, we decided to investigate the ability of the guidelines to explain trial and plea cases separately, and to examine trial and plea cases sentenced by Superior Court judges separately from the rest.

Table 8.6 presents, in summary form, the results of our analyses of the guidelines' ability to account for variation in the incarceration decision, separately for trial and plea cases; and where each type of case was sentenced by a Superior Court judge. We replicated all three of our earlier sets of analyses: first, including all appropriate cases in the data base; next, excluding those cases earlier identified as outlying cases; and finally, including an interaction term for the weapon and injury scores.

Insert Table 8.6 here

The first row of the table presents the results of those analyses that incorporated only the four guidelines factors as predictors of the guidelines sentence. These analyses also used all of the cases in the data base for the regression equations. As can be seen from the regression coefficients produced by the regression equations, the Massachusetts sentencing guidelines are only minimally able to account for variation in the judges' decisions to incarcerate offenders, regardless of the specific sub-sample involved. In fact, the Massachusetts sentencing guidelines do the worst job of explaining the variation in the incarceration decisions for Superior Court trial cases -- that is, the very cases for which use of the guidelines was mandated. And the guidelines factors are only slightly better able to explain the

Factors Included in Regression Equation	All Cases (N=1440)	Trial Cases (N=221)	Superior Court Trial Cases (N=165)	Plea Cases (N=1198)	Superior Court Plea Cases (N=1018)
Massachusetts Guidelines Factors Only					
R Square	.148	.107	.098	.155	.153
F Value	62.41	6.48	4.33	54.87	45.65
Significance	.0001	.0001	.0024	.0001	.0001
Guidelines Factors Only; Exluding Outlier and Life Sentence Cases					
R Square	.171	.146	.148	.188	.189
F Value	71.20	8.22	6.08	67.27	57.37
Significance	.0001	.0001	.0002	.0001	.0001
Guidelines Factors, Weapon-Injury Interaction Term; Excluding Outlier and Life Sentence Cases					
R Square	.173	.146	.150	.190	.191
F Value	57.79	6.56	4.90	54.40	46.65
Significance	.0001	.0001	.0004	.0001	.0001

Table 8.6: Regression Analyses of the Massachusetts Sentencing Guidelines Ability to Explain Variation in the Incarceration Decision for Selected Cases

incarceration decision variation in plea cases than they are able to account for decision variation in plea and trial cases combined; the improvement from an R square of .148 to an R square of .155, however, is negligible.

The second row of the table lists the regression coefficients produced during the second phase of our reanalysis when cases that had received a life sentence without parole, or that had been identified as outliers, were removed from the data base. Though all of the regression coefficients produced by the equations using the guidelines factors on the limited sample show improvement over those produced using all of the cases in the data base, the Massachusetts sentencing guidelines still do not appear to be able to account dramatically for the sentence decision variation.

The final row of Table 8.6 presents the regression coefficients produced by the Massachusetts guidelines factors and the weapon-injury interaction term. Again, there is a substantial improvement in predictive ability of each of these models over the model using only guidelines factors to account for the sentences of all cases in the sample; however, there is only a negligible improvement over the coefficients listed in the second row of the table. The addition of the weapon-injury interaction term does little to <u>further</u> improve the ability of the guidelines factors to determine the incarceration decision outcome; it appears instead that the improvement in predictive ability was most likely the result of the exclusion of odd cases from the data base at the earlier stage of analysis.

Though it is not evident from this table, our separate analyses of trial and plea cases indicated that several of the factors included in the Massachusetts sentencing guidelines are not significant variables in regression analyses performed on each of the groups separately. Two factors consistently showed non-significant values for their contribution to explaining the incarceration decision for trial cases, while one factor was not significant for inclusion in plea case analyses. The offender's scores for injury to the victim and for weapon use proved <u>not</u> to be significant for trial cases. Similarly, the offender's score for injury to the victim was not a significant predictor of the incarceration decision for plea cases. It apeared, from our analyses, that the Massachusetts sentencing guidelines would not have suffered much of a loss of predictive power, at least as far as the incarceration decision is concerned, if these items were to be excluded altogether from the sentencing guidelines.

Somewhat more success was had when the length of sentence decision was analyzed. We noted earlier that the guidelines factors, though initially able to account for only 12 percent of the variation in incarcerated offenders' sentence lengths, were able to account for 36 percent of that variation when outlier cases and life sentence cases were removed from the data base, and 39 percent of the variation when the injury-weapon interaction term was included as an additional factor in the regression analyses. When these same analytical procedures are followed in analyses of trial and plea cases separately, the improvement in the guidelines ability to predict sentence length is similarly marked.

Insert Table 8.7 here

Table 8.7 presents the regression coefficients produced by our three-phase reanalysis of the Massachusetts guidelines for trial and plea cases separately. The regression coefficients indicate that the Massachusetts sentencing guidelines alone are best able to account for variation in the sentence lengths of cases disposed of after a trial; about 18 percent of the variation in that decision can be explained using the guidelines factors. When only trial cases disposed of by a Superior Court judge are used as the sample, the R square for the equation using the guidelines factors drops; only about 14 percent of the variation in Superior Court judges' decisions can be explained by the guidelines factors. Similarly, only about 14 percent of the sentence length variation in cases disposed of after a plea of guilty, whether by a Superior Court judge or not, can be explained using the guidelines factors.

The second row of the table again presents the results of these same analyses replicated on a sample excluding the 47 outliers and cases that received life sentences without parole. Again, the improvement in the predictive ability of the Massachusetts sentencing guidelines is dramatic. The guidelines factors are able to account for almost 57 percent of the variation in the sentence length decisions in trial cases generally, and almost 60 percent of the variation in trial cases disposed of by Superior Court judges. However, the increased ability of the guidelines to explain the sentences of Superior Court judges may suggest that the unusual cases, now excluded from the sample, were actually sentenced by Superior Court judges; once the extreme variation found in the sentences of those cases is removed from the sample, the overall ability of the guidelines to account for the remaining variation improves markedly.

Factors Included in Regression Equation	All Cases (N=774)	Trial Cases (N=175)	Superior Court Trial Cases (N=134)	Plea Cases (N=594)	Superior Court Plea Cases (N=505)
Massachusetts Guidelines Factors Only					
R Square	.123	.184	.138	.137	.140
F Value	27.08	9.55	5.17	23.30	20.31
Significance	.0001	.0001	.0007	.0001	.0001
Guidelines Factors Only; Excluding Outlier and Life Sentence Cases					
R Square	.364	.568	.588	.308	.327
F Value	102.66	47.95	38.98	62.22	57.56
Significance	.0001	.0001	.0001	.0001	.0001
Guidelines Factors, Weapon-Injury Interaction Term; Excluding Outlier and Life Sentence Cases					
R Square	.390	.591	.618	.340	.358
F Value	91.64	41.99	34.88	57.62	52.81
Significance	.0001	.0001	.0001	.0001	.0001

Table 8.7: Regression Analyses of the Massachusetts Sentencing Guidelines Ability to Explain Variation in Sentence Lengths for Selected Cases

The R square coefficients computed for the equations for plea cases also increased significantly after exclusion of outlier and life sentence cases, in comparison to the earlier coefficients. However, the increases, though substantial, were not as dramatic as the increases in the R squares for trial cases.

The final row of Table 8.7 presents the regression coefficients produced when the weapon use-injury interaction term is introduced into the sentence length prediction. The results are fairly consistent across the various types of subgroups that are analyzed here; in general, the weapon use-injury interaction term is able to improve the prediction of judges' sentence length decisions by about three percent regardless of whether the case was disposed of by trial or plea, or by a Superior Court judge.

The relative importance of the various guidelines factors for the specific trial and plea disposition subsamples also showed some differences compared with their performance across those cases combined. One factor was not significant in the regression equations predicting the sentence lengths of trial and Superior Court trial cases; similarly, another factor was not of significance to the prediction of plea and Superior Court plea case sentence lengths. Specifically, the use of a weapon did not increase the ability of the guidelines model to predict sentence lengths in tried cases, both when outlier and life sentence cases were excluded from the sample, as well as when the injury-weapon use interaction term was added to the equation. For plea cases, all of the factors remained in the regression equation as significant contributors even when the outlier and life sentence cases were excluded from the sample. However, the addition of the weapon use-injury interaction term caused the original guidelines factor of injury to the victim to become an insignificant contributor by itself to the prediction of the length of the incarcerative sentence. (This suggests that the interaction term, rather than injury by itself, should probably be deleted for these cases.)

Examining Outliers

In the preceding section we noted that it is possible to increase dramatically the predictive ability of the Massachusetts sentencing guidelines by excluding extremely deviant cases -- those with extreme residual values -- and cases receiving life without parole sentences, from the analysis. We argue that this procedure required no further justification than the fact that the inclusion of such cases will almost certainly lead to mis-estimation of the weights of other factors which may be predictive of sentence length. The extreme values found in a small number of cases could have

resulted (quite justifiably) from any number of things; and in most instances, the justifications would probably be of a kind only rarely present. They can thus distort data analyses based on the "usual" factors thought to be appropriate considerations in reaching sentencing decisions.

Regardless of whether a separate analysis of cases with extreme or unusual residuals can explain the sentences in those cases, we suggest that such a separate analysis is useful. It is possible, after all, that some of the variables in the data set will also display unusual values for those cases -- such as prior record variables, or current offense seriousness variables -- as compared to the average sort of case. Thus, we now look a little more closely at the specific characteristics of the cases that had residual values of plus or minus 50 months or more, when their actual sentences were compared with the sentences called for by the Massachusetts sentencing guidelines.

Table 8.8 presents an overview of the characteristics of the 47 cases with extreme (positive or negative) residuals. It bears repeating at this juncture that a positive residual value indicates that the actual sentence imposed in that particular case was more severe than the sentence called for under the Massachusetts guidelines by the residual amount. Negative residual values, on the other hand, indicate cases where the actual sentence imposed was less severe than that called for by the Massachusetts sentencing guidelines. Cases that had positive residuals are presented first in this table, followed, after the horizontal dotted line, by cases displaying negative residuals.

In reviewing cases where the residual value was positive, one would naturally look for circumstances or factors of aggravation, justifying a heavier-than-usual sentence. The most apparent, and easily identifiable, characteristic of the cases with high positive residual values is that the majority of offenses involved some sort of physical injury, notably sexual assault (and in the first case, homicide). Sexual assault was frequently the offense of conviction for these cases; often this offense took place in combination with other assaultive offenses or robbery, however.

Scanning the case characteristics for the remaining variables, a general pattern of aggravating factors does not emerge. Several offenders have very high prior offense seriousness scores; but several do not. Offender scores for the use of a weapon appear to be rather high; in fact, considering the scores for weapon use, it is rather surprising that the corresponding scores for injury to the victim(s) are not higher. One additional trend that does appear is that the

offenders' prior record scores appear to be quite high for those offenders who have committed robberies. One might surmise that the continuing criminal activity of robbery offenders, in particular, was a reason for judges to aggravate some sentences.

Insert Table 8.8 here

In reviewing the cases with negative residual values, several opposing trends appear in the data. These cases received sentences that were less severe than the guidelines sentence. If the guidelines had been consulted in those cases, and the same sentence been imposed, the judge would presumably have deviated from the guideline sentence, because of some mitigating factors. In comparison with the positive residual cases, where the presence of sexual assault is conspicuously present, it is conspicuously absent for those cases that received apparently mitigated sentences. However, a number of the offenders did commit assaultive offenses, and one offender committed murder; thus the lack of physical injury cannot have been the only reason for sentence mitigation. In general, it does appear that a number of the offenders who received much lighter-than-predicted sentences either committed offenses without causing injury to the victim, or that did not involve the use of a weapon. Several of these offenders did not have large prior offense seriousness scores, and the bulk of the offenders had never before been incarcerated. This last point is of particular interest. A number of judges, during our interviews with them about how they decided upon particular sentences for offenders, indicated a general reluctance to commit an offender to jail or prison who had not previously been incarcerated. The general belief was that incarceration would not rehabilitate the offender; at worst, the experience of incarceration would probably induce further criminal tendencies.

At this point, we consider briefly some kinds of deviations, in the Massachusetts data, from the sentences prescribed by the Massachusetts guidelines (supposedly based on those data). The most obvious type of deviation from the guidelines would be found in the case that received a life sentence without any possibility of parole -- simply because the guidelines do not make any provision whatsoever for this type of sentence. We noted earlier that six cases in the sample had in fact received sentences of "life without parole." What would the Massachusetts sentencing guidelines -- given their current structure -- have prescribed for those six cases? The answer is, sentences requiring incarceration of 6,

Table 8.8 : Characteristics of Cases With Excessive Residual Values

Case Description								DA Recommended	Sentence
Residual Value	T=Trial P=Plea	Summed Seriousness Score Prior Offenses	Summed Seriousness Score Current Offense	Number Prior Incarcerations	Injury Score	Weapon Use	Specific Offenses	Actual Time (Months)	Actual Time (Months)
+226	T	37	16	25	2	9	Murder Assault Robbery	--	312
+156	P	11	4	4	-	2	Sexual Asslt Assault Burglary	64	184
+156	P	2	18	-	-	-	2 Sexual Asslts Assault	56	180
+152	T	-	4	-	2	3	2 Sexual Asslts Assault	120	180
+148	P	-	4	-	4	1	---	--	180
+139	T	10	12	4	2	2	Sexual Asslt 2 Assaults Robbery	168	180
+135	P	11	21	2	-	8	2 Robberies	--	192
+135	P	41	1	9	-	2	---	56	180
+128	T	-	4	-	5	6	2 Assaults Robbery	--	180
+123	T	15	4	9	2	2	Sexual Asslt Assault	12	160
+121	T	6	13	-	1	5	Sexual Asslt Assault Burglary	--	163
+121	T	4	8	8	8	1	Sexual Asslt 2 Assaults Robbery	--	180
+105	P	5	6	3	-	-	Sexual Assault 2 Assaults	--	120

+101	P	-	8	-	12	2	Sexual Asslt Assault	--	180
+ 97	T	2	7	3	12	3	Robbery	--	180
+ 97	P	20	3	12	4	1	---	152	144
+ 88	P	3	10	-	6	3	Assault Robbery Burglary	--	144
+ 84	P	8	15	7	4	5	Sexual Asslt Assault Robbery 2 Burglaries	--	144
+ 82	T	1	21	1	-	4	Sexual Asslt Assault Robbery	160	120
+ 80	P	30	8	9	-	1	2 Robberies	--	120
+ 79	P	10	4	-	-	-	---	96	96
+ 77	P	60	8	19	-	2	2 Robberies	160	144
+ 70	P	16	5	10	-	1	Robbery	--	96
+ 69	P	5	3	6	4	1	---	80	104
+ 68	P	14	35	10	6	11	Sexual Asslt 2 Assaults Robbery Burglary	--	180
+ 67	T	19	13	13	16	7	Assault Robbery	--	200
+ 65	T	8	5	3	3	8	2 Murders Sexual Asslt Assault	--	120
+ 64	T	2	52	9	-	-	Sexual Asslt 2 Assaults	144	120
+ 63	T	-	4	-	-	-	Sexual Asslt 2 Assaults	120	72
+ 59	P	7	11	17	-	-	Burglary	--	80
+ 58	T_o	12	4	4	-	1	Robbery	160	80
+ 58	T	8	3	3	4	1	---	96	96

+ 57	P	-	3	-	4	1	---	12	88
+ 57	T	3	15	-	-	6	---	--	96
+ 56	P	6	4	1	4	2	---	--	96
+ 53	P	14	5	4	-	2	Assault Robbery	--	80
+ 52	P	25	3	17	-	-	---	--	80
- 51	P	1	2	-	10	2	Assault	12	12
- 51	T	3	10	-	6	2	Assault Robbery	12	1
- 52	P	39	6	15	4	2	Assault Robbery	--	15
- 52	P	13	53	3	-	-	Burglary	--	14
- 54	T	76	1	24	-	-	Assault	18	12
- 56	P	-	17	-	6	4	Assault Robbery	--	6
- 65	T	-	-	-	11	3	Murder Assault	--	4
- 83	P	12	118	12	3	27	Assault Robbery Burglary	--	160
-103	P	5	8	-	15	5	2 Assaults	12	3
-132	P	2	6	-	20	9	2 Assaults	12	6

7 1/4, 8, 9, 10 1/4, and 12 1/2 years respectively -- that is, on average, about nine years, Recall that the guidelines sentence represents time to be served before parole eligibility; and that a life sentence with parole has an expected time to be served of 15 years. What is interesting about the guidelines' sentences in these instances is that not one of them was more than 15 years in length. In other words, the prescribed term of imprisonment, in each of these "life without parole" cases, was a good deal less than that entailed by a life sentence of the kind that would have allowed for parole (after 15 years).

The second type of "misses" that we examined in some detail were those cases where the Massachusetts sentencing guidelines call for an "out" sentence -- but in which the sentences actually imposed involved some amount of time spent in jail or prison. As might have been expected, given the factors included in the Massachusetts sentencing guidelines and their associated weights, there were not a lot of cases like this; in fact, only 43 offenders would have received an "out" sentence under the guidelines, who had in fact been incarcerated. Almost half of these 43 incarcerated offenders had received sentences of incarceration for between one and five months in length; periods incarcerated for the remaining offenders ranged from 6 to 15 months, and one offender in this group received a sentence entailing incarceration approximately 40 months.

Finally, the lengths of the prescribed "in" guidelines sentences for those offenders who were not incarcerated were examined. Also as might have been expected, given the factors (and weights) included in the Massachusetts guidelines, there were quite a number of such offenders; 248 offenders who actually received "out" sentences would have been prescribed "in" sentences under the Massachusetts guidelines. What were the lengths of time to be spent in jail or prison prescribed for these offenders? Table 8.9 presents the distribution of lengths of terms prescribed for these cases.

Insert Table 8.9 here

The frequency distribution of guidelines sentence lengths for the "prescribed in, but actually out" cases in this table shows a markedly different pattern from that of the "prescribed out, but actually in" cases just previously described. The Massachusetts sentencing guidelines in fact prescribed a wide range of sentence lengths for cases that judges had decided not to incarcerate at all. Whereas almost half of the offenders who should according to the guidelines

Table 8.9: Cases Not Actually Incarcerated, for whom the Guidelines Prescribe Incarceration (N=248)

Guideline Sentence	N	%
1-5 months	24	9.7
6-10 months	67	27.0
11-15 months	30	12.1
16-20 months	35	14.1
21-25 months	20	8.1
26-30 months	15	6.0
31-35 months	18	7.3
36-40 months	12	4.8
41-45 months	13	5.2
over 45 months	14	5.7
Total	248	100.0

have received an "out" sentence actually received jail or prison time of up to 5 months, the guidelines prescribed sentences of that length for only 10 percent of the cases not in fact incarcerated. About 60 percent of the cases actually not incarcerated would have sentences prescribed under the guidelines of up to about 20 months in length. More alarming, perhaps, is that the upper tail end of the distribution of guidelines sentences for what were actually "out" cases appears to taper off, only minimally, at about 36 months and over. In other words, had these last offenders, who received non-incarcerative sentences as actual dispositions, been sentenced under the Massachusetts sentencing guidelines, their initial guidelines sentence would have been one leading to terms of three years or more. This fact appears to confirm again our earlier view that the Massachusetts guidelines research results were badly distorted by cases evidencing extreme seriousness of one sort or another.

Examining the Guidelines Ranges

Our discussion thus far about the fit of the Massachusetts data to the sentencing guidelines has been only concerned with the relations between the prescribed guidelines sentence and the sentence actually imposed, in each case. The analysis has not, so far, considered another important feature of the Massachusetts sentencing guidelines -- that of the sentence range (of plus or minus 50 percent) allowed around the calculated sentence.

According to the Massachusetts sentencing guidelines, any sentence that is within a fifty percent range, either above or below the initial guidelines sentence, is considered to be "acceptable", i.e., the sentence is not considered a departure from the guidelines. We noted in Chapter 7 of this report, that the Massachusetts sentencing guidelines project said that their choice of fifty percent as the appropriate range was based on their calculation that a smaller range width would result in sentencing guidelines that would not have included most of their sample case sentences. The research team also believed that a range wider than fifty percent would not have been of much help to judges in structuring sentencing discretion, because even extreme cases could be accounted for using a wider range size.

Since the reports of the Massachusetts sentencing guidelines project's research on this particular issue were not available to us, and since there appeared to be some confusion over exactly how many of the original sentences could be covered by guidelines that incorporated a fifty percent range, we decided to investigate this issue of sentence ranges a little more closely. We were specifically

interested in two aspects of the sentence ranges. First, what proportions of the offenders' actual sentences would be within ranges of varying widths around the initial guidelines sentence? Secondly, since sentences within any given range around an initial sentence could fall either above or below that initial sentence, we wanted to detemine where the bulk of the Massachusetts sample were located. In other words, were most of the sentences (in the construction data) within any given range higher or lower than the initial guidelines sentence?

Table 8.10 summarizes the results of our investigation of the first of these two issues -- i.e., of how many offenders' actual sentences would be found within ranges of various widths around the guidelines sentence. The analyses were conducted initially on all of the cases in the original sentencing data; later on, we split the sample into subgroups of trial and plea disposition cases, to determine if the guidelines would capture different proportions of these two types of offenders within the specified ranges. The left-hand column of the table lists the various widths analyzed; as is clear from the list, the range widths begin at zero (or no range at all around the guideline sentence) and continue in increments of 10 percent to include a 100 percent range around the guideline sentence at the upper extreme. Of course, the size of the widths chosen was arbitrary; ranges that increased by five (or any other) percent could also have been used.

Insert Table 8.10 here

The findings reported in Table 8.10 are somewhat alarming, since they show that very wide ranges around the initial Massachusetts guidelines sentence would have been required in order for the guidelines to have actually included a majority of the sentences in the original data. Let us look at the percentages of cases captured within each of these ranges around the guidelines sentence, for all of the cases combined, to make this point more apparent. If the Massachusetts sentencing guidelines project had chosen the sentence calculated by using the guidelines to be the only (presumptive) sentence to be imposed by judges -- that is, if they had provided for a "zero range" around the initial guidelines sentence, unless reasons could be given for a departure -- then only about 7 percent of the offenders in the Massachusetts sentencing data would actually have received an "acceptable" sentence. By specifying a 10 percent range around the initial sentence as the acceptable sentence spread, then about 10 percent of the offenders in the data base would actually have received such sentences -- only a three percent

Table 8.10: Distribution of Massachusetts Actual Sentences Within Ranges of the Massachusetts Guidelines Sentence, All Cases (N=1440)

% Range Around Guideline Sentence	All Cases (N=1440)		Trial Cases (N=221)		Plea Cases (N=1198)	
	N of Cases Included	Cumulative % of Cases Included	N of Cases Included	Cumulative % of Cases Included	N of Cases Included	Cumulative % of Cases Included
No Range	107	7.4	8	3.6	89	7.4
10% Range	149	10.4	16	7.2	123	10.3
20% Range	191	13.3	23	10.4	158	13.2
30% Range	255	17.0	38	17.2	206	17.2
40% Range	325	22.6	51	23.1	262	21.9
50% Range	478	33.2	89	40.3	376	31.4
60% Range	549	38.1	105	47.5	432	36.1
70% Range	668	46.4	119	53.8	537	44.8
80% Range	813	56.5	136	61.5	664	55.4
90% Range	965	67.0	153	69.2	798	66.6
100% Range	1314	91.2	189	85.5	1106	92.3

*If the guidelines sentence was 5 months or less, the permissible range has a lower limit of 0 months.

improvement over allowing no range of sentence at all. At each increment in the range spread, it appears, from the figures provided on Table 8.10, that about 3 percent more of the cases would have had acceptable (i.e., non-departure) sentences under the Massachusetts guidelines.

It will be recalled that the range finally chosen for the Massachusetts sentencing guidelines is one of plus or minus 50 percent. The figures in Table 8.10 show that this plus-or-minus range would include only a third of the sentences in the original data. Given this fact, it is hard to believe the various statements made by the Massachusetts project staff, to the effect that the choice of the final (50 percent) range width was based on calculations that the guidelines would then cover a majority of the cases.

A somewhat different pattern emerges when trial and plea disposition cases are analyzed separately, although in neither case does this much improve the fit of the guidelines to the construction data. As can be seen in Table 8.10, only 4 percent of the actual sentences for trial cases would be included within a zero range width, and about 40 percent of the trial cases' actual sentences would be covered by the guidelines given a 50 percent range width around the guidelines sentence. Comparable figures for plea disposition cases indicate that 7 percent of the actual sentences would fall within a zero range, while about 30 percent of the actual sentences would fall within a 50 percent range width. Aside from the fact that plea cases are a little bit more often right "on the nose" (in terms of having guidelines sentences that equal the actual sentence imposed) than are trial cases, there does not seem to be any particular pattern that differentiates one type of dispositions' sentences from the other.

After determining that the Massachusetts guidelines do not cover a majority of the cases within ranges of up to plus or minus 50 percent (at least), we now turn to the question of where the cases within each range were actually located. It should be obvious that if a setnence falls within any given range, but is not the same exact sentence as the initial guidelines sentence (i.e., right on the nose) then the actual sentence must fall either above or below the initial guideline sentence within the range. Table 8.11 splits the proportions of cases within each of the ranges in Table 8.10, into those that were above the initial guidelines sentence and those below it. The table has been percentaged across its rows to permit comparisons between the various ranges. However, since each range necessarily includes the cases in earlier ranges, as well as the cases added specifically for that range, the figures listed for the 100 percent range deviations are also

the marginal percentages for the total number of cases deviating, in one or the other direction, from the initial guidelines sentence, for all categories of ranges.

Insert Table 8.11 here

The pattern shown in Table 8.11 is that within-range deviations from the initial Massachusetts guidelines sentences are overwhelmingly those that are below the initial guidelines sentences, regardless of the range width specified. Moreover, as the ranges increase in width, the preponderance of deviations below the guidelines sentence becomes even more pronounced. Compare, for example, the deviations for a range of 10 percent and a range of 90 percent: 62 percent of the deviations are below the initial guidelines sentence in the 10 percent range, while four of every five deviations in the 90 percent range group are below the initial guidelines sentence. Thus, even when we allow The Massachusetts guidelines to encompass greater proportions of the original cases by specifying wider ranges, the guidelines sentence becomes increasingly more severe than the sentence that was actually imposed in those cases. Enough has been said, we think, to show that the Massachusetts sentencing guidelines mainly call for more severe sentences than those actually imposed in the past on comparable cases.

Constructing New Massachusetts Sentencing Guidelines

Our final analyses of the Massachusetts sentencing data were concerned with the development of a different version of sentencing guidelines for that state. We have noted in earlier sections of this chapter, as well as in Chapters 6 and 7, that the conception, development and implementation of the Massachusetts guidelines were all heavily influenced by policy decisions made either by the guidelines project's research team, or by their oversight body, the Committee on Probation and Parole of the Superior Court. We also noted, in this chapter, that our reanalysis of the Massachusetts sentencing guidelines suffered to some extent from the fact that we did not know precisely what the research team had done, due to the absence of precise documentation for many of their research analyses. We knew that policy decisions were likely to have modified the descriptive analyses done; as it is, however, we are unable to say how much effect they had. Since we did, however, have access to the data actually collected by that project and supposedly used in part to develop the guidelines, we thought it appropriate to conduct some analyses that would, independent of the Massachusetts project's efforts, yield a description of pre-guideline sentencing patterns in the state

Table 8.11: Percentages of Sentences Above and Below the Massachusetts Guidelines Sentence by Range for All Cases (N=1440)

Range	Sentence Below Guideline		Sentence Above Guidelines		Total	
	N	%	N	%	N	%
10% Range	26	61.9	16	38.1	42	100
20% Range	48	57.1	36	42.9	84	100
30% Range	91	61.5	57	38.5	148	100
40% Range	147	67.4	71	32.6	218	100
50% Range	248	66.8	123	33.2	371	100
60% Range	310	70.1	132	29.9	442	100
70% Range	421	75.0	140	25.0	561	100
80% Range	558	79.0	148	21.0	706	100
90% Range	701	81.7	157	18.3	858	100
100% Range	1037	85.9	170	14.1	1207	100

of Massachusetts. Our intention thus was to develop a model of sentencing practice based on those factors that were most predictive (in the statistical sense) of sentencing outcomes. Would this model look anything like the guidelines eventually developed?

Examining Correlations Among Items

We began our analysis of sentencing practice in Massachusetts by examining the correlations between each of the data elements in the original sentencing data and the "in-out" decision on the one hand, and the length of term decision (for those cases that had been incarcerated) on the other. In addition, we studied the degree to which items in the data were highly correlated with each other; as we noted in earlier sections of this report, it is wise to choose those factors that are highly correlated with the outcome variable, but fairly independent of each other, to use as predictors of that outcome variable. To do otherwise may introduce inefficiency due to multicollinearity.

One additional reason that we began the guidelines development process with an analysis of the correlations between factors and the sentencing decisions was to serve as a check on the manner in which we reconstructed the Massachusetts data set. We noted at the onset of this chapter that several of our statistical programs required aggregate (case-level) data for analysis; thus, it was necessary for us to compute totals for items such as seriousness of current convictions, etc., rather than analyze each current conviction separately. We were unsure, at times, of the extent to which our procedures may have changed the scoring of various items in the original data (such as weapon use, or injury); since the Massachusetts project did not document the manner in which they constructed such aggregates, we could not know if we had followed the same procedures that they did. However, we did have some information from the Massachusetts project on the degree to which certain items in the data correlated with the "in-out" and length decisions; thus, we were able to compare our correlations with theirs as a check on the quality of our aggregate data as compared to their figures. We are pleased to note that the correlations suggested to us that we had aggregated the data in much the same fashion as the Massachusetts project did; all of the correlations were of the direction and strength that, according to their figures, should have been expected.

The Decision to Incarcerate

About 25 items showed at least moderate correlations with the decision to incarcerate; the magnitude of the coefficients

for these items ranged in value (either positively or negatively) from .20 to .46. The five most highly correlated items all pertained to some aspect of the defendant's custodial status before (trial or plea) disposition of the case. These items were: the logarithm of the amount of time detained between arrest and disposition (+.46); whether the offender was at liberty before trial (+.42); whether the offender had been released on recognizance (+.36); the offender's specific liberty status before trial -- i.e., on bail, or own recognizance (+.34); and the amount of the final bail set (+.30). As might be expected, these items were highly correlated with each other; thus in our later regression analyses we used only 3 of them.

Of the next group of factors correlated with the incarceration decision -- were any costs assessed at disposition? (-31); did the offender file for appellate review? (.30); the actual amount of time detained prior to trial (rather than the log of this item) (+.29) and the log of the amount of the total costs assessed (-.28) -- none was used in our later regression analyses. The first two of these items, as well as the last, are not really relevant to the decision to incarcerate, as they occur in time after that decision has been made. The cost disposition items are probably, as well, contingent on a decision by the judge not to incarcerate an offender; hence their negative coefficients. The last item, time detained before trial, was already present, in effect, in the first group of strongly correlated items, in logarithmic form.

Seriousness scores of prior and current convictions were the things next most highly correlated with the decision to incarcerate (+.28 and +.27); followed by the number of prior incarcerations (+.26); offender cooperation with the authorities (+.24); number of current persons offenses (+.24); robbery as conviction offense (+.23); weapon use (+.22); number of prior persons offenses (+.22); neighborhood type (-.22); method of conviction (+.22); number of counts charged (+.21); district attorney's recommendation for incarceration (+.20), and the number of counts at conviction (+.20). Two of these items were not included in our analyses -- offender cooperation and neighborhood type -- because there were too many missing observations for each. The remainder of the items were included at least initially in our regression analyses, as they showed only fairly low to moderate correlations with other factors in the data.

Finally, a number of items, excluded during our early regression analyses, were entered into those equations at later stages, as the initial items proved to be unimportant. These items, and their correlation with the incarceration

decision were: gun use score (+.19); any change in bail status before trial (+.19); age at first juvenile adjudication (+.19); district attorney's recommendation for a suspended sentence (-.18); number of juvenile supervision terms (+.18); seriousness of charged offenses (+.18); any charges pending (+.17); number of weeks employed during last year (+.17); number of prior paroles from incarceration (+.17); and number of juvenile adjudications for serious offenses (+.16). Some of these items were later discarded, however, for lack of predictive power in the incarceration decision analyses.

What can be said about these factors? Most of the items that proved to be highly correlated (in terms of the magnitudes given above) with the decision to incarcerate would have been expected. By and large, the items represented some aspect of the seriousness of the offender's prior adult and juvenile criminal history, or some aspect of the seriousness of the current charged and conviction offenses. However, the items most highly correlated with the prediction of incarceration were not of that kind. That whether or not an offender had been released (e.g. on bail) before trial appears to have the greatest association with ultimate incarceration is a finding that would probably cause some judges consternation, in Massachusetts or elsewhere. This correlation seems to imply that the ultimate decision to incarcerate is strongly influenced by prior decisions made about pre-trial release; and thus that decision should be examined more closely.

Length of Sentence

Fewer items were found to be highly correlated with the judges' decision as to the length of jail and prison sentences; though most of the items that did show moderate correlations were ultimately included in our regression analyses of that decision. In examining the correlations between various offender items and the length of sentences, we used only those offenders who had been incarcerated, and, further, we excluded those 47 cases that earlier analyses had led us to believe were "unusual", and the six cases where the offender had received a life sentence without parole.

The things most highly correlated with lengths of jail and prison sentences were: the amount of injury to the victim (+.49); the district attorney's recommendation as to the length of incarceration (+.44); whether the offender filed for appellate review of the sentence (+.41); the logarithm of the final bail amount (+.39); the number of current persons offenses (+.38); the gun use score (+.36); weapon use score (+.36); and the seriousness of the conviction offenses (+.30). What is interesting to note about these factors, in comparison

with those factors found to be correlated with the decision to incarcerate, is that they are mostly not things concerned with the offender's custodial status prior to trial.

The second group of factors, that proved to show the strongest correlations with sentence lengths, included the amount of time detained between arrest and trial (+.26); the log of that item (+.25); whether the conviction offense was for robbery (+.25); was the offender at liberty at the time of disposition (+.24); whether the offender was prosecuted as a major violator (+.23); the number of counts at conviction (+.18); whether the offender was released on his or her own recognizance (+.18); and the number of counts charged (+.17). Finally, a number of variables that were only slightly correlated with the length of sentence decisions were later included in the regression analyses of that item, as earlier included factors proved to have only slight predictive power and were removed from further analyses. These items were: the seriousness of prior offenses (+.17); the number of prior persons offenses (+.15); the seriousness of charged offenses (+.15); the value of property taken (+.15); whether the victim was known to the offender (+.14); the number of prior paroles (+.14); the number of counts at conviction (+.14); whether there were any pending charges (+.14); and the number of prior incarcerations (+.13). Of the factors listed above, the number of counts at conviction and whether the offender was released on recognizance were not included as predictors of sentence length, because they were highly correlated with other factors already being used -- the number of counts charged and the offender's custody status prior to disposition.

To summarize, the items of information that proved to be most highly correlated with the length of offenders' sentences were most often items that conveyed information about the extent of the offender's prior involvement in the criminal justice system. In contrast, those items most highly correlated with the decision to incarcerate at all appeared to involve the offender's custody status prior to trial.

Finding the Best Fit

The first aspect of our analysis involved inspection of associations between all usable items in the data and the two types of sentence decisions, i.e., the decision to incarcerate and the length of term decision. As we noted in the last section, a number of different items of information about the offender and the offense proved to be moderately associated with those two decisions. In the second stage of our analyses, we incorporated, in a series of regressions, those factors that looked predictive of each decision. Initially,

the highest correlates of the decision (either "in-out" or length of term) were included in the regression equations. Then, factors that proved to account for insignificant amounts of the variation in the decisions in question were removed and other items (usually of lower zero-order association) were substituted for them.

The final results of the regression analyses of the decision to incarcerate, and of the length of sentence for incarcerated cases, are reported here. Tables 8.12 and 8.13 list those factors that proved, from a purely statistical point of view, to be the best predictors of the decision to incarcerate and of the decision as to how long an offender's jail or prison sentence should be.

Table 8.12 lists those factors that are the best predictors of the decision to incarcerate -- shown separately for trial and plea cases, and for cases sentenced by a Superior Court judge, as well as for all cases combined. As can be seen by the check marks appearing in the columns of this chart, a number of factors prove to be the best predictors when all cases are considered that are not predictors of the separate trial and plea case incarcerations. When all of the cases in the sample are used in the analysis, nine factors are able to account for about 33 percent of the variation in the decision to incarcerate. Those factors are: the amount of time detained between arrest and disposition of the case, whether the offender was at liberty (or released) at the time of case disposition, the seriousness of prior and current conviction offenses, whether the case was disposed of by trial or plea, the district attorney's recommendation as to whether to incarcerate, the number of counts at conviction, whether there had been a change in bail status, and the number of juvenile supervision terms.

Insert Table 8.12 here

A number of these factors continued to account for a significant portion of the incarceration decision variation in cases disposed of by a plea of guilty; however, only one of these factors proved to hold any predictive ability for the decision to incarcerate in trial cases. The amount of time detained from arrest to disposition of the case continued to account for a significant portion of the trial cases incarceration decision, as well as for the plea case disposition. The remaining initial factors did not appear to contribute to the explanation of trial case incarcerations. In their stead, three other factors emerged as significant predictors of trial case incarceration decisions. These items

Table 8.12: Rutgers Secondary Guidelines Analysis of the Decision to Incarcerate (N=1440)*

Sentencing Factors	All Cases (N=1050)	Trial Cases (N=180)	Superior Court Trial Cases (N=135)	Plea Cases (N=928)	Superior Court Plea Cases (N=785)
Time Detained: Arrest to Disposition (Logarithm)	✓	✓	✓	✓	✓
At Liberty at Disposition	✓				
Seriousness of Prior Offenses	✓			✓	
Seriousness of Current Conviction Offenses	✓			✓	
Trial or Plea Disposition	✓				
District Attorney Recommended Incarceration	✓			✓	✓
Number of Counts at Conviction (Logarithm)	✓			✓	✓
Change in Bail Status	✓				
Number of Juvenile Supervision Terms	✓				
Offender was Released on Recognizance		✓	✓	✓	✓
District Attorney Recommended Costs as Sanction		✓	✓		
Other Charges Pending		✓	✓		
Robbery as Conviction Offense				✓	✓
R Square	.33	.28	.30	.33	.32
F Value	58.09	17.18	14.01	63.63	73.31
Significance	.0001	.0001	.0001	.0001	.0001

*Different N's appear per category due to listwise deletion of cases evidencing missing information on any included item.

-- whether the offender had been released on recognizance, whether the district attorney recommended costs as a sanction, and whether there were other charges pending -- indicated that judges appeared to be influenced by prior decisions about the offender's potential for default if released prior to trial, and about the offender's appearance for the disposition of other charges that were pending.

Different factors proved to be the most predictive of the incarceration decision for plea cases, as we noted earlier. Again, the amount of time detained between arrest and disposition was found to predict all plea disposition cases. The seriousness of prior and current conviction offenses, on the other hand, was predictive of all plea disposition cases generally, but did not account for any significant amount of the incarceration decision variation for plea dispositions by Superior Court judges. In addition, four other items were found to account for significant amounts of the incarceration decision variation in plea cases: the district attorney's recommendation as to whether to incarcerate, the number of counts at conviction, whether the offender was released on recognizance, and whether the offense of robbery figured as a conviction offense.

Though the nature of each of the items that were found to predict the incarceration decision for various types of cases is of interest, we were most interested at this stage in the amount of variation that these items could explain, when compared to that explained by the Massachusetts sentencing guidelines factors. We noted earlier that the Massachusetts sentencing guidelines factors -- seriousness of prior and current conviction offense, weapon use score, and injury score -- can account for only about 17 percent of the variation in the judges' decisions in the construction data. That figure rose to about 19 percent for plea cases, and dropped to approximately 15 percent for trial cases. In contrast, the items included in the final guidelines model developed by this project are able to explain approximately a third of the incarceration decision variation over all cases, and a little less than a third of the variation in that decision for <u>either</u> plea or trial cases -- a substantial improvement over the predictive power of the Massachusetts guidelines factors.

The improvement in predictive ability over that of the Massachusetts sentencing guidelines was even more pronounced when the length of sentence decisions were reanalyzed. Table 8.13 provides a list of those variables that, after a series of regression analyses, continued to explain significant portions of the length of sentence decisions. The R square values listed for each of the separate regression analyses presented in this table indicate that it is possible to

explain about 44 percent of the variance in sentence lengths for all cases combined, and about 70 percent of the offenders' sentence lengths for trial cases. (The Massachusetts guidelines factors could account for about 35 percent of the variation over all cases, and about 60 percent of the variation in trial disposition cases, in contrast.) However, only about 33 percent of the plea disposition sentences can be explained using the factors listed; this is about the same portion of variation that also can be accounted for using the Massachusetts sentencing guidelines -- thus, the improvement for this class of cases seems, on the surface, to be slight.

Insert Table 8.13 here

Nine factors again prove to be the most important predictors of the sentence length decision for all cases. These are: the amount of injury to the victim; the district attorney's recommendation as to sentence length, the final amount of bail, the weapon use score, the seriousness of the current and prior offenses, the amount of time detained prior to trial or plea, whether the conviction offense was for robbery, and whether other charges were pending disposition. Only four factors entered into the regression analyses for trial cases; these included injury to the victim, the final bail amount, the amount of time detained prior to trial and the number of current persons offense convictions. Most of the original nine factors -- excluding the district attorney's sentence length recommendation and the other charges pending -- were important predictors of the plea case sentence lengths. In addition, it appeared that whether the case had been prosecuted by the Major Violator's Division of the prosecutor's office also contributed to the explanatory power of the model.

The factors that proved to be important in the prediction of the length of term decision (for all incarcerated cases) do include the Massachusetts sentencing guidelines factors. Injury to the victim, weapon use, and the seriousness of current and prior offenses are able to account for significant amounts of variation in the offenders' lengths of term. What is interesting to note about these analyses, however, is that the four Massachusetts guidelines factors, alone, do not best account for sentence length variation; rather, five additional factors are needed to accomplish that end.

Of more interest is that the Massachusetts sentencing guidelines factors are not significant predictors of offenders' sentence lengths for trial, or Superior Court trial, cases. Only one of those items -- injury to the victim

Table 8.13: Rutgers Secondary Guidelines Analysis of the Length of Sentence for Incarcerated Cases, Excluding 47 Outliers, and Life Sentence Cases (N=721)*

Sentencing Factors	All Cases (N=568)	Trial Cases (N=127)	Superior Court Trial Cases (N=96)	Plea Cases (N=424)	Superior Court Plea Cases (N=358)
Injury to Victim Score	✓	✓	✓	✓	✓
District Attorney Recommendation as to Length of Term	✓				
Final Bail Amount (Logarithm)	✓	✓	✓	✓	✓
Weapon Use Score	✓			✓	✓
Seriousness of Current Conviction Offenses	✓		✓	✓	✓
Time Detained: Arrest to Disposition	✓	✓	✓	✓	✓
Robbery as Conviction Offense	✓			✓	
Seriousness of Prior Offenses	✓			✓	✓
Other Charges Pending	✓				
Number of Current Persons Offense Convictions		✓			
Prosecution by the Major Violator's Division				✓	✓
R Square	.44	.67	.70	.32	.33
F Value	49.60	62.81	53.78	24.33	24.69
Significance	.0001	.0001	.0001	.0001	.0001

*Different N's appear per category due to listwise deletion of cases evidencing missing information on any included item.

-- is able to contribute significantly to the prediction of sentence length for trial cases. In contrast, all of the Massachusetts guidelines factors do contribute to the sentence length prediction for plea cases. This suggests that the sentencing guidelines model developed by the Massachusetts guidelines project for initial use on trial cases was in fact one fitted to the judges' decisions made in plea cases (which greatly outnumber trial cases in the sample).

We are not, of course, holding up our "model-building" analyses as an example of what ought to be done by anybody who wants to make up "empirically based" sentencing guidelines. For one thing, we did not validate these results through a separate analysis (any more than anyone else who has carried out this kind of research has done); our results are intended to be illustrative only. What these analysis do show, we think, is that the factors included in the final Massachusetts guidelines bear little resemblance to the set which would be used, if the "best" purely descriptive analysis is all that is wanted. Further, it is clear that (for the Massachusetts data) different combinations of different variables yield the "best" predictions of the "in-out" decision, and of lengths of terms, for different sub-sets of cases in the construction data; this in turn suggests that different decision-making processes were going on in, e.g., trial and plea cases. The "process" variables we used (including such things as custody status at disposition and amount of bail) are probably best interpreted as proxies for "dangerousness" or some similar variable, which in turn is correlated with judicial views of seriousness of offense and prior record; we do not suppose that the Massachusetts judges would use e.g. amount of bail as such as a determinant of sentences. Nonetheless, any purely descriptive analysis of sentencing practice should surely have stumbled across findings like those which we have just presented. Moreover, to the extent that "process" variables do play some part -- never mind what part -- in sentencing practice, they need to be considered at the analysis stage, even if they would never be included in guidelines finally developed for use in the future. The reason for this is that -- like other clearly inappropriate variables, such as race -- things like amount of bail, or length of time from arrest to trial, may bias the estimated effects of appropriate variables such as seriousness of offense and prior record. This possibility cannot even be considered, let alone corrected for, unless those variables are explicitly included in sentencing models at some stage of the analysis.

Examining the Guidelines Ranges

In the last section of this chapter, it was emphasized that it is possible to improve the fit of the Massachusetts sentencing data to the various sentencing decisions if one is interested solely in those items that are best able to predict sentence outcomes. Of course, the Massachusetts sentencing guidelines project was not only interested in modelling the best predictors of the sentence decision, but also had to reconcile ethical considerations about the influence that certain variables should have on sentence outcomes as well. The end result of the Massachusetts sentencing guidelines project's efforts to reconcile ethical considerations with the inclusion of various predictive factors was a set of sentencing guidelines that did not represent the best possible description of sentencing patterns in that state. Rather, as we have explained at length in other sections, those guidelines are heavily influenced by policy decisions made by the project and the judiciary, that made necessary the exclusion of items thought to be undesireable considerations in sentencing decisions.

But suppose, as we have done, that the Massachusetts guidelines project and the judiciary did not consider whether or not certain items of information should influence judges' decisions. Suppose, that the project only considered those items that were the best statistical predictors of sentence outcomes. The Massachusetts sentencing guidelines would, in that case, include the factors listed in Tables 8.12 and 8.13 in their models of the decision to incarcerate and of the length of term decision. And, in some instances, these items are not the same as those used in the Massachusetts sentencing guidelines model.

The higher regression coefficients found for our revised versions of sentencing guidelines for the decision to incarcerate and to determine the length of sentence indicate that these models provide a better fit to the Massachusetts sentencing data than does the Massachusetts guidelines model. But how much better is that fit in terms of the numbers of cases that could be covered by guidelines of different range widths? Also, for those cases that fall within ranges of given sizes, but do not show exact guidelines sentences, how far away from the guidelines sentence are they, and in which direction? To answer these questions, we replicated the analysis of guidelines ranges done earlier. This time, however, we conducted the analysis on incarcerated cases using the guidelines factors produced for all cases from our re-analysis (regardless of method of disposition) as the predictors of sentence length.

Table 8.14 provides the percentages of actual sentences that would fall within ranges of various sizes for incarcerated cases in the Massachusetts sample. Again, the left hand column of the table lists the various widths analyzed; as is clear from this list, the range widths begin at zero (or no range at all) and continue in increments of 10 percent to include a 100 percent range around the guideline sentence at the upper extreme.

Insert Table 8.14 here

The figures presented in Table 8.14 substantiate the claim that our revised version of the Massachusetts sentencing guidelines provides a slightly better fit to the actual sentences received by offenders than does the actual Massachusetts guidelines model. As can be seen from that table, the new guidelines are able to capture about 40 percent of the sample cases within a fifty percent range; the Massachusetts guidelines could only capture about 30 percent of the cases within such a range. Moreover, the number of cases within ranges appears to increase uniformly as the range size increases. This suggests that there is not a concentration of cases that missed the guidelines sentence by a particular amount at any range.

More importantly, perhaps, is the fact that our revised guidelines do not appear to affect trial cases in a significantly different manner from how they affect plea cases. The proportions of cases included within ranges of any given size are similar, regardless of whether the offender was sentenced after a trial or after a plea of guilty. These percentages are also similar to those provided for each range when both types of cases are combined, as would be expected.

The second phase of our analysis of how well the revised guidelines fit the actual sentences imposed addressed the question of whether cases, included within ranges of various sizes, displayed actual sentences that were uniformly higher or lower than the sentences provided by the guidelines. Table 8.15 presents the results of this analysis of sentence length deviations within ranges for all incarcerated cases.

Insert Table 8.15 here

The most striking fact about the results presented in Table 8.15 is that the type of "miss" from the guidelines sentence (i.e., above or below that sentence) varies depending on the range width involved. For the lower ranges, the majority of within-range deviations from the guidelines

Table 8.14: Distributions of Massachusetts Actual Sentences Within Ranges of the Rutgers Revised Version of the Massachusetts Guidelines Sentence for Incarcerated Cases, Excluding 47 Outliers and Life Sentences (N=721)

% Range Around Guidelines Sentence	All Cases (N=721)		Trial Cases (N=151)		Plea Cases (N=565)	
	N of Cases Included	% of Cases Included	N of Cases Included	% of Cases Included	N of Cases Included	% of Cases Included
No Range	29	4.0	7	4.6	22	3.9
10% Range	67	9.3	14	9.3	53	9.4
20% Range	106	14.7	25	16.6	81	14.3
30% Range	162	22.5	40	26.5	122	21.6
40% Range	225	31.2	50	33.1	175	31.0
50% Range*	295	40.9	63	41.7	231	40.9
60% Range	321	44.5	68	45.0	252	44.6
70% Range	386	53.5	80	53.0	305	54.0
80% Range	436	60.5	89	58.9	346	61.2
90% Range	477	66.2	98	64.9	377	66.7
100% Range	503	69.8	105	69.5	395	69.9

*If the guidelines sentence was 5 months or less, the permissible range has a lower limit of 0 months.

Table 8.15: Percentages of Sentences Above and Below the Rutgers Revised Version of the Massachusetts Guidelines Sentence by Range for Incarcerated Cases, Excluding 47 Outliers and Life Sentences (N=721)

Range	Sentence Below Guideline		Sentence Above Guidelines		Total	
	N	%	N	%	N	%
10% Range	14	36.8	24	63.2	38	100.0
20% Range	40	51.9	37	48.1	77	100.0
30% Range	69	51.9	64	48.1	133	100.0
40% Range	114	58.2	82	41.8	196	100.0
50% Range	164	61.6	102	38.4	266	100.0
60% Range	187	64.0	105	36.0	292	100.0
70% Range	234	65.6	123	34.4	357	100.0
80% Range	273	67.1	134	32.9	407	100.0
90% Range	304	67.9	144	32.1	448	100.0
100% Range	312	65.8	162	34.2	474	100.0

sentence are such that the actual sentence imposed was higher than that suggested by our revised guidelines model. In contrast, for ranges of 50 percent and higher, the majority of within-range deviations from the guidelines involve actual sentences that are below the guidelines sentence. The majority of the deviations from the original Massachusetts guidelines (presented earlier in Table 8.11), were always of the sort where the actual sentence was below the guidelines sentence.

What do the figures presented in Table 8.15 suggest about the fit of the guidelines to actual sentences? In the first instance, they suggest that the revised guidelines model proposes sentences that, given ranges of 50 percent or less, will be lower than the sentences actually imposed by judges in the past. With ranges of higher value, the trend reverses itself: then guidelines sentences are, as they were for the actual Massachusetts guidelines model, usually higher than thsoe sentences actually imposed. However, at least in the case of guidelines that call for smaller ranges, the presumption in the revised guidelines tends to be toward lower sentences, and this presumption strikes us as more reasonable than guidelines that prescribe higher sentences than actual in all instances.

Summary and Conclusions

Our purpose in this secondary analysis of the Massachusetts sentencing guidelines has been to point out some common problems inherent in guidelines development, and to pose some issues in the construction of guidelines that to date have been overlooked.

In the first section of the chapter, we summarized the data on Massachusetts offenders and their sentences collected by the Massachusetts sentencing guidelines project. As was evident in that discussion, as well as in the discussion in Chapter 7 of this report, the Massachusetts sentencing data suffered from some severe problems. The most common problem was that a great deal of the data was missing, and thus could not be used for analysis. Of the data present that could be used in the development of guidelines, we, as after-the-fact studiers of the research process, could not be sure exactly how that information was used in researching sentencing patterns in Massachusetts because specific information on the statistical research conducted by the project was never released.

Suffering from a lack of information about the guidelines research process as it was conducted in Massachusetts, we decided to investigate the actual fit of the Massachusetts

sentencing guidelines to sentences received by offenders in that state. As was noted earlier, the Massachusetts guidelines generally do not fit the sentences imposed to a very high degree. In fact, when the guidelines are used to predict the sentences that offenders should get, given their scores on the items used in the guidelines, by and large the guidelines suggest sentences that are higher than those actually received. This fact, along with the fact that the structure of the Massachusetts sentencing guidelines appears to create a strong presumption toward incarceration, suggests that sentences prescribed by the guidelines for future cases will tend to be more severe than those given out in the past.

During our analysis of the fit of the Massachusetts sentencing guidelines to the actual sentences imposed, it was further noted that the Massachusetts project did not investigate two particular issues of importance to a guidelines model. The first of these, that of the existence of "unusual" cases in the data, we discovered probably added to the distortion of the guidelines model, in that the guidelines do not attempt to exclude bias caused by offender scores in these cases from the analyses on which their guidelines were based. Secondly, the Massachusetts guidelines project did not investigate, to our knowledge, the possibility of interactive effects between the items included in the guidelines modal. The addition of interaction terms into the regression analyses of the decision to incarcerate and the decision as to length of sentence clearly improve the predictive ability of the guidelines model.

Another topic ignored by the Massachusetts project was the examination of residual values resulting from the use of the guidelines in particular cases. This is a particularly important area because it allows unusual cases to be identified easily by the presence of extreme residual values. Since the Massachusetts project apparently did not analyze residuals from their regression model, the researchers on that project were probably not aware of the direction in which cases were being "missed" by the sentencing guidelines.

Our analysis of the apparent Massachusetts data also brought to light some interesting facts about sentencing in Massachusetts. Probably the most important of these is that the factors used in the Massachusetts sentencing guidelines are almost certainly not the factors that are best able to describe or predict the sentencing decisions of the Massachusetts judiciary. (Indeed, we suspect that the factors that, in the best of all possible worlds, would have been the strongest predictors of sentence are not even included in the Massachusetts sentencing data -- as we noted in Chapter 7.) The point

behind the development of new guidelines for Massachusetts was not, however, so much to say that the project had chosen the wrong factors to use, but rather to point out that the use of different factors would affect how well the resulting models could be made to fit the actual sentences at hand without uniformly over- or under-stating those sentences. The Massachusetts sentencing guidelines, in comparison, appear uniformly to suggest higher sentences than those imposed in the past for similar offenders who commit offenses of similar seriousness.

Notes to Chapter 8

[1] We mentioned in the preceding chapter, and reiterate here, that the Massachusetts sentencing guidelines project did not make the results of their analyses of past practices or their guidelines development analyses available to the public. Some information, usually in summary form, was made available in reports given to the judges of the Committee on Probation and Parole, and later to the remaining judges on the Superior Court. These reports were: Sentencing Guidelines Project, First Report on Analyses, 2 October 1979; Second Report on Analyses, 15 October 1979; Third Report on Analyses, 19 October 1979; and Fourth Report on Analyses, 22 October 1979. The reports contain summary information about the attributes of defendants, the extent of missing information, and the explanatory power of the various guidelines models. They do not, however, contain statistical information such as correlation coefficients, beta or B weights, standard errors of estimates, etc. Such information would have been meaningless for judges, in most instances, and we assume that this is why it was not included. However, the lack of this kind of information makes it very difficult for us to know for sure exactly what the Massachusetts sentencing guidelines project staff did, in terms of specific analyses.

[2] Again, we must note that without precise information about how the analyses were conducted, we can only make some educated guesses about how the research was conducted.

[3] We received the data from the Massachusetts sentencing guidelines project in coded form; however, we were not able to obtain the system files constructed and used in analysis by the Massachusetts project. Thus, we had to create a data set that merged a number of separate files containing specific information about defendant characteristics, charges, dispositions, recommendations, etc. We assume that the methods we used to accomplish this feat were roughly equivalent to those followed by the Massachusetts sentencing guidelines project. (By the time that we obtained the original data from that project, the project had expired; thus, it was impossible for us to check with anyone on the precise methods used by their research staff.)

[4] For test purposes, the seriousness scores were first assigned to "Accessory after the fact" offenses in correspondence with the Massachusetts sentencing guidelines project's directions. However, the agreement between the seriousness scores computed both ways was so strong that the adherence to the latter method to compute the score did not substantially change the resulting total seriousness score or significantly change the importance of items in predicting the

guidelines sentence. Thus, it seemed rational to use the latter method -- which was more in agreement with statutes on the matter -- to determine the seriousness scores.

[5] As we discussed earlier in Chapter 6, the term "persons offense" as used to determine parole eligibility should not be confused with the general classification by Massachusetts statutes of "Crimes Against Person". While most of the latter are also "persons offenses" receiving two-thirds parole eligibility, a number are not.

[6] As the Massachusetts sentencing guidelines project did not detail the exact analyses that led to the format of any of the guidelines models, we realy do not know exactly how much of the variance would be explained by the guidelines; however, the claims of the judges of the Committee on Probation and Parole led us to believe that the guidelines could account for a substantial amount of that variation. See Chapter 7 for a detailed description of the comments made at the judicial conferences where the first and second versions of the guidelines were introduced.

[7] We went to great pains to point out in Chapter 3 of this report, and have often repeated, that guidelines do not have to follow the traditional "In-Out; If in, for how long" format. Separate analyses of both of these aspects, however, seems to us to be appropriate, if for no other reason than to confirm this fact.

[8] This possibility must be acknowledged, of course. However, although we did not receive the Massachusetts data in a form ready for guidelines reanalysis, but rather had to recode and compute a number of items (including the guidelines factors) using our best judgement and some rather sketchy reports of that project, the similarity of many of our analyses results lead us to believe that differences in the data are negligible. And while minimal coding, etc. differences may produce changes in the weights assigned to factors in a regression analysis, they will be minor; certainly not of the magnitude evidenced by the injury and weapon use items noted.

[9] Logically, of course, there is no absolute limit to the scores that offenders could rack up for weapon use, injury to victims, or prior record, since those should be totally unique to the particular defendant's offense and prior criminal history; however, since none of the offenders in the data base were known to be mass murderers, and since age does, to some extent, limit the number of prior offenses that a defendant could have committed, there is a practical limit to each of those scores.

Chapter 9: Other Examples of Statewide Sentencing Guidelines

In this chapter we describe and analyze the statewide sentencing guidelines recently developed in Minnesota, Pennsylvania and Michigan. Our account of these efforts at sentencing reform will necessarily be somewhat perfunctory, in comparison to the three preceding chapters which described the development of guidelines in Massachusetts, since the resources of our project did not permit us to study pre-guideline sentencing practices, or to survey interview judges or other actors in the criminal justice systems in the other three states.[1] As was the case in Massachusetts, we had the benefit of the friendly and helpful cooperation of the researchers charged with developing guidelines in both Minnesota and Michigan. We also obtained copies of the data sets analyzed by the Minnesota researchers, and have been promised the data used by the Michigan group; we hope to produce some further analyses of these data in the future.[2] We had less contact with those involved in guidelines development in Pennsylvania, owing largely to constraints of time on us and to the relative recency of their guidelines; at the time of this writing, however, we are negotiating with the Pennsylvania group to obtain a copy of their data as well.[3]

Our primary interest, in this chapter, is to compare the form and content of the guidelines developed in Minnesota, Michigan and Pennsylvania with each other, and with those of the Massachusetts and New Jersey guidelines described earlier in this report. In all three of these states, the guidelines purport to be "empirically based"; but as we shall see, the interpretation of that term differs markedly across the three states, and the extent to which their guidelines are meant merely to mimic a model of antecedent sentencing practice correspondingly varies. In all three states, the guidelines are presented in the form of a matrix, like that originally proposed by Gottfredson and Wilkins. We also use this chapter, therefore, to illustrate some numerical and statistical techniques for analyzing such matrices, which we believe facilitate comparison of (and thus, in a sense, evaluation of) sentencing guidelines that can be put in that form.[4]

The Minnesota Sentencing Guidelines

The Minnesota Sentencing Guidelines Commission (MSGC) was created by that state's legislature in 1978, and directed to develop sentencing guidelines which established (a) the circumstances in which the imprisonment of offenders was proper, and (b) a presumptive fixed sentence for offenders imprisoned, based on appropriate combinations of offense and

offender characteristics. The Commission's legislative mandate instructed it to "... take into substantial consideration current sentencing and release practices and correctional resources, including but not limited to the capacities of local and state correctional facilities."[5] According to Martin's recent (1981) account, the legislature's decision to opt for guidelines as a technique of sentencing reform was something of a compromise; there had been earlier, and unsuccessful, attempts to introduce various forms of presumptive sentences, and to abolish the state's parole board, not only because of dissatisfaction with perceived disparity in sentencing (especially prison terms), but also because it was felt by some legislators that sentences were not sufficiently severe.[6]

The Commission is composed of nine members, of whom three are judges (one from the Minnesota Supreme Court, and two from the District Court); the other members include the Commissioner of Corrections and the chairman of the Minnesota Corrections Board (i.e. the parole board), a prosecutor, a public defender and two "citizen representatives" appointed by the Governor, of whom one (Jan Smaby) is the Commission's chairman.[7] The director of the Commission's research staff, Dr. Dale Parent, had previously developed the guideline matrix used by the parole board; the research director, Kay Knapp, had previously worked in the state's department of corrections. The Commission first met in June 1978; research began early in 1979, and the guidelines themselves were presented to the legislature on 1 January 1980. There were no objections to the guidelines in the legislature; accordingly, the guidelines came into force on 1 May 1980.

In order to comply with the legislature's direction to "...take into substantial consideration current sentencing and release practices...", the Commission necessarily had to follow a very different research strategy than that suggested by Gottfredson and Wilkins and other developers of empirically-based sentencing guidelines. Minnesota's sentencing system was a classically "indeterminate" one, in which offenders who were imprisoned were given a maximum sentence by the judge, which could be any term up to the maximum provided by the statute under which they were convicted. The length of time actually spent in prison was then decided by the Minnesota Corrections Board (using, in recent years, parole guidelines). Thus, as the Commission noted in its report to the legislature (MSGC, 1980:2) under the law in force when it began its work, the "real" judicial decision was whether or not the offender should be given a state prison sentence; since, in general, judicial decisions as to length of term did not constrain the parole board's discretion, those decisions were described as being merely "symbolic".

The Commission thus could not have developed sentencing guidelines which purported to describe previous judicial term-fixing -- since, in effect, judges had no say in decisions as to length of term. This meant that two separate samples of cases had to be selected, in order to take account of past practice: a sample of convicted offenders with which the decision to incarcerate or not could be studied, and a sample of cases dealt with by the parole board, from which information on lengths of term could be obtained.

It seems clear that, in any case, neither the Commission nor its staff was in fact much attracted to the idea of guidelines that would merely be descriptive of past sentencing or releasing practice.[8] In its final report the Commission stated that most sets of guidelines developed up to that time had avowedly aimed at replicating existing practice as closely as possible; but it noted that there were several problems which "precluded" this approach in Minnesota. First, most descriptive guidelines had been developed in single-county or metropolitan court jurisdictions, in which a consistent sentencing practice was more likely to exist because norms, culture and clientele of the courts were likely to be less variable than across a whole state, and because judges had better opportunities to communicate with each other; by contrast, the Commission claimed, it was less likely that a single "usual" or customary sentencing policy or practice would be found within the whole of a heterogeneous statewide system, especially one as large and diverse as Minnesota's. Second, the Commission appealed to the concept of the separation of powers, noting that legislatures define crimes and set parameters for punishment, and that sentencing decisions could have financial and other implications both for the state legislature and for local government; it interpreted the intention of the legislature to entail guidelines which took into account these wider implications (MSGC, 1980:3).

It may be thought that the first of these arguments is not a very strong one. In its analysis of past sentencing practice, the Commission's staff did find some apparent evidence for the existence of variation in percentages incarcerated, across the state's ten judicial districts (see MSGC, 1979:12-15); but this evidence is only briefly described in the one report on this subject that staff prepared for the Commission, and in several instances the analysis is based on small numbers of cases; it is far from clear, in fact, just how much variation (in the construction data) there was, after controlling for differences in patterns of current offense and prior record.[9] In any case, even if there had been considerable between-county variation around a statewide average practice, this would not be a conclusive reason not to use that average as a basis for guidelines.

The Commission's second argument against purely descriptive guidelines seems to us to be on much firmer ground; as we saw in an earlier chapter, the separation-of-powers doctrine was raised by Massachusetts legislators critical of that state's judicially-developed guidelines. Whether the argument has the legal force that its Massachusetts advocates thought it had in that context, it clearly has force in the context of guidelines developed at the instigation of the legislature, as was the case in Minnesota. However that may be, the upshot was that the Minnesota Commission took an unashamedly "policy-oriented" approach to the development of its guidelines, rather than trying to develop sanctioning rules that mirrored past practice. "As a result of our systemwide concern, our guidelines have a greater normative content than prior efforts. In developing such guidelines, we have been informed by, but not bound to, current practice" (MSGC, 1980:3).

The Commission was also directed to take into substantial consideration "correctional resources, including but not limited to the capacities of local and state correctional facilities". In drafting its guidelines, the Commission interpreted this part of its mandate to mean that the guidelines should produce prison populations not exceeding the current capacity of state institutions (no mention was made of local jail capacity); to this end it developed a computerized projection model to simulate the prison populations which would likely result from various guidelines structures. Both the designation of certain cells as presumptively "in" or "out", and the presumptive ranges for "in" cells, appear to have been influenced to some extent by these projections.[10]

The guidelines which resulted from the Commission's work (and that of its staff) are shown in Table 9.1, which is reproduced from the Commission's final report to the Minnesota legislature (MSGC, 1980:38). It will be seen that the guidelines take the familiar two-dimensional matrix form, in which the rows reflect ten ordered levels of severity of conviction offense (the offenses mentioned being common examples) and the columns reflect seven categories of "criminal history score"; the resulting seventy cells contain presumptive terms, or ranges, associated with the various combinations of offense and offender history.[11] The heavy line which staggers upward and to the right from row VI is referred to as the "dispositional" or "in-out" line; cases in cells falling above this line are presumptively to be given "out" sentences, whereas cases in cells below it are presumptively to receive state prison terms.[12] In the presumptively "out" cells, only a single term (e.g. 12 months) is given; this is the term which is to be imposed if the presumption against a state prison sentence is overridden

because of the presence of "substantial and compelling" aggravating factors (or if a non-incarcerative disposition is later revoked). In the cells below the "dispositional" line, a range of months is provided, in addition to the mid-range. The Commission's statutory mandate would have permitted ranges of up to plus or minus fifteen percent of the presumptive (i.e. mid-range) term; in the event, it was decided to use much narrower ranges -- typically, it will be seen, they are plus or minus about five to eight percent -- on the ground (for which no evidence whatever was produced) that "broad ranges would increase the treatment of disparate cases and, in a sense, would allow disparity to continue in practice while defining it away in theory" (MSGC, 1980:12).[13] Though the Commission's report is somewhat ambiguous about this, it is stated that the range provided in the presumptive "in" cells is intended to reflect "legitimate, but not substantial and compelling, differences among cases" (MSGC, 1980:12). Thus variations within the ranges in those cells do not count as "departures" from the guidelines, and written reasons are not required in such cases.

Insert Table 9.1 here

It is important to note that the Minnesota guidelines, at least in the form in which they were initially implemented, apply only to state prison sentences, and that offenders who are given what are presumptively "out" sentences according to the guidelines may in fact spend up to a year in a county jail or similar facility. Moreover, if a prison sentence for such a case is stayed through a stay of execution of the sentence (as distinct from a stay of imposition), the judge may impose conditions relating to the stayed term which exceed in duration the prescribed presumptive term of imprisonment, and that could be as long as the statutory maximum sentence. For example, though the guidelines may set a twelve-month presumptive prison term if an offender is incarcerated, the judge could stay that sentence, placing the offender on probation (or giving him a sentence of jail plus probation) of up to three years, for the unauthorized use of a motor vehicle. If the stay were later to be revoked, the twelve-month presumptive prison term could then be invoked.[14]

The offense severity levels used in the guideline matrix were not derived from any empirical analysis of past sentencing practice, e.g. by estimating changes in the probability of incarceration or differences in lengths of terms across the different offense types. Instead, the Commission itself worked out an "offense severity table" in

Table 9.1: The Minnesota Sentencing Guidelines Matrix

(Source: MSGC, 1980:38)

SEVERITY LEVELS OF CONVICTION OFFENSE		CRIMINAL HISTORY SCORE						
		0	1	2	3	4	5	6 or more
Unauthorized Use of Motor Vehicle Possession of Marijuana	I	12*	12*	12*	15	18	21	24
Theft Related Crimes ($150-$2500) Sale of Marijuana	II	12*	12*	14	17	20	23	27 25-29
Theft Crimes ($150-$2500)	III	12*	13	16	19	22 21-33	27 25-29	32 30-34
Burglary - Felony Intent Receiving Stolen Goods ($150-$2500)	IV	12*	15	18	21	25 24-26	32 30-34	41 37-45
Simple Robbery	V	18	23	27	30 29-31	38 36-40	46 43-49	54 50-58
Assault, 2nd Degree	VI	21	26	30	34 33-35	44 42-46	54 50-58	65 60-70
Aggravated Robbery	VII	24 23-25	32 30-34	41 38-44	49 45-53	65 60-70	81 75-87	97 90-104
Assault, 1st Degree Criminal Sexual Conduct, 1st Degree	VIII	43 41-45	54 50-58	65 60-70	76 71-81	96 89-101	113 106-120	132 124-140
Murder, 3rd Degree	IX	97 94-100	119 116-122	127 124-130	149 143-155	176 168-184	205 195-215	230 218-242
Murder, 2nd Degree	X	116 111-121	140 133-147	162 153-171	203 192-214	243 231-255	284 270-298	324 309-339

(1st Degree Murder is excluded from the guidelines by law and continues to have a mandatory life sentence.)

*one year and one day

which (to begin with) all "commonly occurring" felonies were arranged into six categories -- property crimes, crimes against persons, sex offenses, drug offenses, arson offenses, and a miscellaneous category. Commission members then individually ranked each offense in each of these categories, in order of decreasing severity; 104 offense types in all were ranked in this way. Average ranks for all nine Commission members were then calculated, and a further discussion was held among Commission members to try to establish consensus on the most serious, second most serious, etc., offense type in each of the six groups. This process was repeated until all 104 crime types had been placed in an agreed rank order of severity; when differences between Commission members on rankings occurred, "the members articulated reasons for their preference, and sought to persuade other members to their viewpoint" (MSGC, 1980:7). After consensus was obtained, the Commission divided the overall ranking into ten categories, within which offenses were "deemed to be generally equivalent" in severity (loc. cit.).[15]

It should be noted that this procedure, however agreeably rationalistic it may seem on its face, cannot purport to produce anything more than an ordering of the 104 offenses considered, and gives no ground for saying <u>how much</u> more severe offenses in, say, category VI (which were all "deemed to be generally equivalent" to each other) were than offenses in categories V or IV. The resulting ten-category classification looks reasonable enough to us, on its face; but the procedure by which it was derived gives no empirical basis for relating "severity of offense" (as collectively perceived by the Commission) and the penalties provided by the guidelines. Relative severity is, of course, an element which is used to <u>justify</u> variations in severity of penalties. But while the Commission's ranking procedure may help to justify the claim that (e.g.) aggravated robbery and other category VII offenses are <u>more</u> serious than the offenses in categories I to IV, it cannot justify the claim that category VII offenses are, say, <u>twice as severe</u> as those in categories I-IV, though that is approximately what the guidelines penalties provide.

The "criminal history score" developed by the Commission (which was modified somewhat in the course of their deliberations, before taking its final form) was equally judgemental or non-empirical -- which is not, of course, to say that it was arbitrary. In its report to the legislature (MSGC, 1980:7-8) the Commission stated that it sought to develop an index which (a) was consistent with then-current sentencing and releasing decisions, (b) was based on objective and readily available records, (c) was simple to use, and (d) did not rely on economic or social status variables. The

index is computed in the following way. Offenders are given one point for every prior felony conviction for which a sentence was either stayed or imposed, with multiple sentences for a single course of conduct being counted as one point; there is a maximum limit of two points for prior multiple sentences arising out of a single course of conduct in which there were multiple victims. If they occurred five years or more before the current conviction, stayed felony sentences are to be treated as if they were for misdemeanors or gross misdemeanors, and prior felony convictions are not counted at all if (in effect) the offender had had a crime-free period of ten years or more since the expiration of his or her last sentence. Prior misdemeanor convictions are given one "unit" (not one point), and prior gross misdemeanors (excluding traffic offenses) are given two units, subject to certain conditions;[16] four such "units" equal one point in the criminal history index, though no more than one point on the index can be given for such convictions. The offender is also given one point on the index if he or she was on probation or parole, or in an institution, at the time of the current offense; and can receive one point for every two juvenile adjudications for offenses committed between the ages of 16 and 21 that would have been felonies if committed by an adult, subject again to a maximum of one index point for such crimes.

It is perhaps sufficiently obvious -- from the complexity of the conditions by which the criminal history index is defined -- that it was not derived from any sort of empirical analysis of antecedent practice, but was simply made up by the Commission, no doubt after at least as much deliberation as was involved in constructing the offense seriousness score. It would, we think, be difficult even to write down an equation reflecting the criminal history conditions specified in the index, much less to estimate that equation from the kinds of data (discussed below) which the Commission's staff had at its disposal. We should emphasize that we do not intend this statement as being in any sense a criticism of the Commission's procedures; quite the opposite. However, some of the statements made (by the Commission's staff to the Commission itself, and by the Commission to the legislature in its report[17]) are at best misleading on this point. For instance, in the report and several other documents made available to us by the Commission and its staff, it is stated that "the most significant factor in judicial decision making was the criminal history of the offender" (MSGC, 1980:5, emphasis added).[18] But the truth of that claim depends entirely on how "criminal history" is defined; and it may well be that other definitions of that variable -- for instance, excluding convictions between the ages of 16 and 21 -- would have led to a different result. We have not, within the time limits of our project, been able to carry out secondary

analyses of the Minnesota construction data sets which would enable us to settle this point, though we hope to do this in the future. The fact remains that the Commission (and/or its staff) several times stated unambiguously that "criminal history" had previously been the "most significant" factor in prior sentences: yet in constructing the Minnesota guidelines, the Commission went out of their way to propose a set of sanctioning rules which gave a different result -- i.e. sentences scaled mainly to the severity of the offense as they defined it. The empirical data which they presented, to us and the Minnesota legislature, do not give a very strong warrant for their description of antecedent sentencing practice.[19]

The guidelines shown in Table 9.1 did not purport merely to mirror sentencing or term-fixing practices. Even so, how much did what they prescribe differ from prior practice? Before tackling this question we describe briefly the data on which the Commission's staff's analysis of prior practice was based. In order to study the "in-out" decision, staff selected a 42 percent random sample of male offenders convicted and receiving a felony or gross misdemeanor sentence in fiscal 1978, together with all of the females similarly convicted in that year; persons convicted of felonies, but receiving misdemeanor sentences, were excluded from this sample.[20] Counties with large Indian (i.e., native American) populations were oversampled, in order to obtain sufficient cases of this kind for analysis; in all, the "dispositional" sample contained about one-half of all persons convicted of felonies in Minnesota in fiscal 1978 (N = 4,369 in the Commission's report).[21]

As we have already noted, time served in state prison before the guidelines could not be estimated from judicially-imposed sentences. Instead, the Commission estimated this from data on all persons first released on parole (or released not under supervision) in the fiscal year 1978 (MSGC, 1979:2). In passing, we should note that this may, at least in theory, be far from an ideal group of prisoners with which to study the question of time served to first parole. For a variety of reasons, the cohort of prisoners released from prison in any given year may have served very different amounts of time in prison from the cohort of offenders admitted to prison in the same year; and it is surely the latter group whose time in prison one wants to consider, in any meaningful comparison of time to be served by prisoners in the future and time served by prisoners sentenced in the past.[22] In practice, if sentencing and paroling practices remain relatively constant, the differences between admission and release cohorts' experiences may not be great; and it is certainly easier, as a practical matter, to

obtain information on the latter group. Nonetheless, there may be substantial differences between the two; and we do not feel, for this reason, that we can place great weight on the Minnesota Commission's estimates of time served in prison to release or first parole under the guidelines.

Table 9.2 sets out the percentages incarcerated, in the Minnesota court sample, with cases classified according to the Commission's offense severity and criminal history scores; as in Table 9.1, the heavy zig-zagging line is the "in-out" line, above which prison terms are presumptively stayed in the guidelines themselves. It will be seen that this line does in fact reflect a fairly sharp distinction between cases imprisoned and those given probation or other non-prison dispositions; only about 15 percent of those in cells above the line had been sent to prison (in the cases sentenced in fiscal 1978), whereas over two-thirds of cases in cells below the line had been imprisoned. It is also clear that the percentages incarcerated tend generally to increase, as one moves down the rows and from right to left across the columns of Table 9.2, as one would expect. However, there is a fair number of cells in this table in which the presumptive term dictated by the guidelines is different from the modal sentence before the guidelines. Thus in the three cells for offense severity levels VII, VIII and IX, for offenders with criminal history scores of zero, fewer than half of the cases in the pre-guidelines data were imprisoned; but according to the guidelines themselves, these cases are to be treated as presumptively "in". Similarly, in no less than seven of the 27 cells above the line, the majority of cases in the pre-guidelines data had been imprisoned; yet according to the guidelines these are presumptively "out" cells.[23]

Insert Table 9.2 here

Table 9.2 also shows (in parentheses) the numbers of cases in each cell defined by the guidelines, in the 1978 court sample. It is evident that these are very heavily clustered in the upper left-hand corner of the matrix, and that convictions for the more serious crimes -- those presumptively to be given prison sentences under the guidelines -- are relatively few in number, as are offenders with the more serious prior criminal histories. In fact, the eight cells defined by the top four rows and two left-hand columns of the matrix contained (in the 1978 data) about two-thirds of all of the cases in the sample; less than eight per cent of those cases were sent to prison. At the other extreme, the bottom four rows of the table -- containing offenses ranging in seriousness from aggravated robbery up to

Table 9.2: Percentages Imprisoned in Minnesota Court Disposition Data, Classified by Guideline Categories. Numbers in Parentheses are Weighted N's on which Cell Percentages are Based. (Source: Data Provided by MSGC.)

		History/Risk Score							
		0	1	2	3	4	5	6+	Totals
Offense Severity:	I	4 (474)	13 (126)	44 (69)	59 (32)	56 (23)	26 (8)	74 (16)	15 (748)
	II	6 (476)	24 (90)	27 (82)	56 (24)	85 (14)	53 (18)	100 (11)	16 (715)
	III	6 (532)	16 (169)	40 (100)	57 (79)	79 (35)	72 (16)	58 (20)	20 (951)
	IV	6 (566)	19 (186)	42 (139)	44 (34)	86 (30)	62 (23)	77 (11)	19 (989)
	V	17 (131)	37 (36)	78 (14)	83 (13)	80 (10)	100 (2)	50 (4)	33 (211)
	VI	12 (231)	22 (78)	45 (58)	86 (15)	66 (13)	61 (14)	100 (4)	25 (413)
	VII	39 (97)	68 (57)	86 (28)	85 (15)	100 (11)	100 (4)	100 (7)	62 (219)
	VIII	42 (46)	38 (26)	87 (16)	100 (6)	100 (10)	100 (2)	- (0)	58 (106)
	IX	35 (6)	- (0)	- (0)	100 (4)	100 (2)	- (0)	- (0)	68 (13)
	X	100 (13)	100 (5)	100 (4)	- (0)	- (0)	- (0)	- (0)	100 (22)
	Totals	9 (2571)	24 (774)	45 (511)	63 (222)	80 (149)	62 (88)	77 (72)	23 (4387)

second degree murder -- contain only about eight percent of the cases in the table.

In part for this reason, the location of the presumptive "in-out" line may prove to have a substantial effect on the numbers of persons sent to prison in Minnesota. We can see what that effect might be, by assuming that the presumptive sentence is imposed in every case sentenced under the guidelines, and that the distribution of cases within cells remains approximately the same as it was in the 1978 court sample. If this were so, all persons in cells falling below the line would be sent to prison, whereas all those in cells above the line would be given non-prison dispositions; the total proportion imprisoned would then fall to about 10.5 percent, or roughly half what it was in the 1978 sample. Of course, such an extreme assumption is unlikely to be met, in practice; some proportion of cases above the line are likely to be treated as departures from the guidelines, and sent to prison, whereas some proportion of those below the line will probably be given "out" sentences.[24] The point is that the location of the presumptive line -- deliberately chosen by the Commission on grounds of policy, to reflect a "modified just deserts" approach in which severity of current offense would have a substantial effect on sentencing policy[25] -- may lead to a considerable reduction in the numbers of persons committed to prison in Minnesota in the future.

How do the lengths of prison sentences provided by the guidelines compare with time spent in prison by offenders released in 1978? Repeating our earlier warning that that sample of cases may not give a very good indication of expected time to be served by those committed to prison immediately before the guidelines were implemented, we address this question in Table 9.3. This table shows mean times spent in prison to first release on parole, for the cases in the 1978 sample classified as they would be under the guidelines.[26] In comparing this table with the presumptive terms shown in Table 9.1, it must be borne in mind that state prison sentences in Minnesota now provide for "good time" of one-third of the amount imposed;[27] thus the figures in Table 9.1 must be reduced by one-third, to yield an estimate of the time most prisoners will probably serve.[28] Caution must also be used in making comparisons with Table 9.3, since the numbers of cases in many of the cells are small. With those caveats in mind, it may be noted that the typical estimated time to be served in prison under the guidelines is less than that actually served by prisoners released in 1978, in the great majority of cells in the matrix. In only a few cells -- mostly in the lower right-hand corner of the matrix, do the guidelines prescribe longer terms than were served by similarly classified offenders in the 1978 sample.

Insert Table 9.3 here

As already noted, the lengths of terms prescribed by the guidelines, and to some extent the location of the "in-out" line, were based in part on their likely effects on the Minnesota prison population. In its discussion of the probable impact of the guidelines, the Commission's report to the legislature states that "the guidelines were developed so that the average projected population of the state prison system would be 5 percent below capacity" (MSGC, 1980:14). The Commission also stated, however, that the dispositional line and the suggested ranges should have a substantial impact on the types of offenders in state prisons; it estimated that there would be more murderers, robbers, sex offenders and assaulters there than there had been before the guidelines came into force, since those serious offenders against the person typically receive longer sentences than property offenders, most of whom would presumptively be given non-prison sentences. It estimated that over a five-year period the proportion of "person offenders" in state prisons should increase to about 74 percent, compared with 58 percent before the guidelines (MSGC, 1980:15).[29] It also estimated that the prison population would tend to become slightly older, and somewhat more metropolitan in origin, than previously, though there would be few changes in commitment rates by race or sex. Finally, the Commission noted that there would probably be a gradual build-up in the prison system of persons serving sentences of longer than five years -- to 26 percent, as compared with 18 percent previously -- and a concomitant decrease in the proportion serving terms of three to five years. This last effect, it noted, should be "gradual and manageable" (MSGC, 1980:18).

It will be interesting to see whether this turns out to be correct, even if the assumptions underlying the Commission's prison population projections turn out to be more or less met.[30] The computer program by which the Commission's estimates of future population were derived only provided estimates for a five-year period. However, just because the build-up of long-term prisoners resulting from this change in sentencing policy is gradual, it may well be that the full effects of the longer terms provided by the cells in the lower right-hand corner of the matrix will not be felt until much later than that. (An offender sentenced under the guidelines for second-degree murder will, if he has a criminal history score of six or more, stay in prison for 18 years, with full allowance for "good time". Assume that there is only one such offender committed to prison per year; their numbers will not "level off" for 18 years, at which point they

Table 9.3: Mean Times to First Parole in Minnesota Prisoner Sample, Classified by Guideline Categories. Numbers in Parentheses are N's on which Cell Percentages are based. (Source: Data Provided by MSGC.)

		History/Risk Score							
		0	1	2	3	4	5	6+	Totals
Offense Severity:	I	14.5 (23)	9.8 (15)	15.9 (20)	17.2 (26)	15.0 (8)	28.7 (5)	26.8 (2)	15.8 (99)
	II	9.4 (24)	12.6 (18)	12.7 (23)	12.1 (9)	14.9 (7)	24.4 (4)	13.7 (6)	12.5 (91)
	III	18.3 (25)	13.4 (23)	18.3 (25)	20.4 (22)	16.0 (7)	17.4 (6)	19.8 (7)	17.7 (115)
	IV	14.0 (58)	18.2 (33)	18.3 (47)	25.2 (32)	26.1 (12)	21.8 (10)	31.9 (11)	19.5 (203)
	V	28.2 (21)	27.5 (15)	28.3 (9)	33.0 (4)	22.0 (4)	50.4 (1)	38.1 (1)	29.0 (58)
	VI	21.2 (26)	26.5 (21)	28.8 (8)	29.2 (12)	42.6 (6)	33.4 (3)	28.2 (4)	27.0 (80)
	VII	30.8 (54)	32.2 (30)	38.5 (25)	32.7 (16)	50.6 (9)	38.7 (3)	47.9 (7)	34.9 (144)
	VIII	41.8 (16)	55.3 (6)	61.4 (4)	30.2 (3)	42.2 (4)	50.9 (2)	- (0)	45.9 (35)
	IX	73.7 (1)	- (0)	- (0)	42.6 (1)	- (0)	35.9 (1)	- (0)	50.7 (3)
	X	72.1 (6)	87.3 (1)	109.5 (1)	144.7 (1)	- (0)	- (0)	142.3 (1)	91.6 (10)
	Totals	22.9 (254)	22.5 (162)	23.0 (162)	24.5 (126)	28.3 (57)	27.6 (35)	32.6 (42)	24.1 (838)

will be over three and a half times greater in number than a five-year projection would suggest.[31]

At this point we make use of the Minnesota guidelines matrix to illustrate a number of techniques for analyzing the structure of this form of sanctioning rule. By "structure" we mean the relations between the ranges and presumptive terms shown in Table 9.1, without reference to the numbers of cases falling within each cell of the matrix or the extent to which the guidelines are either followed or departed from in practice. The statistical and numerical techniques which we shall use were mostly developed by Tukey (1977) and his colleagues; they are of quite general application, and were not invented for the purpose of studying sentencing or any other topic in the field of criminal justice. As it happens, however, they are ideally suited to one of the problems with which we are primarily concerned: to wit, the relations between the term or terms prescribed by a guidelines matrix, and the elements (such as severity of offense and prior criminal history) by which that matrix is defined.

We have already noted that the Minnesota matrix, like most others used to structure decision-making in sentencing and parole, consists of rows defined by offense severity and columns defined by prior criminal record, with each of these variables being ordered from "best" to "worst"; and that the prescribed terms (or ranges of permissible "normal" terms) tend to increase as we move across either of the two defining dimensions.[32] Thus the prescribed terms for those convicted of the least serious offenses, and who have the less serious prior criminal records, are in the cells in the top left-hand corner of the Minnesota matrix; conversely, the most serious offenses, committed by those with very bad prior records, are in the cells at the bottom right-hand corner of that matrix. A reasonable first step in analyzing the structure of the matrix, therefore, is to ask: exactly how does this passage from the "best" (or, perhaps, "least bad") cases to the "worst" ones come about? Are the "steps" from one row to the next, or from one column to the next, reasonably regular ones? Or are there some cells that are markedly "out of line" with the general picture presented by the rest of the matrix? Can we, for example, say that the prescribed terms in particular cells of the matrix can be well represented by a simple combination of an "offense severity" effect, and a "prior criminal record" effect -- perhaps a simple addition of those two things?

In order to tackle this question, we need a single number (e.g. months for which sentenced, or to be served until release on parole), rather than a range. In the case of the Minnesota sentencing guidelines, it is reasonable to use the

mid-range for those cells that provide for presumptive "in" sentences, since the guidelines themselves appear to treat that as the presumptive term in the absence of "legitimate, but not substantial or compelling" differences between the cases.[33] For those cells in which the presumptive sentence is "out", we will use the single term provided if offenders are sent to prison (i.e. are treated as a departure from the guidelines). This is admittedly not entirely satisfactory, since the intention of the Minnesota commission was that the majority of cases falling into the cells above the "in-out" line in Table 9.1 would not be sent to prison at all. However, the question we wish to address here is the relations between lengths of terms imposed if offenders are imprisoned. How are those lengths affected by different combinations of offense severity and prior record?

We begin by seeing whether the cells of the Minnesota matrix can be represented as the sum of an "offense effect" reflecting levels of severity, and an "offender effect" reflecting variations in prior criminal history. To investigate this, using Tukey's (1977) method, we first calculate medians across each of the ten rows of the matrix, and subtract those medians from each of the cells in its corresponding row, to get a matrix of residuals; we next calculate medians down the columns of the residual matrix, and subtract those medians to get a second matrix of residuals; we go on subtracting medians, first across the rows and then down the columns, of the successive sets of residuals, until the medians all become zero (within limits of rounding error). The successive row and column medians are then added together and displayed outside the final matrix of residuals, as in Table 9.4. The extreme right-hand column in this table (outside the double vertical lines) is the result of taking the median down the column of final row medians, and subtracting this out to give, in effect, an overall median for the whole matrix (which is 32 months, in Table 9.4). This is the more usual way to represent the results of this kind of analysis; but for our purposes the two columns headed "offense effects" in Table 9.4 can be thought of as exactly equivalent ways of presenting the same information.[34]

Insert Table 9.4 here

Table 9.4 asks, in effect, how well we can represent the Minnesota matrix by an additive model in which cell values are predicted by simply adding together a certain number of months associated with level of offense severity and a certain number of months associated with level of prior record. (Alternatively, can we get close to the actual cell values by

Table 9.4: Additive Model Fitted to Mid-ranges of Minnesota Matrix
(Effects and Residuals in Months)

		Criminal History Score							Row Effects	
		0	1	2	3	4	5	6+		
Offense Severity Level:	I	9.5	4.5	.5	0	-6	-12	-18.5	15	-17
	II	7.5	2.5	.5	0	-6	-12	-17.5	17	-15
	III	5.5	1.5	.5	0	-6	-10	-14.5	19	-13
	IV	3.5	1.5	.5	0	-5	- 7	- 7.5	21	-11
	V	.5	.5	.5	0	-1	- 2	- 3.5	30	- 2
	VI	-.5	-.5	-.5	0	1	2	3.5	34	2
	VII	-12.5	-9.5	-4.5	0	7	14	20.5	49	17
	VIII	-20.5	-14.5	-7.5	0	10	19	28.5	76	44
	IX	-39.5	-22.5	-18.5	0	18	38	53.5	149	117
	X	-74.5	-55.5	-37.5	0	31	63	93.5	203	171
Column Effects:		-12.5	- 7.5	- 3.5	0	9	18	27.5	---	32

adding together the overall median of 32 months, plus or minus an amount associated with offense severity and an amount associated with prior record?) Consider, for example, the cell corresponding to offense severity level V, and a prior criminal history score of 2; in the matrix itself (Table 9.1), this cell contains a presumptive term of 27 months. Our additive model combines an offense severity effect of 30 months for that row, with a prior history effect of -3.5 months for that column; the algebraic sum of these is 26.5 months. (Equivalently, the overall median of 32 months minus 2 months minus 3.5 months equals 26.5 months.) The residual (observed minus predicted) is .5 months, which is what is shown in the cell in question in Table 9.4.

If the additive model gave a good fit to the original matrix, we would expect that the residuals in Table 9.4 would all be close to zero. Inspection of Table 9.4 shows that, unfortunately, this is not the case. While some of the residuals are zero, and some are near it, the great majority are not. Moreover, there is an obvious pattern to the residuals in Table 9.4: those in the upper right-hand corner and the lower left-hand corners are negative. This pattern is characteristic of situations in which the additive model does not fit. While there is not, so far as we are aware, a formal goodness-of-fit test for this technique,[35] simple inspection of the residuals -- which show the extent to which the additive model's predictions "miss" -- enables us to reject that model. We cannot adequately account for the variation in presumptive terms across the cells of the Minnesota matrix by a simple additive combination of offense severity effects and prior criminal history effects.

We thus try a slightly more complicated model. Suppose we see whether the cell mid-ranges can adequately be represented by a multiplicative model, in which each cell value is predicted by an offense severity effect times a prior criminal history effect? It is easy to fit such a model, by taking logarithms of the cell mid-ranges in Table 9.1, and then extracting row and column effects in the same way as before; we happen to prefer logarithms to the base e (=2.718281828....), but any other base, e.g. 10, would of course do as well. The results of this exercise are shown in Table 9.5, in which both effects and residuals are in natural logarithms. It will be seen that the residuals in this table are mostly very small, and that they do not display the same patterning of positives and negatives that we saw in Table 9.4; this is a good sign that the multiplicative model fits the data reasonably well. (We have not presented the matrix of logarithms from which Table 9.5 was derived. However, the arithmetic of Table 9.5 is easily checked in the same way as was done for Table 9.4. Consider again the cell corresponding

to offense severity level V and prior criminal history score of 2: this is predicted by a row effect of 3.45 plus a column effect of -.17, which equals 3.28; the cell value is actually 3.29 (the logarithm of the 27 month median), so there is a residual of +.01.)

The fitting of the multiplicative model is perhaps more easily seen in Table 9.6, in which the effects and residuals have been reconverted back to months by taking antilogs. Once again we consider the cell corresponding to offense severity level V and prior criminal history score 2. The overall median (under this model) is 34.12 months; this times a row effect of .923, times a column effect of .844, gives a predicted value of 26.58 approximately; there is thus a residual of +.42 of a month (in Table 9.6 we have arbitrarily rounded residuals of less than .50 to zero).

Insert Tables 9.5 and 9.6 here

What does it mean to say that the structure of the Minnesota guidelines is "multiplicative" rather than "additive"? In the additive model depicted in Table 9.4, the number of months added for a given level of prior criminal history is the same, across all levels of offense severity; conversely, the amount added for a given level of offense severity is the same, across all levels of prior record. Starting from the overall median of 32 months, an offense of level V is worth a further two months, regardless of the offender's prior record; and a prior criminal history score of 5 is worth an extra 18 months, regardless of the severity level of the current offense.

A multiplicative model, by contrast, describes a situation in which the effect of, say, prior criminal history varies, depending on the severity of the current offense. In the Minnesota guidelines, the effect of a bad prior record is generally greater, the more serious the current offense. One way to see this is to divide the prescribed terms in each row by the element in that row's left-hand column; this gives the ratios of penalties prescribed for offenders with varying degrees of prior history, to the penalty prescribed for those with no record (more strictly, those with a prior history score of zero). Thus, for example, for those convicted of offenses of severity level I, the term prescribed for those with a criminal history score of 6 or more is 2.0 times that prescribed for an offender with a prior history score of zero; by contrast, for those convicted of offenses of level IV, the ratio is 3.42 to 1 (41 months, against 12); for offenses at level VII, such as aggravated robbery, those in the worst

Table 9.5: Multiplicative Model Fitted to Mid-ranges of Minnesota Matrix

(Effects and Residuals in Natural Logarithms)

		Criminal History Score							Row Effects	
		0	1	2	3	4	5	6+		
Offense Severity Level:	I	.37	.15	-.02	.04	.04	-.01	-.03	2.67	-.86
	II	.22	0	-.01	.01	0	-.06	-.06	2.82	-.71
	III	.11	-.03	.01	.01	-.02	-.01	0	2.93	-.60
	IV	-.01	0	.01	-.01	-.01	.04	.12	3.05	-.48
	V	-.01	.03	.01	-.06	0	-.01	-.01	3.45	-.08
	VI	0	0	-.03	-.07	0	.01	.03	3.60	.07
	VII	-.15	-.08	-.01	0	.10	.12	.14	3.89	.36
	VIII	-.02	-.01	0	-.01	.03	.01	0	4.34	.81
	IX	.13	.12	.01	0	-.01	-.06	-.10	5.00	1.47
	X	.03	-.10	-.02	.03	.03	-.01	-.04	5.28	1.75
Column Effects:		-.56	-.34	-.17	0	.18	.38	.54	--	3.53

Table 9.6: Multiplicative Model Fitted to Mid-ranges of Minnesota Matrix
(Effects and Residuals in Months)

		Criminal History Score							Row Effects	
		0	1	2	3	4	5	6+		
Offense Severity Level:	I	3.8	1.7	0	.6	.7	0	-.8	14.44	.423
	II	2.4	0	0	0	0	-1.5	-1.8	16.78	.492
	III	1.3	0	0	0	0	0	0	18.73	.549
	IV	0	0	0	0	0	1.1	4.8	21.12	.619
	V	0	.6	0	-1.5	0	0	0	31.50	.923
	VI	0	0	-.9	-2.6	0	.5	2.2	36.60	1.073
	VII	-3.9	-2.8	0	0	6.5	9.5	13.1	48.91	1.433
	VIII	-.8	-.6	0	-.7	3.2	.8	0	76.71	2.248
	IX	12.3	13.3	1.7	.5	-1.6	-12.0	-24.7	148.41	4.349
	X	3.9	0	-3.7	6.6	7.9	-3.1	-13.0	196.37	5.755
Column Effects:		.571	.712	.844	1.0	1.197	1.462	1.716	--	34.12

class of prior criminal history are prescribed terms that are over 4.0 times as heavy as those with a criminal history score of zero.[36]

We are not, of course, claiming to have discovered that the Minnesota guidelines were actually made up with a multiplicative model in view, or that the Minnesota Commission deliberately planned the various prescribed terms so that they would display the multiplicative model entails. So far as we know, they did not.[37] All that we are saying is that the guidelines do in fact generally take that form (the multiplicative model fits quite well, whereas the additive one fits badly). It is of interest to note, however, that several decision-making guidelines matrices developed in recent years seem to take approximately this form. In the next section, we show that this is the case with the Pennsylvania sentencing guidelines; elsewhere it has been shown that the same thing is true for the guidelines matrices used by the Oregon parole board, and the U.S. Parole Commission.[38]

Nor are we suggesting that the Minnesota guidelines (or any others) necessarily should take an additive rather than a multiplicative form, or indeed that they should take any other form. The point of our analysis is not to argue that guidelines in which prior criminal history has the same effect across all categories of offense severity are in some sense better or fairer. The analytical techniques employed here are, in our opinion, useful in that they bring out structural features about the guidelines that bear on sentencing policy which might otherwise not be obvious; they thus make possible critical analysis and evaluation of the matrix. (They also make possible a comparison of guidelines from different states.) We have just seen that the Minnesota guidelines do in fact take a multiplicative form; why this is so, and whether it should be so, are separate questions. Are there sound reasons of the principle for such a structure -- and if so, what are they?

Or again, suppose that a generally coherent and defensible structure can be discerned in the guidelines; but that a few cells here and there are markedly "out of line" with what would be expected from that general model, in that they contain prescribed terms that are markedly heavier or lighter than the rest of the matrix would suggest? Even if that were to happen, it would not necessarily be wrong; it might be that there was in fact something special about those combinations of offense type and prior record, such as to justify a marked difference from the pattern displayed by the rest of the cells.

In other words, using the technique of extracting a general structure from this kind of matrix, and noting marked departures from that structure (i.e., large residuals, whether or not these appear to form some pattern), we can reveal apparent anomalies in the matrix. These may turn out to be defensible, or they may not; but at least they are made more obvious by this kind of analysis than they would be in the matrix itself. For example, we just noted that the ratio of the prescribed term for those with the worst records to those with the best prior records increased, as offense severity increased. There may be good arguments for this (though we cannot, as it happens, think of any). The increase, however, is not a monotonic one. Instead, the ratio of worst-to-best increases up to offense severity level IV, then decreases, then reaches its highest value (4.04) at severity level VII, and then declines again. Why should this be? Why should prior record have a much heavier impact, relatively speaking, for those convicted of aggravated robbery (or other category VII crimes) than it does for those convicted of second-degree murder (or other category X crimes)? It is not obvious what the justification for this might be.

We observed that the residuals left after fitting a multiplicative model to the Minnesota matrix were generally small; in fact, 29 of the 70 are small enough that they can be regarded as being effectively zero. In Table 9.7 we plot the remaining (non-zero) residuals in a way which helps to throw light on any patterns which they may display. Table 9.7 is called an "inside-out" plot, since instead of having row and column labels on the outside of the table it has residual values outside (on the left-hand side), and row labels -- here, offense severity levels, written with Arabic rather than Roman numerals for clarity -- inside; column labels (the prior record scores) have been kept outside the table in order to help us to see if there are different patterns within different categories of prior record.[39]

Insert Table 9.7 here

This table shows quite clearly a number of things. First, the multiplicative model fits least well where offenders with the worst prior criminal history scores (5 and 6) are concerned; there are relatively extreme values, in both a positive and a negative direction, in these two columns. Second, offenses of severity level 9 (a category including third-degree murder) have large positive residuals for those with prior history scores of zero and one, but large negative residuals among those with prior history scores of five and six or more. In other words, persons convicted of offenses of

Table 9.7: "Inside-Out" Plot of Non-Zero Residuals from Table 9.6

	Prior Criminal History Score						
Residuals:	0	1	2	3	4	5	6+
+15							
+14							
+13	9	9					7
+12							
+11							
+10						7	
+ 9							
+ 8					10		
+ 7				10	7		
+ 6							
+ 5							4
+ 4	1,10				8		
+ 3							
+ 2	2	1					6
+ 1	3	5	9	1,9	1	4 6 8	
0	8	8	6	8			
- 1				5	9		1 2
- 2		7		6			
- 3	7			10		10	
- 4							
- 5							
- 6							
- 7							
- 8							
- 9							
-10							
-11							
-12						9	
-13							10
-14							
-15							
-16							
-17							
-18							
-19							
-20							
-21							
-22							
-23							
-24							9
-25							

this kind are being given much heavier sentences than the general structure of the matrix would lead us to expect if their prior histories are relatively good; but they are being given lighter sentences than would be expected, if their prior records are of the worst kinds. A similar, though somewhat less erratic, pattern can be observed for offenses of severity level 10 (the most serious crimes covered by the guidelines). Third, for offenses of severity level 7, precisely the opposite pattern is observed: these cases receive lighter-than-expected sentences where the prior criminal history score is zero or 1, but heavier-than-expected scores when the prior criminal history score is 5 or 6. Finally, most of the residuals in Table 9.5 are positive; when the multiplicative model "misses", it seems that it predicts a sentence which is too low (compared with the observed value) rather than too high.[40]

A final disclaimer is in order concerning this kind of "model-fitting". The additive and multiplicative models described here do not by any means exhaust the range of possible structures for guideline matrices. On the contrary, it might well be that we could, with a little patience, find some more complicated model that would reduce the residuals in all cells in a matrix like Minnesota's to zero. One suggestion made by Tukey is to fit a model according to which expected cell values were given by

$$\frac{\text{Row effect} + \text{Column effect} + k\,(\text{Row effect}) * (\text{Column effect})}{\text{Overall effect}}$$

where k is some constant (cf. Tukey, 1977, chapters 11-12).

We doubt, however, that there is much point in trying to do this, in the present state of our knowledge. For we have no reason whatsoever to think that such complicated models were in the minds of those responsible for designing the Minnesota (or any other) guidelines; nor is it clear what it would mean if such a model did give a good fit to the observed prescribed terms.[41] In Chapter 3 we discussed briefly the problems of statistical modelling of judges' sentencing practice, as a preliminary to constructing guidelines; and we stressed the importance of having at least the beginnings of some theory which would guide such modelling efforts -- in contrast to the "empirical" approach which others working in this area have advocated.[42] The same point, or a related one, applies here. It is not, after all, as if we knew nothing at all about the factors that are supposed to determine variation in prescribed terms in decision-making guidelines. Quite the contrary: the Minnesota Commission was perfectly clear and explicit that severity of current offense

and prior criminal history (as they defined those things) were supposed to be the sole determinants of sentence length. The present analysis shows, in effect, the ways in which those two things were generally combined, so as to produce the presumptive terms in the guidelines; it also suggests that in a few cells -- those with large residuals -- the Commission set terms that were not consistent with the general pattern displayed by the rest of the matrix.

Before leaving the subject of the Minnesota guidelines, we report briefly on our observations of some aspects of the implementation of those guidelines. As we have noted in earlier chapters, the process of implementation may in many ways be crucial to the success of this type of sanctioning rule. Martin (1981) has described in some detail the Minnesota Commission's skillful (and generaly successful) efforts to win approval of its guidelines in the state legislature; as we shall see in the next section of this chapter, these efforts were in marked contrast to those of the Pennsylvania Sentencing Commission. We were unfortunately not able to obtain any information on this essentially political aspect of the Minnesota Commission's work, during our project; and, of course, legislative acceptance was not a part of the agenda for those responsible for developing the Massachusetts guidelines (though, as we noted in the last chapter, the problem of legislative objections to the guidelines is one which the Massachusetts Superior Court may have to grapple with in the future). But successfully neutralizing potential political opposition to guidelines among legislators is by no means all that is involved in getting those guidelines successfully implemented, even in states like Minnesota and Pennsylvania where the guidelines have a statutory mandate. In addition, those responsible for using the guidelines on a day-to-day basis -- not only judges, but also prosecutors, defense counsel and (perhaps) probation officers -- must understand and accept the guidelines. Unless this happens, there seems a high probability that this kind of sentencing reform will be subverted, however powerful its legislative mandate may be on paper.

During a site visit to Minnesota early in 1980, we attended three one-day sessions at which the Commission explained the guideliness to judges, prosecutors and public defenders respectively. At these meetings the guidelines were explained, questions about them were answered, and responses from the various audiences considered and discussed. By any standards, these sessions were impressive examples of lucidity and rationality: the contrast between them, and the introduction of the Massachusetts guidelines to judges and others in that state (described in the preceding chapters) could scarcely have been greater.

Typically, the Minnesota meetings were firmly under the control of the Commission's chairman (Jan Smaby), who presented clearly the rationale, objectives and legislative history of the guidelines; the guidelines themselves were then clearly described by the research staff (Parent and Knapp), and the basis of particular policy choices made by the Commission (e.g. in respect of the "in-out" line, and the exclusion of social factors from History/Risk score) were explained. The written materials accompanying these presentations were of an extremely high quality, as were the audio-visual aids used to display the guideline matrix; and Commission members and staff were superbly prepared for the questions which their exposition might have elicited. It seems to us to have been important that the three groups involved in these meetings -- judges, prosecutors and public defenders -- were dealt with separately; their concerns are naturally somewhat different, and by meeting with each group separately, the Commission avoided conflicts of a kind which could easily have arisen had all three been addressed together. (The fact that each group was represented on the Commission itself must also have helped: in particular, it was our observation that the initially hostile reaction from some prosecutors was largely nullified by the skillful presentation made to them by Mr. Steve Rathke, the prosecutor member of the Commission.) It may seem cynical to describe these meetings as "performances". We do not intend that description in that vein, but we do not think it inaccurate; and as performances, the three we witnessed were surely first-rate.

We have no way of knowing how far these meetings may have helped to insure the eventual success of the Commission's work in the state legislature; but it seems at least plausible to suppose that by informing such potentially powerful constituencies as judges and prosecutors, the Commission helped to nullify objections which might otherwise have arisen. Of at least as much importance, in our opinion, is the fact that the sessions which we observed almost certainly did a lot to recruit not only judges, but also other actors in Minnesota's criminal justice system, to support of the guidelines. This may not guarantee the guidelines' eventual acceptance in practice; but we do not think that it can have done any harm.

The Pennsylvania Guidelines

In Pennsylvania, statewide sentencing guidelines were developed by a Sentencing Commission created by statute in 1978.[43] As had been the case in Minnesota and Massachusetts, guidelines were first advocated as an alternative to mandatory sentencing; they were first proposed

by a state legislator, Representative Anthony Sirica, in 1976. According to Martin (1981), Sirica had learned about guidelines when he participated in a sentencing institute with Judge Marvin Frankel; he took Judge Frankel's federal sentencing guidelines commission bill as a model for legislation which he introduced, though unsuccessfully, in 1976. Two years later, after much legislative infighting with the numerous advocates of mandatory minimum sentences, Sirica's bill became law, giving the Sentencing Commission a four-year existence.[44] The Commission contained four judges, four legislators, and three other members -- a prosecutor, a defense attorney and a law professor -- appointed by the Governor. Its research director, Professor John Kramer of Pennsylvania State University, had previously conducted an evaluation of the "determinate" sentencing law in Maine.[45]

In order to understand the structure of the Pennsylvania guidelines, a brief description of the state's sentencing system is necessary. Under the criminal code which became law in 1974, offenses are graded into eight major categories, each of which has a legislatively-set maximum sentence; within that framework, judges have virtually unfettered discretion, being able to impose no penalty, a fine, probation, or partial or total incarceration. For incarcerated offenders, the judge imposes both a maximum and a minimum term; the minimum, which may not be more than half the maximum, sets the earliest time that the offender may be considered for parole if sent to prison. Offenders serving more than a five-year maximum term must go to a state prison; those serving less than a two-year maximum must go to a local jail; those serving maximum sentences of between two and five years may be sent to either a state prison or a local jail, at the discretion of the judge. There is no "good time" in the Pennsylvania system; parole release for state prisoners is decided by the state's Board of Probation and Parole, while for jail inmates it is determined (at any time during the sentence) by the sentencing court. About 80 per cent of all imprisoned offenders are released at the expiration of their minimum sentence.

It is important to note that the Pennsylvania Commission's legislative mandate did not require it to take any account of past sentencing practice, nor did it require any consideration of the capacity of correctional facilities. Nor did it require the Commission to monitor the use of the guidelines after their introduction. It merely stated that the Commission should develop guidelines which "(1) Specify a range of sentences applicable to crimes of a given degree of gravity; (2) Specify a range of sentences of increased severity for defendants previously convicted of a felony or felonies, or convicted of a crime involving the use of a

deadly weapon; (3) Prescribe variations from the range of sentences applicable on account of aggravating or mitigating circumstances."[46] In fact, the Commission did conduct a study of past sentencing practice; and in its final report to the legislature, it stated that it would monitor the use of the guidelines and make "necessary revisions."[47]

Research on past practice appears to have played virtually no part at all, in fact, in the creation of the Pennsylvania guidelines. Under Kramer's direction, a 12 percent random sample of cases sentenced in 1977 was selected (N = 2,907), and an analysis was made of both decisions to incarcerate and lengths of minimum terms. However, according to Martin (1981:75) the findings of this study were not available until the meeting at which the Commission drew the "disposition lines" and set the terms in the guidelines matrix; apparently the Commission was, in any event, not much interested in the data, and not inclined to rely on them as a basis for decision-making. Some attempts were made by staff to estimate the impact of various policy options on state prison populations; but no projection model like that used in Minnesota was developed. (In any case no data were available on actual time served by those sent to jail.)

The data collected on past practice may in fact have turned out to have a negative effect on the Commission's work and its product, since they called attention to a fact that was probably pretty widely known to many people, but might otherwise have been ignored: this is the wide range of regional variation in sentencing which existed in Pennsylvania. According to Martin, over the state as a whole, 39 percent of those convicted in 1977 were incarcerated. But in Philadelphia and Alleghany counties, which have the highest crime rates in the state and the highest proportions of violent crimes, the percentages incarcerated were only 29 and 24, respectively. In suburban areas the incarceration rate was 44 percent; in small cities it was 47 percent; and in rural areas it was 53 percent. Philadelphia judges also gave shorter sentences, particularly for misdemeanor offenses, than suburban and rural judges.

A single set of statewide guidelines would thus have necessitated substantial changes from past practice. Given the demands for a tougher attitude toward crime which pervaded the state, the idea of incarcerating fewer offenders was bound to enrage almost everyone in small-town and rural jurisdictions. The idea that Philadelphia and Alleghany county (Pittsburgh) judges should incarcerate more offenders would no doubt have been politically acceptable to many; but it would also have led to a substantial increase in prison and jail populations, and no one wanted to pay for the new prison

and jail capacity that would have been needed.[48] As matters turned out, this appears to have been one of the factors contributing to the rejection of the guidelines by the Pennsylvania legislature (in the form in which they were initially submitted to it).[49]

Two other factors, both related to the law on sentencing in Pennsylvania, further complicated the Commission's task. First, as noted earlier, the law requires judges to impose both a minimum and a maximum term for offenders who are incarcerated; the minimum cannot be more than half the maximum.

It would have been possible to create guidelines for both minimum and maximum terms. In the event, however, the Commission never got around to doing this.[50] It produced guidelines for minimum terms only; and it seems to have hoped that nobody would notice the implication that these would have for maximum sentences.[51]

The second problem confronting the Commission was that it was required to produce guidelines for both felonies and misdemeanors, and thus to regulate both state prison sentences and jail terms. The Minnesota Commission was able to sidestep this problem, since, as we have seen, its guidelines refer only to state prison sentences; an "out" sentence in Minnesota may still entail the offender's spending up to a year in jail. About 80 percent of those incarcerated in Pennsylvania in 1978, however, were sent to local jails rather than to state institutions (Martin, 1981:48). Creating guidelines to cover jail sentences meant, in effect, that the Pennsylvania Commission was making a statewide policy concerning the use of local correctional facilities. Given the variety of local policies, and the differences between counties in jail facilities, this was a politically dangerous thing to try to do. According to Martin (1981:71) there were several members of the Commission who felt strongly that the guidelines should regulate sentences for misdemeanors, since these were often very disparate; they thus rejected the suggestion (originally made by Rep. Scirica) that matrix cells for misdemeanor offenses should provide a range of "0 - 12", thus effectively leaving judges complete discretion.[52] The guidelines initially devised by the Commission (and the revised ones ultimately submitted to the state legislature) contained a number of cells in which the presumptive sentence was an "out" sentence, e.g. to probation; they also contained a number of cells providing for optional incarceration (e.g. "0 - 6"), as well as cells that provided for probable incarceration. This appears to have pleased nobody; the idea of an explicit policy that certain (minor) offenders could not be sent to jail, except as a "deviation" requiring written justification, was politically highly unpalatable.

The Pennsylvania guidelines, like those of Minnesota, take the form of a two-dimensional matrix the rows of which are defined by offense seriousness and the columns of which are defined by prior record. Neither of these two elements was derived from any kind of empirical analysis of past practice. The classification of offenses by seriousness was first done by a subcommittee which considered each offense in the Pennsylvania Crime Code and placed it into one of ten categories; no procedure such as the card-sorting technique used by the Minnesota Commission was employed, and statutory categories were deliberately rejected as a basis for ranking. According to Martin (1981:79) the results of this exercise were inconsistent; staff were asked to carry out a separate ranking, which it did partly on the basis of "principles" suggested by Professor Andrew von Hirsch, who acted as a consultant to the Commission.[53] The two sets of rankings were finally harmonized, though apparently not to everybody's satisfaction.[54] The Commission's legislative mandate was interpreted to require it to add two further levels of offense severity: basic rankings were to be increased by one level for the use of a deadly weapon, and one level for crimes involving serious bodily injury, where these things were not included in the definition of the offense of conviction. Inchoate offenses were to be moved down one level of severity (except for those in the least serious category). The resulting offense severity scale thus had twelve levels.

The offender score was similarly judgemental, being made up initially before the data from Kramer's study were available. The scoring system involved assigning varying numbers of points for prior felonies and misdemeanors, up to a maximum of four; one extra point was to be added for prior convictions of offenses of seriousness levels six or seven, and two points for a prior conviction of an offense of seriousness levels of eight or higher. There was considerable debate among Commission members about the inclusion of social status variables (such as employment status); ultimately it was decided to exclude these, because of their probable correlation with race. As we have already seen, the same conclusion (based on the same argument) was reached by the Minnesota Commission -- though they had some data by which they purported to show that these factors played little part in previous sentencing practice. There was, however, considerable feeling among Pennsylvania judges that some social status factors were appropriately considered in sentencing;[55] thus, while these were not allowed to play an explicit part in calculating offense or offender scores, they were not explicitly excluded from consideration in applying the guidelines, as is the case in Minnesota.

The guidelines matrix which resulted (late in October 1980) from the Commission's deliberations is shown in Table 9.8. It will be seen that its 72 cells are divided into three groups by zig-zagging lines similar to the "in-out" line in the Minnesota matrix. (As presented by the Commission, Pennsylvania's matrix has the highest offense-score row at the top, and the lowest at the bottom; Minnesota's -- cf. Table 9.1 -- is the other way around. This makes no difference substantively, of course.) The lower of the two lines in Table 9.8 is the "in-out" line; the 13 cells below this line, which contain asterisks, prescribe presumptively non-incarcerative sentences, e.g. probation. The middle tier of 21 cells prescribes jail terms (which may be of up to 12 months); note that seven of these cells allow either a jail term within the stipulated range, or an "out" disposition. Above the upper line -- called, we hope colloquially, the "state/county" line -- the ranges stipulated are for state prison sentences.

Insert Table 9.8 here

The Pennsylvania guidelines, like those in Minnesota, provide for aggravation or mitigation of sentences on a number of stated grounds not dissimilar to those used in Minnesota (except that, as noted earlier, there is no explicit mention of race, social-status factors, or such things as pleading guilty). However, Pennsylvania's guidelines embody a feature not found in Minnesota: the amounts by which sentences may be increased or decreased on these aggravating or mitigating grounds are expressly limited. In the Minnesota guidelines, the aggravating and mitigating circumstances permit departures from the prescribed "normal" ranges; but nothing is said about how much of a deviation from those ranges is allowed. (Indeed, in none of the statewide guidelines thus far developed has the question of the amount of deviation in departure cases been addressed.) In Pennsylvania, by contrast, the amount of variation permitted is limited: in the case of aggravation, the judge may move to the adjacent (right-hand) cell from that in which the case would otherwise fall; if the case falls into the right-hand column, the judge may move one cell up. The rules for mitigation are exactly the reverse. The burden of proof in establishing aggravating circumstances is on the prosecution; mitigating circumstances may be brought to the court's attention by either side.

It is important to note, however, that this shift to an adjacent cell on the ground of a stated condition of aggravation or mitigation is not treated, under the Pennsylvania guidelines, as a "departure" from the guidelines.

Table 9.8: Original Pennsylvania Sentencing Guidelines

(Source: Pennsylvania Bulletin, Vol. 10, No. 43, Saturday, October 25, 1980)

OFFENSE SCORE	OFFENDER SCORE 0	1	2	3	4	5	6
12	78-84	84-90	90-96	96-102	102-108	108-114	114-120
11	60-66	66-72	72-78	78-84	84-90	90-96	96-102
10	48-54	54-60	60-66	66-72	72-78	78-84	84-90
9	36-42	42-48	48-54	54-60	60-66	66-72	72-78
8	24-30	30-36	36-42	42-48	48-54	54-60	60-66
7	8-11½	12-17	17-22	22-27	27-32	32-37	37-42
6	4-7	6-9	8-11½	12-17	17-22	22-27	27-32
5	0-3 / *	3-6	5-8	8-11½	12-15	15-18	18-21
4	*	0-3 / *	0-5 / *	5-8	8-11½	12-15	15-18
3	*	*	*	0-3 / *	2-5	5-8	8-11½
2	*	*	*	*	0-3 / *	2-5	5-8
1	*	*	*	*	*	0-3 / *	0-3 / *

* - Alternatives to Incarceration Including Probation or Probation with Special Conditions

NOTE: All Sentence Ranges are in Months

According to s.303.6 of the draft law in which the guidelines were presented to the legislature, any sentence other than one "suggested" by the guidelines -- including those suggested by the adjacent-cell rule for mitigation or aggravation within the guidelines -- is to be regarded as a "departure", which should only be imposed for "compelling" circumstances. The guidelines also contain rules regulating the use of consecutive and concurrent sentences, and providing for appeal of sentence by both defendant and prosecutor.

As was required by Pennsylvania law, the guidelines were published in the Pennsylvania Bulletin (the state's equivalent of the Congressional Record), and public hearings on them were held at various places around the state in December 1980, at which a total of 75 persons testified. The great majority of these witnesses, it seems, were critical of the guidelines; according to Martin (1981:86), the loudest and most visible opposition came from judges and prosecutors. The former objected to what they saw as a fettering of their discretion (though it would of course have been possible for them to go outside the guidelines, provided they could find something to describe as a "compelling" circumstance). The latter objected to the guidelines on the ground that these were not severe enough.

Accordingly, the Commission made a number of changes in the guidelines before submitting them to the legislature. The revised version is shown in Table 9.9. It will be seen that in this version most of the ranges and mid-ranges in the upper part of the matrix prescribe longer terms than did the original version; in addition, there are only five cells in the lower left-hand corner of the matrix in which the presumptive sentence is "out". In addition, some changes were made in the guidelines which are not obvious from inspection of the matrix. For example, the severity rank of rape was raised from 8 to 9; it is the only offense to have this seriousness ranking in the revised guidelines, which thus treats rape as more serious than kidnapping where the victim is harmed or released in an unsafe place, involuntary deviate sexual intercourse with a victim aged under 14, robbery with serious bodily injury, "causing a catastrophe", arson endangering persons, and a second or subsequent conviction for manufacturing or distributing e.g. heroin to a person under 18;[56] the only offense classified as more serious than rape under these guidelines, in fact, is third-degree murder.[57] Seriousness rankings for several kinds of assault were also increased, the prior record calculations were modified, and penalties were to be increased for any offense in which a firearm was used and discharged. Finally, the language describing the use of the guidelines was hardened, in at least two respects. First, it was no longer indicated that the

prescribed ranges, etc., were presumptive: for example, a statement in the initial guidelines that "The guidelines authorize limited variations from the normal sentence range for aggravating and mitigating circumstances...." was changed to read "The guidelines suggest limited variations...." Furthermore, the necessity to find a "compelling" circumstance to justify departure was changed to a requirement to find that the guidelines sentence would be "clearly unreasonable".

Insert Table 9.9 here

Thus revised, the guidelines were submitted to the Pennsylvania legislature in January 1981; they would take effect if the legislature did not positively reject them within 90 days. Martin (1981:87-88) has summarized what happened:

> The Commission had not built a constituency, had no visible supporters, and had designed no legislative strategy to speak of. It had counted on a disinterested public and the Commission's legislative members to prevent resolutions rejecting the guidelines from reaching the floor of both chambers. Instead, the supporters of mandatory minimum sentences re-emerged, submitting several bills for consideration with much fanfare.
>
> ...The coup de grace was delivered by Representative Lois Hagarty, a Republican ex-district attorney who, ironically, held Scirica's old House seat. She was upset at what she viewed as the guidelines' leniency, and was the first to announce plans to submit a resolution to reject the guidelines. Scirica convinced her to submit a resolution that called for the Commission to revise the guidelines to meet specific criticisms regarding their leniency and resubmit them in six months."

This resolution was subsequently adopted by enormous majorities in both houses of the Pennsylvania legislature, and the Commission was in effect sent back to try again. At the time of this writing, the outcome of their efforts at revision is not known. But in view of the pressures toward increased severity which prompted the rejection of their earlier efforts, and the Commission's general ineptness at selling this brand of sentencing reform, one cannot be confident either that they will produce reasonable guidelines, or that the legislature would accept such guidelines even if they were reasonable.

Table 9.9: Revised Pennsylvania Sentencing Guidelines

(Source: Pennsylvania Bulletin, Vol. 11, No. 4, Saturday, January 24, 1981)

OFFENSE SCORE	OFFENDER SCORE 0	1	2	3	4	5	6
12	78-90	84-96	90-102	102-114	108-120	114-120	120
11	60-72	66-78	72-84	84-96	90-102	96-108	102-120
10	48-60	54-66	60-72	72-84	78-90	90-102	90-114
9	36-48	42-54	48-60	54-66	60-72	66-78	72-90
8	24-36	30-42	36-48	42-54	48-60	54-66	60-78
7	8-11½	12-17	17-22	22-27	27-37	32-42	37-52
6	4-7	6-9	8-11½	12-17	17-22	22-32	27-37
5	0-*6	3-6	5-8	8-11½	12-15	15-18	18-24
4	0-*3	0-*6	0-*8	5-8	8-11½	12-15	15-18
3	*	0-*1	0-*1	0-*6	2-6	5-8	8-11½
2	*	*	0-*1	0-*1	0-*6	2-6	5-8
1	*	*	0-*1	0-*1	0-*3	0-*6	0-*6

All sentences ranges are in months, and are minimum sentences.

As we had only slight contact with the Pennsylvania project during the life of our research,[58] we can say little if anything about the apparent failure of the Commission to design and implement their guidelines. No doubt, as Martin's valuable account suggests, there were many reasons why the Pennsylvania Commission failed where its Minnesota counterpart succeeded. We have already described the excellent job done by the Minnesota Commission and its staff -- in particular, by Smaby, Parent and Knapp -- in presenting their guidelines to judges, prosecutors and public defenders, and dealing effectively with questions, points of criticism and suggested modifications; it is evident that the same skill was not displayed by the Pennsylvania Commission chairman (Judge Richard Conaboy) and his research staff.[59] As Martin (1981) also suggests, Pennsylvania's political culture probably hindered chances for reform, much as Minnesota's helped. The timing of Pennsylvania's efforts -- which coincided with a "get tough on crime" movement symbolically helped by the attempted assassination of President Reagan -- could not have been worse; even the relative dispersal of the Commission's members across the state, which made fully-attended meetings difficult, hampered its work.

In our opinion, however, the difference in the two Commissions' legislative mandates is a crucial factor explaining Minnesota's success and Pennsylvania's failure. We do not refer to the notion of considering past practice, which was explicitly mentioned in Minnesota but adopted (ineffectively) by the Pennsylvania Commission as well. The really important difference is that the Minnesota Commission was required to consider existing correctional resources (which it successfully interpreted as referring to bed-space in the state prison system), whereas the Pennsylvania Commission had no such direction. Surely, however, the only way in which the demands for increased severity of sentences in Pennsylvania could have been countered was by showing that such lengthy sentences would have entailed new prison and jail construction which nobody was willing to pay for. Some attempts were made by the Pennsylvania research staff to estimate the impact of the revised guidelines on institutional populations; but these were both belated and ineffectual, and served neither to influence the Commission nor to meet the arguments of legislative critics.

Given the uncertain status of the Pennsylvania guidelines, an analysis of their structure (as of January 1981) is necessarily something of an academic exercise. It is, however, worth undertaking a brief analysis of the Pennsylvania matrix, since it provides some interesting comparisons with the one adopted in Minnesota. To begin with, it may be noted that in the Minnesota matrix (Table 9.1) there

is no overlap between the ranges provided in adjacent cells prescribing prison sentences; both across the rows and down the columns, the ranges are separated by a few months, thus providing for clear distinctions between the different levels of offense seriousness and prior criminal history. This is not the case with the Pennsylvania matrix; Table 9.9 shows that most adjacent cells across its rows overlap substantially, and there is also some overlap between adjacent cells down the columns. Our copy of a Pennsylvania staff memorandum[60] states (p. 17) that "in fixing the ranges, the Commission tried to avoid overlapping ranges between adjacent cells. Such overlaps are undesirable as they permit less serious cases (in terms of the crime and criminal history) to receive longer terms than more serious ones." In fact, however, this was not quite true, even of the original version of the Pennsylvania guidelines. As Table 9.8 shows, adjacent cells across the rows of that matrix are bounded by the same numbers of months; for example, for Offense Score 8 the cell values are 24-30, 30-36, 36-42, 42-48, and so on. The term of 30 months may thus be given either to someone with an Offender Score of 0 or an Offender Score of 1; the same is true for the other boundary terms.

The same thing is true for several pairs of adjacent cells in the revised matrix (Table 9.9); for example, for Offender Score 1, the term of 54 months applies to both Offense Scores 9 and 10. Moreover, in most places in the revised matrix, there is substantial overlap across the column: in row 8, for example, a term of 54 months could be given, without invoking aggravation or mitigation, to offenders with Offender Scores of 3, 4, or 5. Such an arrangement is not necessarily wrong; it might be argued, for example, that there was so much variation in the seriousness of prior criminal histories that was not captured by the Offender Score (for example, variation in the seriousness or type of prior felonies) that such overlapping ranges were necessary. (The same sort of argument might be made for overlaps between adjacent cells in the columns.) That suggests, however, that the scales used to define the rows and columns of the matrix were themselves defined in ways that simply failed to capture those <u>legitimate</u> differences between offenses and offenders which the Commission presumably wanted to preserve. 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. (Note that this does not mean that there should be <u>no</u> overlap at all between two or more cells of the matrix. It might be reasonable to prescribe exactly the same range for a combination of a less serious offense and a bad prior record, and for a more serious offense coupled with a good prior record.) In our earlier discussion of the Minnesota guidelines

(which do clearly demark terms for differing offender scores), we noted that their criminal history score was quite complex in construction. Though we do not have supporting empirical evidence for this suggestion, it may be the case that the clear demarcation of terms could have been a function of the complexity of that score's construction.

Further, it can be seen from Table 9.9 that 18 cells bordering the "in-out" line toward the bottom of the Pennsylvania matrix contain ranges the lower limit of which is zero months, and the upper limit is three or six (or, in one case, eight) months in jail. As we noted in an earlier chapter,[61] a prescription of this kind imposes only minimal structure on judges' discretion. Offenders falling into these cells -- which, in the Pennsylvania matrix, include a variety of miscellaneous and rather trivial misdemeanors, coupled with an averagely bad prior record -- may be given either a jail sentence or a non-incarceration sentence such as probation, without there being a "departure" from the guidelines requiring special justification. Of course the jail terms allowed in such cases are not long ones; nonetheless, the choice between sentences that are "in" and those that are "out" is one that has been felt, in many jurisdictions, to be one of the most important choices with which sentencing judges are confronted, in carrying out the unpleasant task of sentencing convicted offenders. If guidelines are to play any role at all in guiding judicial decisions, surely this should be one of the choices which they help to inform and to regulate?

It was also noted earlier that the "in-out" decision, by virtue of being dichotomous, is one which for the concept of guidelines is intrinsically ill-suited -- especially if the guidelines purport to be empirically based.[62] It is not easy to see how a percentage incarcerated in the past can be turned into a sanctioning rule prescribing a proportion incarcerated in the future -- even if (as seems unlikely) some further factors can be found that separate the "ins" from the "outs" with some degree of precision. In the first (October 1980) version of the Pennsylvania guidelines, there were only seven cells which prescribed ranges of the form "0 - x"; the maximum term in those cells (found in the cell corresponding to Offense Score 4 and Offender Score 2) was five months in jail. The larger number of such cells in the revised (January 1981) guidelines no doubt reflects the views which the Commission encountered in its public hearings, which appear to have included the idea that disparity in sentencing was actually a good thing;[63] perhaps this change in the guidelines, before submission to the legislature, was also intended to cater for the fact (noted earlier) that suburban and rural judges had in the past sent much higher proportions

of convicted offenders to jail than had judges in Philadelphia and Pittsburgh. Whatever the explanation, the increase in the number of cells containing ranges of the form "0 - x" seems regrettable. Such a range may appear to provide an element of control over judges' discretion; in reality, however, it does little to promote consistency in sentencing.[64]

What of the structure of the Pennsylvania guidelines? Can the ranges (or mid-ranges) in the cells of this matrix be reproduced reasonably well by simply adding together a row (Offense) effect, and a column (Prior Record) effect? The answer is no. An additive analysis of the matrix -- similar to that done with the Minnesota matrix, as described in the preceding section -- shows clearly that such an additive model does not fit the observed values in the matrix; not only are there many large residuals, but they show the characteristic pattern of positive and negative signs shown in Table 9.2, indicative of a bad "fit" to the data. (The results of this analysis are not reproduced here.)

Unfortunately -- at least for those for whom tidy-mindedness is an important virtue -- a multiplicative model of row and column effects does not fit well either. We record this failure in Table 9.10, in which the row and column effects, and the residuals from the model, are expressed in months as in Table 9.7 above. (For ease of comparison, we have reversed the rows of the Pennsylvania matrix, so that more serious offenses are at the bottom.) It will be seen that, first, there are some very hefty residuals, especially in the lower corners of the matrix; and, second, that there is an evident pattern in those residuals, with those in the bottom left-hand and top right-hand corners being mostly positive, and those in the lower right-hand and top left-hand corners tending to be negative.

Insert Table 9.10 here

Confronted with this failure, it might seem natural to try to fit more complicated combinations of Offense and Offender effects (like the one set out for Minnesota above); surely if the Pennsylvania Commission were at all rational there must be some not-too-arcane combination of their Offense and Offender Scores that will account for the cell mid-ranges in Table 9.9? A moment's reflection, however, suggests that more complicated arithmetic would be a waste of time. The Pennsylvania guidelines are intended to apply to all offenses in the state's criminal code: they cover both felonies and misdemeanors, and thus prescribe terms for a much more heterogeneous group of offenses than do the Minnesota

Table 9.10: Multiplicative Model Fitted to Mid-Ranges of Pennsylvania Matrix
(Effects and Residuals in Months)

		Offender Score							Row Effects	
		0	1	2	3	4	5	6+		
Offense Score:	1	-.12	-.31	0	-.07	.74	.59	1.87	.566	.03
	2	-.12	-.31	0	-.07	2.24	2.09	5.37	.566	.03
	3	-1.17	-1.15	-2.17	-2.50	-.02	1.70	3.78	3.00	.16
	4	-1.03	-.57	-1.78	.01	1.05	3.12	3.58	6.49	.35
	5	-.81	.12	-2.20	-.03	.39	.85	1.54	9.78	.52
	6	-.13	1.81	-3.10	.06	.15	3.90	8.26	14.44	.77
	7	.28	1.14	-2.12	.21	-.55	1.86	3.84	24.29	1.30
	8	11.49	9.89	-.25	.53	-9.61	-15.95	-34.47	47.47	2.53
	9	18.70	15.14	.83	.26	-14.05	-23.58	-37.88	59.74	3.19
	10	24.97	19.06	-.25	3.56	-15.05	-23.10	-46.14	74.44	3.97
	11	31.59	23.47	-.52	1.77	-22.23	-39.17	-64.58	88.23	4.71
	12	41.97	30.73	.08	.23	-30.14	-55.43	-94.46	107.77	5.75
Column Effects:		.39	.55	.89	1.00	1.34	1.60	1.99	--	18.73

guidelines. For this reason, any model of their structure which is derived from the entire matrix is almost bound to fit badly, unless it takes this into account. Table 9.10 shows that there is a substantial jump in the row (Offense) effects, between rows 6 and 7; this is of course the point at which the guidelines begin to prescribe state prison rather than local jail terms. The large residuals in the lower rows of Table 9.10 reflect the fact that the Offense Seriousness effects in that table were averaged over the felony and misdemeanor crimes. One analytical strategy which might be used in such a case is to estimate row and column effects separately for different parts of the matrix, e.g. for the top six rows and the bottom six. We have not done this, however, in part because the "state-county" line slopes in such a way that some cells in even row 4 have prison terms as their normally prescribed ranges.[65]

Even without such more complicated analyses, some interesting facts about the Pennsylvania guidelines matrix can be seen. In particular, inspection of the mid-ranges of the cells of the matrix reveals that the effect of prior criminal record (the Offender Score) on prescribed ranges differs, for different levels of Offense Seriousness. Neglecting the bottom rows of the table, a comparison of mid-ranges in the left-hand and right-hand columns of the matrix shows that the Offender Score has progressively less effect on prescribed terms, the more serious the offense. Thus, for example, at Offense Score 5, the mid-range of the cell corresponding to an Offender Score of zero is three months; that for an Offender Score of six is 21 months, i.e. seven times as great. This ratio falls steadily, so that at Offense Score 9 it is about 1.93 to 1 (42 months for Offender Score zero, versus 81 months for Offender Score six). At the top of the matrix, for Offense Score 12, the ratio of highest to lowest mid-range is only about 1.43 (84 months versus 120 months).

This pattern of variation is precisely the reverse of that displayed by the Minnesota guidelines matrix, in which the effect of prior record was greater for the more serious offenses. That arrangement -- which delivers an extra heavy thumping to the "worst" cases in the matrix -- is likely to turn out to be of little consequence, given the relative rarity of cases falling into the cells at the lower right-hand corner of the Minnesota matrix. It may well have been adopted for reasons of political expediency (satisfying some real or imagined public hunger for "getting tough on crime"); we have already suggested that there seems no principle by which it can be justified.

The Pennsylvania arrangement seems even more difficult to justify. To see its effects, consider the following two

cases. A burglar with no prior record might (according to the guidelines) expect to receive about three months in jail, if he were not armed and were fortunate enough not to have a confrontation with anyone on the premises (18 Pa.C.S., s.3502; Commission seriousness rank 5). Given three prior felonies and three prior misdemeanors, however, he could expect a prison sentence of about 21 months; that is, about seven times as much as a first offender might get, and in a different type of institution. A rapist, by contrast, might expect a prison term of about 42 months, if he were a first offender (and if there were no mitigating or aggravating factors); given the worst possible record, he would receive a minimum prison sentence of about 81 months, which is less than twice the term given to the first offender. It might be argued that to say that prior record makes more of a difference with the less serious offenses is really the same as saying that prior record makes less of a difference with the more serious offenses; and that that is really what the guidelines should provide. But why should this be so? It is true that the top row of the Pennsylvania guidelines is constrained at its right-hand end by a statutory limitation on minimum terms; thus, given that a range of 78-90 months is thought to be the appropriate punishment for such offenses even if committed by someone with no prior record, that penalty can only be increased by about 50 percent by the worst possible prior record. But it does not follow from that fact, that the ranges in lower rows of the matrix should provide proportionately much greater terms; nor was this arrangement entailed by the Commission's legislative mandate, which merely directed it to "specify a range of sentences of increased severity for defendants previously convicted of a felony or felonies...."[66] Moreover, the multiplicative arrangement of the Pennsylvania matrix is likely to have considerable practical consequences, since there are probably many offenders in the lower rows of the matrix with substantial prior records. We have no data on sentencing in Pennsylvania, comparable to the data for Minnesota presented in Tables 9.2 and 9.3 above; but it seems likely that the impact of the Pennsylvania guidelines on institutional populations (as projected, somewhat crudely, by the Commission's staff[67]) would be more a consequence of the augmentation of terms for property offenders with prior records, than of the length of terms for violent crimes.

A further fact about the Pennsylvania guidelines matrix is that it contrains discretion much more narrowly for the more serious offenses and offenders than for the less serious. In general, the ranges in the Pennsylvania matrix are commendably "tight", if they are thought of in plus-or-minus percentage terms around the mid-ranges. They are not quite so narrow as those in the Minnesota matrix; but they are

certainly much narrower than the permissible variation (plus or minus fifty percent) which, as we saw in earlier chapters, is allowed as a sentence "within the guidelines" in Massachusetts. Inspection of the Pennsylvania matrix reveals, however, that there is much more "running room" (relative to the middle of the cell range) in the lower left-hand corner of the matrix, than in the upper right-hand corner. There is thus much more "normal" variation permitted for the less serious offenses and offenders, than for the most serious ones. Thus, for example, in the cell at the intersection of row 12 and column 5 of the Pennsylvania matrix, the prescribed range is from 114 to 120 months in the state prison; that is a range of plus or minus three months around a mid-range of 117 months, or a little under three percent. One may wonder what point there is in allowing sentences to vary by three percent in either direction -- especially since judges, left to themselves, have seldom tried to make such fine-grained distinctions even without guidelines. Why not just say "117 months" and be done with it?[68]

By contrast, in the lower left-hand corner of the Pennsylvania matrix -- ignoring those cells in which the presumptive sentence is a non-incarcerational one, and those cells (discussed above) with prescribed ranges of the form "0 - X" months, the ranges provided are much wider relative to their mid-ranges. For example, the cell at the intersection of row 3 and column 4 of the matrix prescribes a range of between two and six months in jail. That is a range of five months, around a mid-range of four months, or permissible variation of 63 percent around that mid-range -- which is more, relatively, than is allowed in Massachusetts! Of course, in that quadrant of the Pennsylvania matrix, the absolute numbers of months involved are not great (although the absolute numbers of offenders may be -- and months do add up for a large caseload). Thus if there were two offenders convicted of driving under the influence of alcohol, each of whom had (say) four prior convictions for non-serious felonies, one might be sent to jail for two months, the other for six months, within the guidelines. It is possible, of course, that neither would see this as being very much of a difference; yet one would have received three times the jail term of the other.[69] Or again, consider the cell at the intersection of row 10 and column 0 of the Pennsylvania matrix, in which the range prescribed is from 48 to 60 months. That is a range of 13 months, around a mid-range of 54 months, or approximately plus-or-minus 12 percent. An offender who had committed a crime of the same seriousness level, but who had a prior record giving him an Offender Score of six, would confront a prescribed "normal" range of from 90 months to 114 months: that is a range equivalent to a mid-range of 102 months, plus or minus (approximately) only seven percent.

What reason of principle can account for that sort of nonsystematic variation, within the structure of the revised Pennsylvania guidelines? We strongly suspect that there is none.

Because of variations in legal definitions of offenses, and probable variations between states in such matters as charging, plea bargaining, and the like, it is difficult to make direct comparisons between the sentences prescribed by the Pennsylvania sentencing guidelines and those in Minnesota. Martin (1981) has made some efforts at a few broad comparisons of this kind, with appropriate reservations; her main conclusions, while cautiously stated, seem to us to be generally correct. The prison terms prescribed by the Minnesota guidelines look much more severe in many cases, especially for the more serious offenses; but in fact they are generally less severe, first because they will be reduced in most cases by one-third, for "good time" and second, because the terms prescribed by Pennsylvania's matrix are minimum terms, not maximum terms.

In summary, the Pennsylvania sentencing guidelines in their original (October 1980) form had a number of commendable features -- several of which were removed, or diluted, by the January 1981 revision. The ranges originally provided were commendably narrow in comparison with other guidelines except Minnesota's; and there was only minimal overlap between the ranges in adjacent cells. The form of the guidelines, with the "state-county" line, would probably have reduced some variation in place of confinement permitted for offenses of intermediate gravity (those with maximum terms of between two and five years), for which judges can sentence either to jail or to state prison. The list of aggravating and mitigating factors was clearly and precisely drawn; and there were limits placed on the amount of variation which could be justified by those factors. In the revised guidelines, prescribed ranges were often made wider, and overlap between adjacent categories increased; there was also an increase in the number of cells making use of a minimum of zero (i.e. ranges of the form "0 - X"), which effectively provides very little constraint on judicial discretion. Moreover, the structure of both the initial and the revised guidelines allowed prior criminal record to have a much heavier impact on prescribed terms, for those convicted of the less serious offenses; we have already noted that this was in no sense necessary, and we can see no particularly good reason for it.

The Pennsylvania sentencing law, with its provision for both maximum and minimum terms, seems to have been at the root of some of the reasons for the guidelines' failure in the legislative arena; this failure was probably also made more

likely by the extremely wide variation in the use of incarceration across the state. Finally, the absence of any express requirement to consider institutional populations probably left the Commission completely unable to withstand legislative demands for more severe sentences that would have shown the "toughness" that legislators wanted (but might have been unwilling to pay for). But even this requirement might not have secured the guidelines' acceptance, unless the Commission had been more adept politically than it seems to have been.

Sentencing Research and Guidelines in Michigan

Research on sentencing practice, with the eventual objective of producing sentencing guidelines, was begun in the state of Michigan in April 1978. In contrast to Minnesota and Pennsylvania, this did not come about as a result of a legislative mandate; instead, the research was supported by a grant awarded (by the Michigan Office of Criminal Justice) to the State Court Administrative Office. In this respect the Michigan guidelines project was closest in concept to the original ideas of Gottfredson and Wilkins discussed in Chapter 3 above, and to the New Jersey effort described in Chapter 4: its animus was judicial rather than legislative, and the resulting guidelines were meant to be primarily descriptive or empirically-based.[70] The project had a 19-person Steering and Policy Committee which included eight judges as well as three legislators, the chairman of the parole board, director of the department of corrections, four representatives of the legal profession, and the State Court Administrator, Mr. Einar Bohlin. The director of the project was Professor Marvin Zalman, who had priviously been a member of the law faculty at Michigan State University.

Michigan had (and still has) an "indeterminate" sentencing system of a fairly conventional kind. In felony cases, judges have discretion either to grant probation, or to impose jail or a minimum state prison sentence (which last by law may not exceed two-thirds of the statutory maximum for the offense[71]). Release from prison, at any time between the judicially-set minimum (less good time) and the maximum is at the discretion of the parole board. By 1977 there had been three proposals to replace this system with a more "determinate" one. A "flat" sentencing scheme, with mandatory minimum terms and no probation or parole, had been proposed by the Wayne County prosecutor; the Criminal Code Revision Committee of the State Bar had recommended presumptive sentences, and the Department of Corrections had proposed a modification of the existing system using parole guidelines. In addition, in 1977 a two-year mandatory minimum sentence for possession of a firearm during the commission of a felony

became law; early in 1978, harsh mandatory minimum sentences for drug offenses were re-introduced (they had been eliminated in 1971); and in November 1978 a law passed by referendum in effect abolished "good time" for prison sentences for over 80 offenses (Zalman et al., 1979:15-18).

It was in the light of this background -- the local manifestation of the swing away from indeterminacy, and the beginnings of a demand for greater severity -- that the Michigan project proposed to create sentencing guidelines. These were said to provide "the best alternative to the current sentencing morass"; it was said that a guidelines system "can provide an understanding of sentencing practice, unambiguous guidance to judges, flexibility in decisions, and a method to continuously monitor the sentencing process. This system provides a level of rational policy input, oversight, and accountability that is not available in other sentencing alternatives" (Zalman et al., 1979:17). As we have seen in earlier chapters, such hyperbole is not at all unusual among the proponents of guidelines.

The Michigan researchers set themselves initially to answer two research questions:" (1) Are there patterns (i.e. policies) among judicial sentencing decisions in Michigan? And if so, describe them, and (2) Is there sentence disparity in Michigan?" (Zalman et al., 1979:56). In passing, we may note again that the equation of "patterns" and "policies" in the first of these questions is mistaken; and that the definition of "disparity" here and elsewhere in the report leaves something to be desired.[72] However that may be, the Michigan team's pursuit of these objectives was carried out in a serious and careful fashion. From the Criminal Case Conviction Register maintained by the Michigan Department of Corrections, a 25 percent sample of all persons convicted of felonies and sentenced in the calendar year 1977 was selected; this sample contained a total of almost 6,000 cases. A disproportionate stratified random sample design was used, with dual stratification for geographic region and offense seriousness. This yielded three geographic strata (metropolitan, urban and rural groups of counties), and five categories of offense severity, or a total of 15 strata from which cases were selected at random. The design effect resulting from this strategy is not directly reported, but from unpublished appendices it appears to be reasonably small.[73] The sampled cases were weighted to reflect the total population of 26,116 sentences from which they were selected, and most of the findings in the report are based on the weighted data.

Data on sentenced offenders were collected from pre-sentence reports. As we have noted earlier, this

procedure is not ideal; but it is virtually the only one available when sentencing research is done retrospectively.[74] It is stated that "in virtually all instances" complete reports were available; the few specimens which we saw, during our site visit to Michigan in 1979, were indeed remarkably detailed. The data collection forms used, admittedly over-inclusive of variables, contained a total of 421 items concerning the offense and offender, plus 12 items for each prior recorded offense. Over 30 persons were employed to code the data, with ten percent of the cases being re-coded as a reliability check; the overall inter-coder agreement was just over 88 percent, being much higher on most of the important variables. Though we have not yet analyzed these data ourselves, the codebook and its related documentation, and the data collection procedures as described, seem to us to reflect a very high level of social science research practice -- not just by the standards of research on sentencing, but in general.[75] The resulting data set should thus be an extremely valuable resource for secondary analyses.

The statistical analyses reported also reflect, in our opinion, a high level of technical competence. Though we shall offer some conceptual criticisms of those analyses, it is important to emphasize that they represented the state of the art (among guidelines developers) at the time they were done; several of the problems which we shall note were not fully appreciated at that time, and some are even now not fully resolved.

As a first step, the offenses of "primary convictions" in the sample (i.e. presumably the most serious offense of conviction, in multiple-count cases[76]) were grouped into 12 categories, containing the follwing numbers of cases:

Category	Sample N	Population N
Homicide	301	951
Assaults	338	1,701
Sex Crimes	444	1,169
Robbery	694	2,218
Drugs	758	4,266
Burglary	1,684	6,391
Larceny	608	3,787
Fraud	533	2,015
Property Destruction	119	540
Weapons	240	2,066
Escape	45	216
Miscellaneous/Unspecified	145	796
Totals	5,909	26,116

Inspection of the offense titles in these categories suggests that they are mostly rather homogeneous; certainly this is a much more homogeneous grouping of crime types than the "general" or "generic" formats suggested by Gelman, Kress and Calpin (1979), which were specifically rejected by Zalman et al. as too broad. Nonetheless, inspection of the statutory maxima in each category shows that there is still quite a lot of within-category heterogeneity in some cases; for example, the category of Sex Crimes contains no less than 26 distinct crimes (according to Michigan law), ranging from rape and first-degree criminal sexual conduct, each of which is punishable by life imprisonment, down to contributing to the neglect of a child and soliciting for the purpose of prostitution, each of which is apparently punishable only as a misdemeanor.[77] This illustrates a quite general problem of research on sentencing: even with a large sample -- and Michigan's is to our knowledge the largest sample so far used in this kind of research, since in New Jersey the whole population of sentenced cases was used -- it can be very difficult to get enough cases of particular types to assess adequately all of the factors which may legitimately influence sentencing practice. For example, the Michigan sample included 82 cases of criminal sexual conduct in the first degree, and 193 cases of criminal sexual conduct in the second or third degrees (each of which carries a statutory maximum of 15 years in prison). But second and subsequent offenses of these types carry mandatory minimum terms of five years; presumably the sentences actually imposed in such cases would be even heavier. Any global assessment of the perceived seriousness of such crimes might be distorted by the presence of repeated offenses; yet the numbers of such cases would probably be far too small for separate analysis. This is one reason why careful inspection of the residuals from sentencing models is important.[78]

After quoting the now-familiar dogma that "there are two basic sentencing decisions: whether to incarcerate, and if so, how long" (Zalman et al. 1979:72), attempts were made to forecast the dichtomous "in-out" decision using all cases in the sample, and to estimate lengths of prison or jail terms for those incarcerated. Some attention was also given to variations across the three regional strata (metropolitan, urban and rural).[79]

The data-analytic strategies used can perhaps best be described as Sophisticated Kress-Gelman. First, variables which showed insufficient variation were excluded; then, a group of judges were asked to identify which of the remaining variables were potentially relevant to sentencing decisions. These remaining variables were grouped in offense, offender and "other" categories, with it being assumed that each of

these could affect sentencing outcomes while being themselves unrelated (see the diagram in Zalman et al. 1979:88). Though it was thought that the "other" variables had an impact on sentences, these were excluded in the first analysis on the ground that they were not explicitly considered by judges. The first model fitted, therefore, took the form

$$S = a_i A_i + D_j B_j + e,$$

where S is the expected sentence, the ai are weights associated with the Ai offense factors, the bj are similar weights associated with the Bj offender variables, and e is an error term.

It appears that no models other than this linear additive one were considered (Zalman et al., 1979:89). Though the various categories of offense were analyzed separately, an Offense Seriousness variable was for some reason added; this took the value of the statutory maximum penalty for the (primary) offense with which the offender had been charged.[80] The parameters of the model were estimated using ordinary least-squares (OLS) regression; while it was noted by Zalman et al. that other techniques, such as logit or probit analyses, are theoretically more efficient where the dependent variable is dichotomous, it was found that these did not, in practice, produce much different results.[81]

Because of computer core limitations, it was necessary first to regress outcome variables on the fairly large number of offense variables, selecting those that were statistically significant for further analysis; and to repeat that operation for the offender variables as a group. The surviving variables were then combined into a single equation, and the final model included all of the variables which attained a high level of statistical significance (p .01) in that analysis. It appears that no interaction terms were ever fitted; and some possible interactions were of course necessarily excluded from consideration (i.e. those involving variables excluded by virtue of the first, separate, offense and offender regressions). Finally, the weighted offense variables and weighted offender variables were separately summed, and sentence outcomes regressed on them; the standardized coefficients of these composites were taken to reflect the relative contributions of the groups of offense and offender variables to the sentencing outcomes.

The outcome of this process is illustrated in Table 9.11, which is based on Table 3.5 of Zalman et al. (1979:95). This table shows that for the category of Sex Crimes, three offense variables and 10 offender variables were included in the final model. Though all of the coefficients in this table are

evidently statistically significant, some of them equally evidently make no sense; given the direction of coding, all should be positive, whereas in fact three (bodily beatings, type of work, and violent felonies as a juvenile) are negative. The total variance accounted for is R square = .31, though as the authors note this is not a very useful statistic where the outcome variable is dichotomous. They thus calculated a "predicted score" equal to the sum of their composite (standardized) regression coefficients; somewhat arbitrarily, they classified cases as predicted "in" if this sum exceeded 50 and "out" otherwise.[82] The results of this exercise, classified against observed "in-out" decisions, are shown in Table 9.12. It will be seen that while the model correctly predicted all of the 173 cases actually given "in" sentences, it was wrong twice as often as it was right for the predicted "out" cases, which is not a very encouraging result.[83] Finally, from the standardized weights for their offense and offender composites, they suggested that offender variables were more than twice as important as offender characteristic, for the "in-out" decision for Sex Crimes. For reasons which we shall try to make clear in a moment, we believe that this conclusion is unwarranted and probably wrong.

Insert Tables 9.11 and 9.12 here

Similar though sometimes a little more favorable results were found in parallel analyses of the "in-out" decision for the other offense categories. Values of R square ranged from .55 for property destruction to .25 for burglary; different though somewhat overlapping combinations of offense and offender variables found their way into these equations, and while the coefficients of those finally included were all statistically significant, they were sometimes counter-intuitive. In the prediction of "in-out" sentences for other offense categories, percentages correct were between 60 and 80 percent for most major categories -- the exception, apart from Sex Crimes, being Robberies, where only 34 percent of observed decisions were predicted by the model. In most cases the errors of prediction involved cases predicted to be "out", but actually incarcerated.[84] Finally, the procedure of comparing standardized weights for summed variables suggested that offender characteristics were more important for all categories of offense except homicide.

The procedures just described were then repeated with length of sentence as the dependent variable, using only those cases that had been incarcerated. The results obtained for Sex Crimes are shown in Table 9.13; it will be seen that the

Table 9.11: Statistically Significant Variables in the "In-Out" Regression Equation for Sex Crimes in Michigan Sentencing Research (Based on Zalman et al. 1979:95)

	b	beta	F
Offense variables:			
Seriousness (stat. max.)	.0009	.186	49.1
Extent of mental trauma	.1390	.086	12.0
Bodily beatings	-.0720	-.080	9.8
Offender variables:			
No. of incarcerations	.093	.198	56.5
Relation to CJ system	.093	.189	55.4
"Good moves" since arrest	.204	.218	69.8
Type of work	-.085	-.130	25.7
Reason for leaving school	.108	.108	18.1
Drug use status	.093	.079	9.6
Alcohol use	.045	.084	10.7
No. juv. viol. felonies	-.318	-.087	12.2
Residential stability	.042	.077	8.8
Detainers outstanding	.133	.071	7.7

Adjusted R^2 = .31

Table 9.12: Cross-Classification of Predicted Versus Actual Sentences for Sex Crimes in Michigan Sentencing Research (Based on Data in Zalman et al. 1979:97)

Predicted Sentence	Actual Sentence		Total	%
	"In"	"Out"		
"In"	173	0	173	(86)
%	(100)	(0)	(100)	
"Out"	627	369	996	(14)
%	(63)	(37)	(100)	
Totals	800	369	1169	(100)
%	(69)	(31)	(100)	

seven offense variables and ten offender variables in this table together account for 65 percent of the variance in observed terms. Again, however, while the F-values are all substantial, the directions of five of the included coefficients are counter-intuitive; why, for example, should bodily beating and number of offenders lead to lower sentences? (That sex offenses with younger victims should in general receive longer terms would perhaps not be surprising; however, that is not what that coefficient means.[85]) It is worth noting that only two of the offense variables in the regressions for lengths of sentence, and only three of the offender variables, were also included in the "in-out" regressions -- though this does not, we think, lend much support to the view that two entirely separate decisions are made by judges in such cases -- nor, of course, does it show that the factors included in these equations are the ones actually considered by the judges who dealt with the cases in this sample.

Insert Table 9.13 here

Overall goodness-of-fit was fairly good -- R square = .50 or better -- for most categories of offenses; however, it was comparatively poor for robbery (R square = .34), for which, of the 11 offense variables and 10 offender variables attaining statistical significance, five were signed in the opposite direction from what one would commonsensically expect. (This was also the case in most of the other regressions.)

The next step in the Michigan researchers' analysis was the construction of what they called "descriptive sentencing matrices", in which they cross-classified cases in each offense category according to combinations of the offense and offender variables uncovered at the first stage of their analysis. However, before turning to those "sentencing matrices" -- which were intended to serve as the basis for subsequently-developed guidelines -- we note some problems with the regression procedures we have described.

(1) To begin with, the Michigan researchers' procedures were not based on any well-articulated model of sentencing policy -- their crude categorization of variables on which they had collected data into "offense", "offender" and "other" types being scarcely worthy of that name, especially since they did not consider possible interrelationships between candidate variables in those categories (e.g. the possibility that those committing more serious current offenses would have less lengthy prior records). In the absence of any such model or theory about how judges actually do decide what sentences

Table 9.13: Statistically Significant Variables in the "Length" Regression for Sex Crimes in Michigan Sentencing Research (Based on Zalman et al. 1979:111)

	b	beta	F
Offense variables:			
Type of weapon	27.44	.300	123.1
Seriousness (stat. max.)	.33	.378	214.1
Bodily beatings	-16.42	-.099	18.4
Age of primary victim	-23.19	-.124	26.3
Long-standing feud	-51.42	-.073	10.3
Victim-offr. relationship	8.09	.068	8.7
No. of offenders	-8.20	-.058	6.4
Offender variables:			
No. adult felony convictions	3.62	.071	6.4
Pending charges elsewhere	19.54	.080	12.3
Disposition, last parole	9.98	.103	14.6
"Good moves" since arrest	20.34	.108	21.6
Mental health	18.33	.097	18.8
Detainers outstanding	37.57	.117	27.7
Disposition, last probation	-6.01	-.080	12.1
Has job to go to	13.50	.105	19.6
Type of work	-8.83	-.069	9.1
Sum months minimum terms	.06	.069	7.9

Adjusted R^2 = .65

they impose, they were necessarily left to snuffle around, in a crudely "empirical" fashion, among those items of information (in themselves of unknown reliabilities[86] or validities) which happened to be contained in pre-sentence reports. As Zalman et al. note in several places in this section of their report, this essentially atheoretical procedure may well be responsible for some of the counter-intuitive findings which we have already mentioned, e.g. negative coefficients where common sense would have led one to expect positive ones. For example, Sex Crimes cases in which there was some degree of "beating" might have been given very light sentences on some other ground, not represented by a variable included in their final equation. This problem is, if anything, likely to be compounded by some of the coding choices which Zalman et al. made, e.g. treating ordinal-level variables as if they were measured at the interval level. (For example, the "bodily beating" variable just mentioned was coded zero if there were no such beatings, 1 if there was one (whatever that might mean), and 2 if there were more than one; extreme cases of violence would thus in effect have been underscored.)

(2) In their initial attempts to model previous sentencing practice, Zalman et al. deliberately excluded the whole category of "other" variables, including not only such clearly invidious things as race but such arguable ones as social status and sex of offender. According to their report, they did not do this because the judge who they consulted said that these things were totaly irrelevant; but that does not matter. The point, which we discuss further below in dealing with Zalman et al.'s analysis of racial and other effects, is that the exclusion of such factors from the modelling equation (thus consigning them, for practical purposes, to the error term) may affect the coefficients associated with other variables -- e.g. those relating to offenses or offenders -- that are included in the equation.

(3) Perusal of the variables which ultimately found themselves included in the "in-out" and "length" regressions developed by Zalman et al. shows that there are a number which are probably highly correlated with each other. For example, in the "in-out" regression for assaults, "residential stability" has a positive coefficient; this variable is coded +1 if the offender's residence was "unstable", and -1 if it was "stable". Also included is a variable called "supports spouse/offspring", which has a negative coefficient and is coded +1 if the answer is "No", and -1 if the answer is "Yes". Assuming that those with "unstable" living arrangements do not in general support their spouses, these should be positively correlated with each other -- perhaps to a substantial degree. If so, they should not both be included in the same equation,

as they are merely splitting the available variance in the dependent variable between them. There are a number of other examples, mostly to do with prior-record variables, in which two or more variables which are probably correlated (either de facto, or because they are really ways of measuring more or less the same thing) are in the same equations. In situations of this type, it is likely that at least one of these variables is in effect redundant; it could well be eliminated from the equation (even if, as it happened, its coefficient were statistically significant). (See, for a discussion, Mosteller and Tukey, 1977, Chapter 13.)

(4) Even more important than this, it appears that the Michigan researchers seriously fitted only one model -- namely, the linear and additive one given above. They did not try to fit any interaction terms, so far as their published report reveals; nor did they experiment with reasonable transformations of any of their candidate explanatory variables; nor, despite considerable evidence of skewedness in lengths of term for many of their offense categories (see Zalman et al., 1979:75-84) did they apparently consider logarithm or other transformations of their length-of-sentence measure.[87]

(5) No analysis of residuals from the various models fitted to their data was carried out by Zalman et al.; or at least none is presented. Yet -- especially in view of the counter-intuitive results which they obtained in some cases, and the comparatively poor performance of their models in estimating probabilities of incarceration and lengths of term for some offenses -- they should surely have looked rather carefully at the ways in which those models failed to fit their data, to see whether this was due to a general failure of the model itself, or to problems of measurement, or to the existence in the data of a few wildly abnormal cases (cf. our discussion of the Massachusetts data in the preceding chapter). Such a residual analysis might, at a minimum, have helped to explain the curious results, (e.g. coefficients with odd signs) which they found from their regressions.

(6) Perhaps most important of all, it seems from their published report and other materials which we obtained on their project, that Zalman and his colleagues made no effort whatever to validate their findings, in the statistical sense of seeing what proportion of the findings from their first analysis might have been due to chance variation. It is true that, even given the numbers of cases in their sample, they might have encountered some problems if they had tried to split the data into two (or more) random sub-samples, so as to test the stability of the regression weights they obtained. But something could have been done here; and -- given the

unusual coefficients which they obtained in some cases -- it would have increased confidence in their results, if such a statistical validating procedure had been followed.[88]

(7) Finally, we are in some doubt as to the correctness of the procedures followed by Zalman et al. on the way to their conclusion that for most offense categories, their bundle of "offender effects" was collectively more important (in the sense of contributing more to predictions of sentence outcomes) than their bundle of "offense effects." To begin with, the data supporting this claim come from analyses of proportions incarcerated, or lengths of term, within the several offense categories which Zalman et al. created. Inspection of the data presented in their report, however, shows that there is considerable variation between these offense categories, both in percentages of "in" sentences and in lengths of term for those incarcerated. Thus, for example, from Figure 3.1 in Zalman et al. (1979:73) it can be calculated that for the 10 major categories of offenses they created (excluding "Escape" offenses and the "Miscellaneous/unspecified" category), about 52 percent overall were incarcerated. But the use of incarceration (either prison, jail or a "split sentence" of both probation and jail) varies considerably, from a high of about 80 percent for the 2,215 persons convicted of robbery and 78 percent for the 951 persons convicted of homicide, down to 41 percent of the 2,015 fraud offenders, and 28 percent of those convicted of weapons offenses. Similarly, the mean length of sentence for those imprisoned is just under 36 months, across all of the 13,095 convicted offenders in the ten categories just mentioned; but mean sentences vary between these ten groups, from a high of 162 months for homicides, 70 months for robberies, and 64 months for sex crimes, down to a low of just under 13 months for weapons offenses and larcenies. Thus, taking all felony sentences as a group, what are manifestly "offense" variables -- built into definitions of the different crime categories involved -- account for a substantial amount of variation in sentencing, which is by definition excluded from the within-category analyses on which Zalman et al.'s statements are based.

Furthermore, the data presented in the Zalman et al. (1979) report make it clear that lengths of terms are multi-modal for most offense categories, and extremely skewed, typically (with the exception of the homicide category) to the right; though we cannot be sure of this, it seems more likely that this is a consequence of "offense" factors, rather than "offender" ones. For example: lengths of term within the Robbery category created by Zalman et al. are (by their Figure 3.4, p.77) drastically multi-modal, with extreme peaks at 6, 12, 24, 36, 48, 60, 72, 98, 180, 240 and 300 months.

According to their Table 3.1 (p. 64 of their report), this group of cases represents a population of 2,218 cases over half of which were convicted of crimes with a statutory maximum of like imprisonment (coded by Zalman et al. as 300 months); other "peaks" in the visibly skewed distribution of maxima are at 15, 10 and 5 years. In the equation by which they predict length of sentence, the "offense seriousness" variable for robberies (actually, as we have noted, the statutory maxima in months) has the lowest unstandardized regression coefficient associated with them (.176, compared with e.g. 40.11 for a "yes" answer to "was there injury to eye(s)", which is a dichomotous dummy variable). But the standardized partial regression coefficient (beta coefficient) associated with offense severity is .200, compared with .081 for the eye-injury item just mentioned. This suggests that the statutory maximum term -- Zalman et al.'s measure of "offense severity" within categories -- is in fact having much more effect than its unstandardized partial regression coefficient would suggest. (Indeed, in the case of robbery offenses, the beta coefficient for offense severity is greater than any of the other coefficients included in their final model -- whether these are factors associated with offense or offender.) In short, we suspect that the rather crude measure of "offense severity" used by Zalman and his colleagues may in fact have had much more of an effect, even within offense categories, than their reported results suggest; we are at present unsure what effects the truncation (e.g. through dichotomization) and restriction of ranges and possible values of their other variables may have on their results.

In any case, we are uncertain as to the algebra by which Zalman et al. reach their general conclusion (relating both to the "in-out" decision and lengths of term) that "offender" variables are more important to within-category variation than those relating to the offense itself. Quite apart from uncertainty as to the durability of the variables in either category (because no validation on a separate sample from the same population was performed), the problem is as follows. The "Offense" and "Offender" variables which Zalman et al. use are in the first instance computed by summing the products of the variables in each category and their unstandardized regression weights; e.g., the "Offense" variable is defined as

$$\text{OFFENSE} = \sum_{j} a_{ij} A_{ij}$$

where the summation is over all of the i variables in their first-stage regression equation (Zalman et al., 1979:93).

Now, if we consider a typical element within such a summation, say the "offense severity" variable for sex crimes,

this element equals ajAij for the ith case; if that case is a case of rape, for which the statutory maximum is life (= 300 months on the coding scheme used by Zalman et al.), this element has a value of (.0009)(300) = .27, or (in this case) a 27 percent probability of being incarcerated, from this variable alone, controlling for all others in the equation. Now, part of the increment to the probability of incarceration in this case is due to the (unstandardized) regression weight of .0009; but evidently it is only a very small part. Part, however, is due to the magnitude of the variable itself, i.e. the statutory maximum, which in this case equals 300 (months). If the sum which includes this composite is again standardized -- as is in effect the case, if beta coefficients for the Offense and Offender composites are taken as indicators of their relative predictive power where sentencing decisions are concerned -- the original offense seriousness variable (i.e the statutory maximum, in months) is further diluted by being expressed in standard-deviation units. The same thing is of course true for the raw variables which go to make up the "Offender" composite. These variables, however, have nowhere near the range that the statutory maxima have; many, if not most, are in fact dichotomies. It would seem, therefore, that a better measure of the relative magnitudes of those variables themselves, i.e., Σ aiAij and Σ bjBkj, and not the beta weights resulting from a further regression of sentences on those variables.

Some data presented elsewhere in the Zalman et al. report suggest that the composite Offense and Offender variables -- that is, the sums of the variables in each category, weighted by their unstandardized coefficients -- do in fact give precisely the opposite result from that suggested by the authors: that is, they are consistent with Offense effects being more important than Offender effects. On p. 128 of the report, means for the Offense and Offender composites for the category of Sex Crimes are given, for a different purpose.[89] The mean for the Offense variable is 43.4; the mean for the Offender variable is about 18.2. If our arguments are correct, these figures show that aggregate weighted Offense effects are more than twice as influential in the prediction of sentences (in this case, lengths of term) than Offender effects, even where within-category variation is concerned.

In summary, it seems to us that the Michigan researchers' analyses of past sentencing practice have some shortcomings. To begin with, it must be re-emphasized that they could at best have modelled sentencing practice, not sentencing policies; that data with which they were forced to work could say nothing whatever about the sanctioning rules which Michigan judges took themselves to be following, nor could they throw any light on the judges' interpretations of the

things recorded -- correctly or not -- in the presentence reports with which Zalman et al. worked. Moreover, Zalman and his colleagues based all of the analyses which we have so far discussed on a linear and additive model of "Offense" and "Offender" effects -- this despite the view (expressed on p. 89 of their report) that "other" variables did in fact have an impact on sentences. They did this because they believed that offense and offender variables were "explicitly employed" by judges; this may (or it may not) be true, but it does not license the exclusion of "other" variables from a statistical model of sentencing practice. After testing their model[90] on their data, they concluded that it did not fit; from this finding -- whether or not it is correct -- they concluded, at several places in their report,[91] that there was considerable "disparity" or unexplained variation in sentencing in Michigan in 1977. But nothing at all in their report justifies this sweeping conclusion -- however welcome the conclusion itself may be to those wishing to implement sentencing guidelines. That the additive model of the Offense and Offender variables tested by Zalman et al. gave a poor fit to their data may show nothing at all, except that they had an incorrect model; since it appears that they did not try any other models, that conclusion seems at least as plausible as their conclusion that sentencing in Michigan was unduly variable. (In fact, as we shall see shortly, some other analyses later presented by Zalman et al. show quite clearly that their additive "Offense" and "Offender" model was wrong; these analyses themselves seem to us to have some serious flaws, but if they are at all correct they should have led Zalman and his colleagues to reject their model, instead of using it as a basis for criticizing the sentencing practice of Michigan's judiciary.)

It is not, of course, unreasonable to suppose that some of the variables providing an adequate description of sentencing practice should relate (in some sense) to the offenses for which sentences are imposed, and that some of them should relate (in some sense) to the offenders on whom they are imposed; nor is it unreasonable to suppose that different items should figure in each of those categories, where different types of offenses are concerned. But, as we have seen, in several cases the results of the analyses done by Zalman and his colleagues were counter-intuitive if not downright nonsensical; and in several other cases, two or more variables included in their modelling equations were probably highly intercorrelated. Most important of all, it appears that nothing whatever was done by the Michigan researchers to validate (in the statistical sense) the model(s) which they fitted to their data. Some of their less plausible findings (and some of their more plausible ones as well) may in fact have been due to nothing more than capitalizing on chance

variation. Finally, we see little reason to agree with their conclusion that their "Offender" effects were more important (in the sense of predicting sentencing outcomes) than their "Offense" effects. It appears, in fact, that the opposite is the case.

Let us, however, suspend disbelief for the moment, in order to consider what Zalman et al. make of the results of the analyses which we have so far described. Their next step is to create what they call "empirical sentencing matrices", one purpose of which was to illustrate the form which guidelines based on their findings might take.[92] For each of the ten major offense categories, the Offense and Offender variables derived earlier were used to define the rows and columns of a matrix. These scores -- the sums of whatever variables had entered the earlier equation, weighted by their unstandardized regression coefficients -- were subjected to a z-score transformation; the resulting scales were then chopped, only a little bit arbitrarily, into five categories.[93] On the assumption that these scales were orthogonal,[94] they were used to form a five-by-five matrix. Cases were then cast into the 25 cells of this matrix, on the basis of their (transformed) Offense and Offender scores; a moderately complicated set of scoring sheets was provided for this purpose. The sentences imposed on offenders in each cell were then summarized and displayed within the cell; separate matrices were constructed for the "in-out" and length variables, for each crime category. Table 9.14 illustrates the result of this process, using the matrix describing length of incarceration for those convicted of Sex Crimes.

Insert Table 9.14 here

The procedures just described should ideally have led to matrices containing about the same numbers of cases, with sentences tending to increase in severity from left to right and top to bottom. Inspection of Table 9.14 shows that both of these things are approximately true for it; but only approximately. In this matrix, and in the other 19 contained in Zalman et al.'s report, there are some cells in which mean or median lengths of term do not behave properly; in particular, there are several matrices in which the cases are not evenly distributed across the 25 cells, owing to the distributions of the Offense and Offender scores, and some correlation between them.

Inspection of Table 9.14 will also show (as Zalman et al. point out) that there is considerable variation in lengths of terms within some of the cells; in general, though not

Table 9.14: "Empirical Sentencing Matrix" for Sex Crimes, Michigan Sentencing Research (Based on Zalman et al. 1979:135)

Offense Score	Offender Score 1-2	3-4	5-6	7-8	9-10	Totals N	%
1-2	18 1.8 1-40	31 3.1 0-16	21 11.1 2-12	29 12.6 3-72	2 12.0 12-12	113	13
3-4	55 6.0 0-54	124 6.4 1-79	55 12.4 3-120	43 29.5 1-120	27 41.0 12-120	304	38
5-6	20 5.6 0-180	24 72.0 1-120	35 35.6 6-96	36 29.2 6-90	19 58.8 3-120	134	17
7-8	21 23.5 3-78	39 44.0 6-180	14 36.0 1-60	40 79.4 30-300	27 123.8 24-300	141	18
9-10	8 58.1 60-87	32 121.0 80-360	16 177.9 36-192	8 120.0 60-300	52 297.1 80-720	116	15
Totals: N	122	250	141	156	127	796	100
%	15	31	18	20	16		100

NOTE: 1st row in each cell represents the number of cases.
2nd row in each cell represents the median sentence term.
3rd row in each cell represents the sentence range.

invariably, the terms are skewed to the right, and there is often considerable overlap between the ranges of adjacent cells. Zalman et al. (1979:136-42) carried out a careful analysis of the frequency distributions of terms in their matrices, on a cell-by-cell basis; in our view this procedure is not only correct but essential, if the empirical method of matrix construction is used. From this analysis, however, they concluded that "...within each cell judges appear to be handing out almost any type of sentence; that is, no matter where the individual lies on the OFFENSE and OFFENDER dimensions, almost any sentence (up to the statutory maximum) is possible. We would have to conclude that there is not much predictability in sentencing, since similar cases are being treated very differently" (1979:136-42). They note that for other crime categories, e.g. Robbery, the situation is even more variable.

This conclusion is, of course, completely unwarranted. It is by no means necessarily true that Michigan judges' sentences are "unpredictable"; the most that can be said is that Zalman et al. failed to predict them with the additive model which underlies their matrices. Furthermore, while many cells do contain a few extreme outliers -- usually with very much longer terms than those in the middle of the distribution -- that does not by itself show that "similar cases are being treated very differently". At most it shows that some cases that are similar on Zalman et al.'s Offense and Offender scores receive different terms from most of the rest. But even if there were no doubts abouts the validity of those scores as descriptors of sentencing practice -- and we hope that we have made it clear that there <u>is</u> some doubt about that -- they plainly do not exhaust the sources of possible variation between cases, which might make much longer terms perfectly justified in a few cases. The original scores, after all, reflect weights which <u>on the average</u> minimized sums of squared deviations; that does not license the conclusion that those scores would predict sentences <u>in each and every case</u>. It must be pointed out again that Zalman et al. apparently did not analyze the residuals from their original regressions in any detail. Similarly, though they note the existence of a few "outliers" in several matrix cells, they did not examine those cases in any detail (or at least their report gives no indication that they did). As we saw in the preceding chapter, it may be appropriate to exclude some such cases, when fitting an equation designed to describe the majority; indeed, if this is not done, the result may be serious distortion of the regression coefficients. In the case of "empirical sentencing matrices" like the one in Table 9.14, the same thing is true. Extreme cases <u>may</u> be the result of disparity (in some sense of that term); but they may also reflect perfectly legitimate factors of a kind that would

justify a "departure" from sentencing guidelines based on the descriptive matrices. As Zalman et al. note (1979:153) both percentages incarcerated and median lengths of term are generally well-behaved, in most of their matrices. Given the limitations of their analytical methods, it may be thought that this is enough to ask.

In the penultimate chapter of their report, Zalman et al. consider some questions concerning "disparity" in sentencing.[95] Thus, they present the results of analyses designed to see whether the low R square values in their earlier regressions were due to aggregation of specific crimes into categories (they claim that they were not). They also investigate variation in sentencing between a small number of judges; we do not understand these analyses, and so do not commment on them.[96] They then examine variation in sentencing in relation to race, and across the three strata (metropolitan, urban and rural) from which they drew their sample. These analyses do seem to us to deserve some comment, since (a) they are only partly correct, but (b) they nonetheless invalidate much of Zalman et al.'s earlier analyses.

The analysis of racial variation begins by showing, somewhat superfluously, that some observed zero-order differences between whites and non-whites in percentages incarcerated and lengths of term are statistically significant; in fact, half of them (i.e. in ten of their 20 categories) are, with non-whites generaly being treated more severely. They next carry out analyses of covariance which, in effect, show that the coefficients associated with the Offense and Offender variables used in their earlier regressions display rather different patterns for whites versus non-whites. For example, in predicting the "in-out" decision for Sex Crimes using their variable 54 -- a dichotomous dummy variable scored +1 if "severe" mental trauma were inflicted -- the unstandardized regression coefficient is a statistically significant .222 for whites, but a non-significant .076 for non-whites.

What this suggests -- though it does not prove it, by a long shot[97] -- is that judges were applying different sanctioning rules when dealing with white offenders, from those used to sentence non-whites. (Even if Zalman et al. had not found such differences, it would not follow that judges were sentencing non-whites and whites in exactly the same way, of course: they might be taking exactly the same account of Offense and Offender variables for both groups, so that the partial slopes for those variables were the same, but just giving non-whites heavier sentences across the board. If this were so, it would show up as a difference in the constant

terms, or intercepts, of the two groups' equations. Since Zalman et al. nowhere report the intercepts of any of their regressions, we cannot see how far this might be true. (Cf. our analysis of the Massachusetts data in the last chapter.) At the very least, Zalman et al.'s analysis suggests that if race had been included in their original model, i.e. if they had fitted

$$S = \Sigma_i a_i A_i + \Sigma_j b_j B_j + cR + e,$$

where R is a dummy variable for Race, then the coefficient c associated with that variable would have been significant.

Surprisingly, Zalman et al. seem never to have considered this approach at all. It is important to do so, however, since if there is a significant effect of Race (or any other inappropriate variable present in the data), its inclusion in the model may well alter the values of the a's and b's, i.e. of the coefficients associated with Offense and Offender variables. We are not, of course, suggesting that a Race factor should be built into sentencing guidelines. If, however, there has been racial variation in sentencing in the past, and if it is desired to exclude such variation in guidelines for the future, then what is needed is an estimate of Offense and Offender effects, controlling for the effect of Race; and this can only be obtained if Race is included in the regression model, rather than being lumped into the error term as it was in Zalman et al.'s original model.

As they stand, then, the "empirical sentencing matrices" constructed by Zalman et al. are based on Offense and Offender effects which may be biased, perhaps to a substantial degree, by the effect of race. For example, suppose that non-whites tend to have very much worse prior records, but that over and above that, non-whites tend to receive more severe sentences because of (conscious or unconscious) racial prejudice, or for some other reason.[98] Then it may be that the coefficient associated with prior record will be different -- in this case, it will be higher -- if race is not included as an endogenous variable in the model. There will still be an important policy decision to be made when it comes time to construct sentencing guidelines, if the effects of race on prior sentencing practice are to be excluded: should we, in the future, sentence all offenders in the way in which whites were sentenced in the past? Or should we prescribe sentences for both whites and non-whites, that are like those given to non-whites in the past? Exactly how such an adjustment will have to be made, in practice, will depend on the exact nature of the race effect on previous sentencing practice. The important thing, however, is that the Offense and Offender

effects obtained from a model which includes race will themselves be uncontaminated by that effect.[99]

Precisely the same thing is true for effects of regional variation, or indeed any other variable which may in fact have influenced sentencing practice in the past -- whether legitimately or not. If there are such factors, it is important to identify them; and this obviously cannot be done if they are not explicitly included in the model or models used to describe sentencing practice. If they are morally iniquitous variables like Race, they will obviously not be explicitly included in sentencing guidelines. But they may nonetheless influence the guidelines, through their effect on Offense or Offender variables.

The issue of regional variation is an especially important one, of course, for proponents of statewide guidelines; as we say in the preceding section, this seems to have been one of the factors leading to legislative rejection of the Pennsylvania guidelines. In a table presented earlier in their report, Zalman et al. (1979:86) show that there was indeed such variation across their three strata (which may not themselves have been completely homogeneous). This variation was less marked than in Pennsylvania; and it displayed an interesting pattern. Over the whole population of 1977 sentences, about 48 percent were "out" sentences; 23 percent were to jail, and 30 percent were to prison. The metropolitan judges used "out" sentences in 52 percent of their cases; the urban judges did this in 43 percent of their cases. The rural judges, however, used "out" sentences only 36 percent of the time. At the other extreme, however, it was the metropolitan judges who made the greatest use of state prison sentences -- 34 percent, against 27 percent for the urban judges and 20 percent for the rural ones. Thus while judges in rural areas (like their Pennsylvania counterparts) used incarceration more frequently, they were much more likely to impose a jail sentence than a prison term; the metropolitan judges were just the reverse.

In their analysis of covariance across the three regions, Zalman et al. (1979:260-61) show that for the "in-out" decision for Sex Crimes, the coefficients of their model's variables are not too dissimilar as between urban and metropolitan judges; they suggest that the difference in percentage incarcerated across these two strata is thus to be found in the constant term of their fitted equation, though again they do not say what this is. For rural judges, the model fitted much less well. When they carry out a similar analysis for length of sentences for sex crimes, they find a somewhat similar pattern, though they note that the explanatory power of their model (as measured by R square)

varies directly with mean sentence length; that is, "the longer the average sentence given in a stratum, the greater the explanatory power of the model" (1979:261). It may be, however, that this is a direct consequence of the much greater use of jail rather than prison sentences by the rural judges -- a fact which in turn may not be due to different sentencing practices as much as to the greater availability of jail facilities (or the remoteness of prison facilities). In such a situation, it might be well to regress prison and jail sentences separately on candidate independent variables; if this were done, the differences in practice might not be very great. Again, it would seem that a term for "region" ought to be explicitly included in the model.

It is of interest to note that Zalman et al. follow this strategy part of the way in their analysis of what they call "other indicators of disparity" (1979:265-69). The three variables which they consider here are type of attorney, custodial status at time of sentencing and method of conviction (i.e. trial versus plea). What they do, however, is to regress the residuals from their Offense and Offender model on these other variables, finding several significant coefficients for the "in-out" decision though very few for lengths of term. What they do not do is to look simultaneously at Offense, Offender and "other" variables, to see whether the latter affect the former to an important degree.

In summary, it seems to us that Zalman et al.'s empirical analysis of felony sentencing in Michigan has certain shortcomings. We repeat that in many ways their work displays a high degree of sophistication and competence; and that some of the mistakes which we believe they made have also been made by many others engaged in this kind of research. We have discussed their work at some length, not merely because of the conceptual and methodological problems which it illustrates, but because (in contrast to Minnesota and Pennsylvania) it was intended to provide an empirical basis for sentencing guidelines; yet in many ways this is precisely what it does not do. It says nothing whatever about sentencing policies, i.e. the sanctioning rules actually (whether or not consistently) used by judges in Michigan in 1977; nor is it even a plausible external description of their sentencing practice, except at a very crude level. Of course the most important items used by judges in choosing sentences are related (in some sense) to the instant offense and the characteristics. But what are those items, and how are sentences related to them? Linear additive models are usually the first thing one should try, and occasionally they provide an acceptable description; but they are not guaranteed to do so, and they are seldom the only thing that one should try.

Furthermore, the production of counter-intuitive findings (like the negative coefficients associated with the "bodily beatings" variable) are not something to be dismissed as an aberration or a mere curiosity; at the very least they should be excluded as artifactual, and if there are enough of them they are a sign that the model itself is wrong.

We also believe that much more consideration should have been given by Zalman et al. to the residuals from their regressions (and to cases falling outside the cells of their ingeniously-constructed matrices). Finally, if the notion of empirically-based guidelines is to have any meaning at all, then the empirical analysis of prior practice should aim to give as complete and accurate a description as possible -- which means including, in at least the first such modelling attempts, such variables as race, sex, social status or region if there is any reason at all to believe that these may (consciously or not) have influenced sentencing practice.[100]

Following the submission of Zalman et al.'s report to the Michigan judiciary in October 1980, its main recommendation -- that sentencing guidelines be developed and implemented -- was accepted. Beginning in March 1981, the guidelines -- which we describe briefly below -- were introduced in three counties, on a pilot basis; a comparison with pre-guidelines sentencing is to be carried out, though at the time of writing no information is available to us about the results of this, or about the possible use of the guidelines in the future.[101]

The guidelines themselves differ in several respects from the "empirical sentencing matrices" described above. For one thing, they deal with eleven offense categories rather than ten, since homicide and negligent homicide are treated separately in them.[102] Moreover, the shape of the guidelines matrices differs from that of the empirical sentencing matrices: the former contain three rows and six columns, whereas, as we have already seen, the latter are five-by-five. The scoring and coding instructions (to be used by judges and probation officers) are somewhat simpler than the scoring sheet used to classify cases into the empirical matrices; in particular, they do not require multiplication of case variables by regression weights calculated to three decimal places. But the variables by which the Offense and Offender scores used with the guidelines are defined are in many cases very different from those listed in the 1979 report; and different matrices are provided for sub-sets of the offenses included in each of the "empirical sentencing matrices". We illustrate these differences by considering the guidelines for Sex Crimes, which may be compared with the empirical matrix in Table 9.14.

There are in fact five separate guidelines matrices covering the offenses in the Sex Crime category, each one relating to a different set of those offenses. It appears that this arrangement was adopted because (as noted earlier) there is considerable variation among the crimes in that category, in respect of statutory maximum sentence; while it would perhaps have been possible to combine all of those crimes into a single, large matrix, this would obviously have been cumbersome to use. The five matrices are however summarized in a convenient composite table in the Michigan guidelines, which we reproduce here as Table 9.15. The three rows of this table reflect low, medium and high offense severity, and the six columns reflect the categories of the offender score; the five ranges in each of the cells correspond to the statutory maxima shown at the left of the composite matrix.

Insert Table 9.15 here

Several things may be noted about these guideline matrices. First, in pretty well every case, there is substantial overlap between the ranges in adjacent cells across the columns; for example, in the medium seriousness row for crimes with 180-month maximum terms, the cell in column B prescribes a range of 18-36 months, whereas the cell in column C prescribes a range of 24-48 months. (Moving down the rows of the matrices, however, there is only minimal overlap, at the end-points.) Second, there are not separate guideline matrices for the "in-out" decision, and for lengths of term for those incarcerated; instead, a single matrix tries to convey information about both of these decisions. It does so in a way which we have already argued[103] is extremely unsatisfactory, viz. providing ranges of the form "0 - X", where x is the normal upper limit in months for those incarcerated, and zero indicates that a non-incarcerative sentence is within the guidelines. Not only does this leave the "in-out" decision virtually uncontrolled in such cells; but in addition, the matrices provide judges with no indication whatever of the proportions of offenders in each cell who had previously been incarcerated. No doubt judges will realize that this proportion tended to increase, as one moved down the columns and across the rows of the matrix; but they can have no idea of how much the percentage incarcerated rose, from one cell to the next; nor are they even told the modal sentence ("in" versus "out") in each cell, in the past. What kind of guidance is that?

Next, let us consider the ways in which cases are defined, by Offense and Offender variables, so as to be

Table 9.15: Comparison of Guidelines Within Each Cell by Statutory Maximum for Sex Crime Groups, Michigan Sentencing Guidelines 1980 (Source: Zalman et al., 1980)

STAT MAX		A	B	C	D	E	F
Life/years	LOW	0-36	12-48	12-48	36-60	48-84	72-120
180 months	LOW	0-18	6-18	12-24	18-36	24-48	36-72
120 months	LOW	0-12	0-18	6-24	12-36	24-48	36-72
60 months	LOW	0-12	0-12	6-12	6-24	12-30	18-36
24 months	LOW	0-3	0-6	0-9	3-9	3-9	6-9
Life/years	MEDIUM	36-72	48-84	60-108	72-120	96-180	120-240
180 months	MEDIUM	0-36	18-36	24-48	36-48	42-60	60-96
120 months	MEDIUM	0-18	6-24	12-30	24-48	36-48	48-72
60 months	MEDIUM	0-12	0-12	6-24	12-30	18-36	24-36
24 months	MEDIUM	0-6	0-9	3-9	3-12	6-12	9-12
Life/years	HIGH	72-120	96-180	120-240	180-240	180-300/life	180-300/life
180 months	HIGH	24-48	36-60	48-72	60-96	84-120	96-120
120 months	HIGH	12-36	12-36	24-48	36-60	48-80	60-80
60 months	HIGH	6-18	6-18	12-30	18-36	24-40	32-40
24 months	HIGH	0-9	0-9	3-9	3-12	6-12	12-16

classified into the cells of the appropriate matrix. It will be recalled that in the case of Zalman et al.'s empirical matrices, cases were classified by their scores on Offense and Offender variables, weighted in each case by the unstandardized regression coefficients derived from their analyses; these weighted scores, once summed, were normalized by a z-score transformation which was then grouped to produce an approximately equal distribution over the five categories. In the guidelines, by contrast, the Offense variables used for a particular group of crimes are given integer scores; these are summed and then trichotomized to produce the High, Medium and Low levels of Offense Severity. The Offender scores are, in general, calculated from the same variables;[104] here too, integer weights are assigned, and the resulting sums split into six (rather than five) categories.

These two sets of variables, and their associated weights, seem to us to bear almost no resemblance to each other. For example, in the empirically derived "in-out" matrix, the statutory maximum (in months) is weighted by .00085; in the length matrix it is weighted by .329; in the guidelines rules it does not appear at all. "Type of weapon" is, in the empirical matrix for length of terms for Sex Crimes, a variable with a range of zero to four, weighted by a factor of 27.441; this variable does not figure in the "in-out" scoring at all. In the guidelines rules for counting Offense score, however, a variable called "Weapon: Presence, Type and Use" is included; this is a four-category variable (values of zero to three), and it is subsequently unweighted. In the empirically-derived scoring sheet for lengths of term, provision is made for four variables which had negative regression weights; these are not included in the guidelines.

The Offender variables in the two sets of matrices are even more baffling. In the empirically derived matrices for Sex Crimes, there were ten variables used to predict the "in-out" decision, and ten used to predict lengths of terms; in each case, eight of these had positive weights, and two had negative weights. Only two of these (one positively-weighted, the other negatively-weighted) are the same for both decisions. Moreover, while there is some overlap between the Offender variables used for Sex Crimes, and those used for other categories, there are many differences; and for those other offenses the variables used for "in-out" prediction and for predicting lengths of term generally show about as little overlap as those used for Sex Crimes.

In the guidelines, by contrast, the same set of Offender variables is used for all types of crimes (except negligent homicide); presumably this is for ease of calculation. The five variables included in the normal case are Prior Felony

Convictions, counted one or two each, depending on whether the offense in question was of "high severity"; Prior Similar Felony Convictions, scored either two or four apiece, again depending on severity; Prior Juvenile Delinquency Adjudications, scored two, one or zero depending on their numbers; Prior Misdemeanor Convictions, again categorized and scored zero, one or two; and Current Relationship to the Criminal Justice System, a zero-one dichotomy reflecting current incarceration, parole, probation, or pending criminal proceedings. None of these variables is given a further weight, in calculating column placement in the guidelines themselves.

The Offender variables used in the Sex Crime cases, by contrast, all carry regression weights (eight are positive and two negative) for both "in-out" and lengths of terms. These variables are as follows: aggregate number of incarcerations, sum of months of minimum prior terms, aggregate number of felony convictions, disposition of most recent parole, relation to criminal justice system at time of present offense, pending charges in other jurisdictions, detainers outstanding, "good moves" since arrest,[105] whether the offender has a job to go to, "mental health", degree of alcohol use, drug use status, reason for leaving school, and residential stability -- all signed positively in the equations they entered; and type of work, disposition of most recent probation, and number of violent felonies as a juvenile -- all signed negatively.

To give some indication of the effects of these changes, consider an offender who has committed an offense of Criminal Sexual Conduct in the second degree (180 months statutory maximum); he has used a knife, inflicted some bodily injury, and carried away his one victim to some other place; he had exploited the victim's vulnerability (we may suppose she was drunk, elderly and/or frail), had inflicted mental trauma but was not part of an organized gang nor did he act in concert with others; vaginal penetration but no sodomy or other penetration occurred. The offender himself, we may assume, had two prior felony convictions of "high severity" (as defined); one of these, for which he was on parole, is similar to the current offense. Then by our (admittedly uncertain) calculation,

his Offense Score is:

```
  3 for use of weapon
 +4 for some bodily injury
 +3 for asportation of victim
 +3 for exploitation of vulnerability
 +4 for mental trauma
 +3 for vaginal penetration
----
=20 total Offense Score, which is in the High severity row of
    the guidelines;
```

and his Offender Score is:

```
  4 for two prior convictions of "high severity"
 +4 for one of those offenses being "similar" to the current one
 +1 for being on parole
----
 =9 total Offender score, which puts him into Category F (the
    most serious)
```

in the guidelines, so that his prescribed sentence would be within the range 96-120 months in prison (minimum).

Where would such a case have fallen in the empirical sentencing matrices from which the Michigan guidelines are ("in part") supposed to have been based? In order to answer this question we shall have to invent some more facts about the case. Let us suppose that, in addition to the things just supposed, the offender knew the victim, but not well; that they had no "long-standing feud"; and that all other facts about the offense were as already described. Let us also suppose that the offender had made some "good moves" since arrest; that the total of his previous minimum prison terms was 150 months; that he had no pending charges in other jurisdictions; that he had not previously been on probation; that he had no juvenile convictions; that his mental health appeared satisfactory; that he had no drug or alcohol problem; that he lived with his parents; and that he had not left school under a cloud; and that his work was unskilled. If these (not wildly implausible) facts be assumed, the arithmetic for the empirical "in-out" decision in his case would have been as follows:

```
300 x .00085 for the statutory maximum              =   .2550
  1 x .1392 for the infliction of "mental trauma"   =  +.1392
  2 x .0719 for "bodily beatings"                   =  -.1438
                                                       -------
  For a total Offense score of                          .2504, and
```

2 x .034 for prior incarcerations	=	.068
1 x .093 for current criminal justice status	=	+.093
-1 x .203 for "good moves"	=	-.203
0 x .108 for manner of leaving school	=	0
0 x .0925 for drug status	=	0
0 x .045 for alcohol use	=	0
-1 x .042 for living with parents	=	-.042
-1 x .085 for unskilled work	=	-.085
0 x .318 for violent juvenile felonies	=	0

For a total Offender score of		-.169

Carrying out the prescribed z-score transformations on these scores locates the case in the cell at the intersection of row 4 and column 2 of the "in-out" matrix.[106] In the 1977 data (Zalman et al., 1979:155) the percentage incarcerated was, as it happens, only 20 percent: but this is almost certainly a function of the small numbers of cases (ten) falling into this cell, and inspection of adjacent cells -- especially the others in the same row -- suggests that a 70 percent chance of incarceration would not be unreasonable. This would clearly make the case presumptively "in". What about length of term? In this instance, the arithmetic (which, to spare the reader's patience and ours, we do not reproduce here[107]) leads to the following result: the case falls into the cell at the intersection of row 4 and column 2 of the "length" matrix; the mean term for the 39 cases in that cell in the 1977 data was 44 months, the median was 53.3 months, and the range was 6 - 180 months.

We find, therefore, that our admittedly hypothetical case would have been located, in the "empirical sentencing matrices" of Zalman et al. (1979) in a cell which on its face had an incarceration rate of 20 percent, but for which common-sense adjustments suggests a presumptive "in" sentence; the median term to be expected was about 53 months, within a range of 6 - 180 months. The guidelines, by contrast, prescribe a range of 96 - 120 months in the "normal" case. The middle of that prescribed range is about 108 months; the spread around that mid-range is about 11 percent. Thus the prescribed mid-range is something over twice the median for "similar" cases in the 1977 data.

We think it important to stress that our hypothetical case is by no means a fanciful one. Yet for cases of that kind -- and the phrase "of that kind" obviously has a pretty flexible meaning, in this context -- the middle of the prescribed range is over twice the middle of the antecedent distribution of lengths of terms -- even if we neglect the probability that such a case might exceptionally be given a departing "out" sentence, which is not suggested by the

guidelines themselves. Furthermore, as we have seen, it was not possible to classify this case in any straighforward fashion under both the guidelines and the empirical matrices, without adducing -- or, in this case, inventing -- a number of extra classificatory facts needed to put the case into the empirical matrix but not to cast it within the guidelines grid. We may well approve of the policy decisions which led Zalman and his colleagues to discard such social or personal variables from the structure of their guidelines. The point is that they obviously were policy decisions; there seems to us no way in which they could have been "empirically derived" from the statistical analyses in the Zalman et al. (1979) report.

Of course it may be that the hypothetical case which we have just used to compare the Michigan guidelines and the empirical matrices would be highly atypical, and that the majority of those sentenced for felonies under the guidelines would be classified in cells with prescribed ranges nearer (in location, if not in dispersion) to the sentences imposed on similar cases in the 1977 sample. Since we do not at present have a copy of the 1977 Michigan data, and since Zalman and his colleagues have not (to our knowledge) conducted or published analyses which would throw light on these possibilities, the issue is somewhat moot; there seems little point in inventing further hypothetical cases with which to make comparisons like the one we have just done.

Even if the Michigan guidelines, when implemented, do not lead to increases in the severity of sentences, they seem to be open to criticism on several grounds. As we have noted, the frequent use of cell ranges of the form "0 - x", and the often overlapping ranges, impose much less by way of constraint then say, the Minnesota guidelines. In addition, the Michigan cell ranges themselves are relatively wide -- they average plus or minus 25-30 percent around mid-ranges, in most cells -- so that there is still room for considerable variation without a departure. Against these points, the Michigan guidelines matrices do not display the clearly multiplicative structure which we saw characterized the guidelines in Minnesota and Pennsylvania, which we argued gave disproportionately great weight to the Offender Score or prior record. We have not carried out row-and-column analyses of all of the Michigan matrices; but some of them, at least, have been shown elsewhere by one of us[108] to be basically additive in structure, with only occasional aberrant residuals.

It is also fair to note that those responsible for developing the Michigan guidelines contemplated that they would be used in conjunction with a meaningful system of

appellate review, and that provision would be made for review and evaluation of the use of the guidelines, on the basis of routinely collected information on post-guidelines sentencing.[109] These are important elements in any reasonable system of sanctioning rules; and to the extent that the Michigan guidelines increase the consistency and rationality of sentencing in that state, they are plainly to be welcomed. It seems to us misleading at best, however, to describe the Michigan guidelines as "empirically based", since any resemblance between the guidelines matrices and the antecedent empirical analysis is best regarded as purely coincidental. Given the several defects of that analysis which we have noted, that is perhaps a good thing.

Summary and Conclusion

We would repeat, in conclusion, that we have not intended in this chapter to present an exhaustive review of all current or planned attempts to develop statewide sentencing guidelines. Instead, our purpose has been to illustrate some important features of the concept of sentencing guidelines, and some problems concerning their development on an allegedly "empirical" basis. Comparison of the Michigan, Minnesota and Pennsylvania guidelines with those of New Jersey and Massachusetts can illuminate a number of issues which it is important to consider in the evaluation of sentencing guidelines, especially those intended to be used on a statewide basis.

All three of the guidelines discussed in this chapter are presented in some variant of the familiar matrix form (two-dimensional in Minnesota and Pennsylvania, three-dimensional in Michigan). This is not, of course, intrinsic to the concept of guidelines; as we saw in preceding chapters, those developed in Massachusetts take a different and in many ways more natural form. The Michigan guidelines, developed as they were from regression equations, could easily have been presented to the judges in a similar form to those of Massachusetts; indeed, it seems to us arguable that it might have been better if they had been.[110] But the Michigan researchers' procedure for translating their regression results into matrices seems to us an interesting one, which has an important consequence for the guidelines which may result. Having derived Offense and Offender scores for each of their crime categories, they transformed these to z-scores, and divided the resulting variables into five categories: the result of this exercise, in most cases, was to secure a fairly even distribution of cases over the "empirical sentencing matrices" thus created.[111] In Minnesota and Pennsylvania, by contrast, the rows and columns of the matrices were defined judgementally; a consequence of

this procedure, whatever its merits, was a very uneven distribution of cases within the matrix, e.g. a heavy concentration of cases in the upper left-hand corner of the Minnesota matrix. (Of course, this may have been exactly what the Minnesota Commission wanted, for political reasons.)

It should be clear that in all three of the states discussed in this chapter, empirical research on antecedent sentencing practices, though carried out at some expense, played virtually no substantive role in the creation of the guidelines ultimately developed. Does this mean that that research played no role at all, in the development of those states' guidelines? Of course not; what it means is that the role played by social science research in those states (and, in all probability, in New Jersey and Massachusetts as well) was primarily a political role, rather than the substantively important role envisaged by Gottfredson and Wilkins when they proposed the idea of "empirically based" guidelines in the first place. To put it bluntly: in all five of the states whose sentencing guidelines have been considered in this report, the idea of starting with empirical research on past sentencing practice was used to sell the idea of guidelines as a technique of sentencing reform, and nothing more pretentious. To be sure, the salesmanship varied considerably in quality, and the real customers were not necessarily those who were ostensibly offered the "product". (Indeed, as we saw earlier in Massachusetts, though the judiciary were to use the product, the empirical approach was lauded, at least in part, to placate the legislature.) Also, for instance, in Pennsylvania, the research on past sentencing practice done for the Sentencing Commission appears to have had little impact on critics in the state legislature (or anybody else), at the time of this writing; if Martin's (1981) account is correct, it appears to have had an adverse effect, e.g. by revealing the extent of regional variation in sentencing in the state. But poor salesmanship is not the same thing as no salesmanship.

In Minnesota, the Commission's staff did just enough research on past sentencing and releasing practice to meet the terms of the legislative mandate; the really important research done for that Commission was the estimation of the impacts of different policies on the state's prison population. In Michigan, the research by Zalman et al. (1979) was done with a view to showing that sentencing guidelines were needed; the guidelines themselves scarcely mirror past practice, in any way that we can see. No doubt there is merit in having some idea of past practice, even if guidelines are then made up entirely by fiat. But in none of the three states discussed in this chapter was the research which might have provided that idea properly designed and executed; nor

did those who wanted to introduce sentencing guidelines seem to know or care that that was so.

Consider, for example, the question of the possible effects of race, sex and social-background factors on past sentencing practice. The Minnesota Commission asserted that there had been some such effects, but produced little evidence that this was so; the Pennsylvania Commission's staff seems not to have examined the question at all; the Michigan researchers asserted that there had been evidence of racial (and regional) variation in past sentencing, though as we have noted their evidence is not entirely persuasive. Particularly in none of the three states, so far as we can determine, was there any attempt to estimate the indirect effects of race (or anything else) on sentences; the Offense and Offender scores may thus have built in some racial bias, instead of enabling it to be removed.

As we have noted in earlier chapters, the question of how wide the prescribed ranges in matrix cells should be is necessarily a judgemental one; in all three of the states discussed in this chapter, the ranges finally adopted were much narrower than the plus-or-minus 50 percent adopted in Massachusetts. The Michigan research shows, however, that empirical research can be of help in making judgements about range widths. Having created their "empirical sentencing matrices", Zalman et al. (1979:137-41) carried out cell-by-cell inspection of the frequency distributions of lengths of term; this is an extremely important operation, even though it was not, in our opinion, done as thoroughly by Zalman et al. as it might have been. It is in fact necessary to do this kind of analysis, if one is to evaluate the impact of guidelines on sentencing practice; without careful inspection of "outliers" both before and after the introduction of guidelines, it is impossible to know whether cases outside the prescribed ranges are justifiably treated as "departures". We will return to this point in our concluding chapter.

Throughout this report, we have stressed the point that sentencing guidelines -- whether or not they purport to be "empirically derived" -- are a species of sanctioning rule: that is, they are intended to be used by judges in imposing sentences in the future, whether or not this is made explicit. A range of terms and a probability of incarceration for the "normal" case are central elements in this kind of sanctioning rule; but they need not be the only elements, as two of the states discussed in this chapter illustrate. In Minnesota, a clear and reasonably detailed list of permissible and impermissible aggravating and mitigating factors accompanies the matrix, and thus constrains departures from the prescribed

ranges; Pennsylvania's rules (at least in their original form) went even further, and limited the amount by which judges could depart (unless "compelling reasons" could be found). In Michigan, by contrast, the guidelines now being used on a pilot basis provide no rules of any kind concerning departures; as in Massachusetts and New Jersey, judges are merely supposed to have some reason for imposing a sentence different from those permitted for the "normal" case. It is clearly no accident that in the last three jurisdictions the guidelines were judicially rather than legislatively developed.

Finally, it must be remembered that we can say nothing at all yet about the likely impact of the sentencing guidelines discussed in this chapter, on future sentencing practice in the three states. But the Minnesota research shows that it is important to estimate possible future effects (e.g. on prison populations) at the time when the guidelines are being constructed. It may be politically important to do this (as Pennsylvania's experience shows). But whether or not this is so, such estimation can provide information about a whole variety of consequences of different sanctioning rules. Some of the estimation and modelling techniques involved also make possible comparisons between different jurisdictions, of a kind which we illustrate in the next chapter.

Notes to Chapter 9

[1] In addition to the materials on states other than Massachusetts, New Jersey and the three states discussed in this chapter, we have some information (much of it of doubtful reliability and/or validity) on developments in other states, at the time of this writing.

[2] See, for some cross-state analyses of data from Massachusetts and Minnesota, Chapter 10 of this report. We have been informed in a personal communication by Dr. Marvin Zalman, director of the Michigan research, that their data are to be archived at CJAIN in Ann Arbor; this is also the case with the Massachusetts data which we analyzed, and the Minnesota data and programs. Contrast the situation regarding the New Jersey data described in Chapter 1 above.

[3] One of us (Sparks) testified before the Pennsylvania Sentencing Commission late in 1980; it is our understanding that the Pennsylvania data may be made generally available to researchers after the state legislature and Sentencing Commission have completed their deliberations. (Personal communication from Dr. John Kramer, research director for the Pennsylvania Commission.)

[4] As we have noted in earlier chapters, the guidelines of New Jersey and Massachusetts cannot easily be represented in matrix form. For some structural analyses of parole decision-making guidelines matrices, see Chapters 9 and 16 of Messinger et al. (1981).

[5] Minnesota Statutes ch. 244 et seq. (1978); Minn. Laws 1978, ch.723.

[6] One of the earliest proponents of "determinate" sentencing in Minnesota, State Sen. McCutcheon, had been a deputy police chief; it appears that he had a fairly substantial following among those in the law enforcement and other communities in Minnesota who were in favor of a "get tough on crime" posture. However, as Martin (1981:9) notes, McCutcheon was also a fiscal conservative and was thus opposed to increasing prison populations (which his 1975 determinate sentencing bill would have done). During our visits to the Minnesota Commission's introductory sessions (discussed later in this section), we heard very little support for increased "toughness".

[7] There were however several changes of personnel during the life of the Commission: see, for a detailed discussion of these and their impact on the Commission's work, Martin (1981)

[8] Martin (1981) cites an unpublished staff memorandum written in around October 1978 to the effect that the proper course for the Commission to follow was (in essence) policy-oriented rather than purely descriptive. However, a memorandum from about the same time (made available to us by Ms. Kay Knapp) suggests that, at least initially, the Commission's staff had argued for a "descriptive" rather than a "prescriptive" approach; it is not in fact clear exactly what this meant, or how the memoranda in question influenced the Commission's own views.

[9] We hope to present the results of some analyses of the Minnesota data, relating to this and other points, in the future.

[10] According to MSGC 1980:10, "consideration of system impact (i.e. availability of prison beds) played a rather passive role in determining the position of the dispositional line". The Commission's simulation program (MSGC, 1981) is however designed so that either proportions incarcerated or mean lengths of term can be varied under a variety of assumptions.

[11] The Minnesota Commission's staff tends to refer to its product as a "grid"; one of them, indeed, has been known to wax exceeding wroth at those referring to it as a "matrix". We do not, however, understand this linguistic idiosyncracy; and since the term "matrix" is generally used in this field we will stick with it.

[12] Initially the guidelines reflected a distinction of Minnesota law between a "stay of imposition" and a "stay of execution": see MSGC, 1980). This distinction (marked by a second staggering line, above the one shown in Table 9.1, is not made in the final matrix, though there are some rules concerning it.

[13] Compare the discussion of "running room" under the Massachusetts guidelines, in Chapters 6 and 7 above.

[14] See MSGC (1980:12-13). It seems misleading to say, however -- as Martin (1981:110) does -- that for such cases the Minnesota guidelines provide more severe sentences than their Pensylvania counterparts. The assessment of such conditional liabilities (e.g. imprisonment for probation or parole violation) is complicated: does one take the expected term times the probability of its imposition, for example? For empirical evidence on inmates' views, see Chapter 5 above.

[15] No details on the bases for these judgements appear in published reports, nor did we obtain any information on

them during our visits to Minnesota during the course of our project. The resulting division seems to us reasonable, though it would be interesting to have data (e.g. mean or median original ranks for the several offenses) on this point.

[16] See MSGC (1980:28-29). Only "statutory" misdemeanors (presumably the more serious kind) are to be counted; there is a restriction on multiple sentences for a single course of conduct; and no convictions antedating a five-year conviction period will be counted (for felonies this period is ten years).

[17] In its summary report on sentencing and releasing practices (MSGC, 1979:4), the Commission had in fact said that seriousness of offense and criminal history were "of roughly equal importance"; compare the statement from MSGC (1980:5) quoted in the text. This suggests that offense seriousness and criminal history (as each might be defined) would have about equal coefficients in a regression equation; the statement in the final report suggests that the criminal history coefficient would be greater. It is obvious that either of these findings might depend crucially on the ways in which the two variables were defined (a question which we plan to explore, using the Minnesota data, in the future). But in any case, to our knowledge, no such analyses were ever carried out by the Commission's staff.

[18] However, in its earlier summary report (MSGC, 1979:23) the Commission stated that the Minnesota Corrections Board placed more weight on offense severity than on prior record, in its parole decisions; the evidence for this claim (see Table 9.3) is not very strong either. In a staff memorandum dated 14 September 1979, on "Philosophical Implications of Drawing IN/OUT Lines", Parent contrasted the implications of policies based on incapacitation with those based on "just deserts"; the Commission itself eventually adopted a policy which it thought reflected more of the latter. The quoted statement from its final report is thus somewhat baffling.

[19] Nor did the Commission call attention to the negative association between offense seriousness and criminal history, though this is evident from tables in their 1979 report.

[20] Also excluded from the MCB sample were persons convicted of first-degree murder and other offenses not covered by the guidelines (e.g. incest) as a result of Commission policy decisions: see MSGC, 1979:19.

[21] In the data made available to us by the Minnesota Commission, the weighted N is 4,387; the sample N is 2,339. Our numbers differ slightly from those published by the Minnesota analysts, in some cases because of missing data and slightly varying definitions of variables.

[22] For a more detailed discussion of this "cohort problem" see Chapter 9 by Sparks in Messinger et al. (1981).

[23] Cf. Zimmerman and Blumstein (1979).

[24] The assumption made by the Minnesota staff in projecting future institutional populations was that 10 percent of cases sentenced would involve departures; the cash value of this assumption in relation to lengths of terms is not clear to us. The assumption itself may be thought to be somewhat optimistic, in view of the variances of most pre-guidelines sentences in most jurisdictions.

[25] We base this statement on correspondence between Dr. Dale Parent and Prof. Andrew von Hirsch, which was made available to us by von Hirsch; he had served as a consultant to the Minnesota Commission.

[26] The N's in this table (which is based on MCB data supplied to us by the Commission, and analyzed by us) also differ slightly from those in the 1979 report, probably because of minor coding differences or missing values.

[27] Minnesota Statutes, ch. 244.04; Minn. Laws 1978, ch. 723, Art. I, s.20.

[28] Few data are available from any jurisdiction on the numbers of prisoners losing good time, or the amounts of time lost. Most probably, however, the frequency distribution of losses is extremely skewed, around a very low mean.

[29] The implications of this change in the prison population for institutional management appear to have received very little consideration -- by the Commission or anybody else. According to Martin (1981:43), under the Minnesota Community Corrections Act, counties are expected to keep felons sentenced to terms of less than five years in local jails; a subsidy is provided to encourage this. Yet felons dealt with in this way will presumably be mostly those convicted of property crimes; the concentration of persons convicted of violent felonies in the state prison system may thus be even greater than the Commission's projection model suggests.

[30] Briefly, the model on which this projection program works (Knapp and Anderson, 1981) is deterministic rather than stochastic; and while it appears to be very flexible to use (in the sense that parameter modifications are easily made at each run), it seems that separate (and possibly quite complex) analyses may be needed to take into account demographic changes in the base population (i.e., the population "at risk" of conviction, imprisonment, etc.), over even a five year period. Since the prison population is very much a "downstream" consequence of earlier demographic and criminal justice processes, it may be very sensitive to very small changes in those things. See further, Chapter 10 below.

[31] See, for an illustration of this point using data from the English prison system, Sparks (1971), esp. Chapter 4.

[32] It is important to re-emphasize that this feature of most guidelines matrices is by no means necessary, nor is it required by the analytical techniques used in this section. On the contrary, from the point of view of a judge using the guidelines, it is merely a matter of convenience. Since "seriousness of offense" is used (in guidelines and other kinds of sanctioning rules) to justify sentences, it would perhaps be better if guidelines matrices were not ordered according to some presumed level ranking of seriousness or perceived seriousness; it is easy to infer from such a ranking that the rows of the matrix represent equal intervals, which is certainly not necessarily the case. This matter is further discussed in a forthcoming paper by one of us (Sparks) on "Empirical Evidence and the 'Just Deserts' Model of Sentencing".

[33] See MSGC (1980). Use of the mid-range seems most reasonable for such structural analyses in any case, though it is important to note that it is not regarded as a presumptive term for "normal" terms in all guidelines, e.g. those used by the U.S. Parole Commission and the Oregon parole board. See further, Messinger et al. (1981), Chapter 8.

[34] For further discussion of these techniques, see Mosteller and Tukey (1977); and Fairley (1978). The method is applied to parole guidelines by Perline and Wainer (1981), and by Sparks in Messinger et al. (1971). It is not necessary to use the row and column medians for such analyses; the mean can also be used, and is arithmetically easier for three-level and higher-order tables. However, as Tukey (1977) points out, the mean has the property of spreading the residual variance around the table, whereas the median -- to be preferred in most cases anyway, since it is more resistant -- shows residual "outliers" more clearly. We know of no comparative analyses on this point.

[35] The technique was in fact invented to cope with situations in which the assumptions of classical analysis of variance do not hold: see McNeil and Tukey (1975). It is of course an empirical question whether substantially different results would emerge from an ordinary ANOVA; again, we know of no comparative analyses of this question.

[36] It is of course open to argument whether ratios (rather than e.g. simple differences) are the best measures to use here: see the discussion by Sparks and von Hirsch in Messinger et al. (1981).

[37] Both the accounts given by Martin (1981) and by Commission and staff members to us during our project, are clear on this point: while there was indeed a conscious attempt to shift term-fixing (including the "in-out" decision) away from what was believed to be a somewhat incapacitative policy toward one of "modified just deserts", there is no indication that the relations between current offense and prior record displayed by the Minnesota matrix were ever recognized, let alone deliberately decided upon, by the Commission. It is, we feel, one of the strengths of techniques like those we illustrate here that they can help to bring to light relationships which may be unintended but nonetheless undesirable (or at least unwanted) consequences of policies.

[38] The U.S. Parole Commission's matrix is analyzed in Perline and Wainer (1981); the Oregon parole board's by Sparks in Messinger et al. (1981).

[39] Plots of this kind are sometimes referred to as RIOT plots -- an acronym of Ramsey's Inside-Out Trick, after their inventor. We are indebted to Howard Wainer for much information on this and kindred graphic techniques: see Wainer and Thyssen (1981) for a comprehensive review.

[40] There seems to be no theoretical reason why the residuals from fitting such a model should sum to zero, though they will probably be fairly close to it. But even in ordinary least-squares regression, where the residuals will sum to zero algebraically, it may well be the case that there are many more negative ones of comparatively small absolute magnitude, which are offset by a few large ones. Cf. our analysis of the Masachusetts data in the last chapter; and for a general discussion, Berk (1981).

[41] It may also be that more complicated models will give a better overall fit to such a table; but that near-zero residuals from the simpler model become much larger or change sign. If the purpose of the analysis is purely exploratory

(e.g. if nothing whatever is known or suspected about components of the variance of the response variable) more intricate fitting may be appropriate; here, however, we see no reason for it.

[42] E.g. Gelman, Kress and Calpin (1977); see above, Chapter 3. We also discussed this point in reference to Massachusetts in Chapter 6.

[43] Act 319, 1978 Pennsylvania Laws 1316.

[44] Though as Martin (1981) has noted, the law also provided that half of the members of the Commission would have one-year terms, and the remainder two-year ones; it is not clear to us (or to Martin) why this was done.

[45] Martin (1981:68) also states that Kramer's selection reflected the Commission's preference for "an individual without extensive knowledge or clear ideas about guidelines over someone who they feared had 'fixed' ideas and who might prove to be too controlling of the Commission." In fairness to Kramer we must note that we had no opportunity to obtain his views on this or any other matter, and that he may well take a different view of some of the issues discussed by Martin.

[46] Note that this last clause apparently does not require the Commission to specify permissible or impermissible circumstances (as it in fact did); it merely requires the stipulation of variations if some such circumstances are found. The Commission's apparently generous interpretation of its mandate seems to us commendable.

[47] It is not clear what this means, since from the legislation there appears to be no delegation of authority to the Commission (as distinct from the legislature itself) to deem certain revisions "necessary". (Contrast the Minnesota legislation discussed in the preceding section.) Since the initially-developed Pennsylvania guidelines needed legislative approval (which, in the event, they did not get), it seems to us surprising that subsequent revisions of those guidelines should be left entirely to the Commission.

[48] For an analysis of the effects on institutional populations of an earlier proposed reform in Pennsylvania, see Blumstein and Miller (1979); a more detailed treatment is presented in Blumstein (1981), of models for estimating the impact on institutional populations of system changes. It is worth noting, in passing, that much of this issue was effectively dodged in Minnesota, since (a) the guidelines only referred to those convicted of felonies, who (apart from the

provisions of the Community Corrections Act) were to be sent to state facilities, and (b) the Minnesota guidelines were explicitly devised to limit prison populations.

[49] In fact, the analyses which we present of the revised (January 1981) version of the Pennsylvania guidelines, later in this section, suggest that they would if anything have made the situation even worse where prison populations are concerned, than the original (October 1980) version. We have not been able to carry out any precise estimates on this point, however, since we have no information on the distribution of cases within the cells of the Pennsylvania matrices.

[50] See Martin (1981), who gives no indication why this was not done; given the structure of the basic Pennsylvania sentencing laws, it is surprising that the need for maximum-term guidelines as well as, or instead of, minimum-term ones, was not appreciated.

[51] In the report in which its revised guidelines were presented to the state legislature, it is stated merely that "the guidelines make no recommendations concerning maximum sentences" (Pennsylvania Bulletin, vol. 11, no. 4, 24 January 1981, p. 465). No justification for this procedure is offered in this report, or anywhere else to our knowledge.

[52] See Martin (1981). We have already criticized the "0 - x" format for guidelines ranges, in Chapter 3 of this report; see also our discussion of the Michigan guidelines in the penultimate section of this chapter.

[53] These "principles" not surprisingly turn out to be not dissimilar to those suggested in von Hirsch (1976), and in memoranda which he submitted to the Pennsylvania Commission in his capacity as a consultant or advisory board member for that project. See further, note 60 below.

[54] According to Martin (1981:79), there was still considerable dissensus within the Commission on this matter, and "further individual changes in seriousness rankings were made at virtually every meeting" after May 1980 (when von Hirsch's "principles" were accepted formally).

[55] According to Martin (1981:80-81), some Commission members argued that exclusion of such factors would be inconsistent with consideration of the "characteristics of the defendant" in appellate review; others felt they should be included, as one of the few ways to mitigate sentences. Martin refers to a Commission "survey of judicial attitudes" which indicated that over 80 percent of the judges regarded

such things as income and employment as appropriate considerations in sentencing; Martin gives no reference to this survey, and we have not seen any other information about it. It is very odd, however, that both in Pennsylvania and Minnesota "social status variables" (such as income) were excluded because they would have had "racial" implications. No doubt blacks in both jurisdictions (assuming that there are any blacks in Minnesota) have on average lower incomes and higher unemployment rates than whites. But has it ever been seriously suggested that poverty and unemployment are justifiable grounds for aggravation of sentences? Or that it is wrong in principle to take into account, as a mitigating factor, the wretched conditions in which many native Americans in e.g. Minnesota and Michigan are forced to live, when assessing their culpability? According to Martin, one of the strongest opponents of the inclusion of "social status" factors in Pennsylvania was State Rep. Norman Berson of Philadelphia, whose largely black constituents would have wanted them not to be included. This must surely rank as one of the most misbegotten judgements ever made: for while it may (or it may not) be true that judges in Pennsylvania or elsewhere have in some instances imposed heavier sentences on blacks than on comparable whites, it is bizarre to suppose that they would do this explicitly; excluding "social status" variables allows them to smuggle such racism in. Cf. our discussion of this point in the preceding chapter.

[56] This absurdity appears to be a direct consequence of testimony (in the Commission's public hearings) by a number of feminist groups. At the hearing at which one of us (Sparks) testified in Philadelphia, such an advocate also appeared; she appeared impervious to the counter-arguments of the Commission's chairman, which were admittedly not forcefully put.

[57] First-degree and second-degree murder are not covered by the Pennsylvania guidelines, since both carry mandatory sentences.

[58] The bill creating the Pennsylvania Commission was signed into law in November 1978; however, the Commission itself did not begin to meet until April 1979, and produced little of substance until much later in that year; by that time, our project's resources were committed elsewhere.

[59] We have been informed by Prof. Alfred Blumstein (a member of our project's Advisory Committee, and chairman of a current National Academy of Sciences panel on sentencing, that the Pennsylvania Commission has been reorganized, with Rep. Scirica as chairman; the Commission is supposed to present its revised guidelines in October or November 1981. The present

Governor of Pennsylvania, however, has meantime sought to introduce legislation that would, among other things, provide for a number of presumptive and/or mandatory sentences, while altering the existing legal relations between maximum and minimum sentences and (possibly) abolishing parole.

[60] The copy on which we rely was made available to us by Prof. A. von Hirsch, who, as we noted earlier (cf. note 53 above) served as a consultant to the Pennsylvania Commission. On internal evidence it appears that this draft copy was largely written by von Hirsch himself, though it may have been revised before submission to the Commission; we do not possess a copy of this document in its final form.

[61] See above, Chapter 3; and our discussion of the Michigan guidelines in the last section of this chapter.

[62] See above, Chapter 3, and 4 of this report. As an afterthought to our discussion of the New Jersey guidelines (in Chapter 4 above), it is ironic that it is only in these guidelines that real guidance -- in the form of percentages not incarcerated -- has been given to judges: this has not been done in any other states whose guidelines we have studied on this project.

[63] See Martin (1981:89-90), who notes that "Sentencing 'uniformity' now became a negative label; critics opposed uniformity for uniformity's sake because, they asserted, each county has its own particular problems and should be permitted to adopt its own policies for coping with them." As we have seen in earlier chapters, the issue of the proper locus of legislative authority has reared its head, in different guises, elsewhere; we return to it in our concluding chapter.

[64] Cf. our discussion of the Minnesota guidelines in the preceding section. As we have noted, those guidelines appear to provide that the mid-range (in those cells prescribing other state prison term ranges) is to be treated as presumptive in the absence of "legitimate, but not substantial or compelling" factors; under such a rule, even a range of the form "0 - x" might provide some constraint, if expressions such as "legitimate..., etc." were to be defined, either in rules accompanying the guidelines or in appellate review. But no such provision is, so far as we can see, a part of the Pennsylvania guidelines; there thus appears to be no requirement, in those guidelines, that judges show that e.g. non-incarcerative sentences are justified by any special factors.

[65] The strategy of extracting overall row-and-column effects may also lead to somewhat misleading results if some

cells of the matrix are structurally or definitionally empty (as is the case with presumptive "out" cells for which no alternative non-zero term is provided). Cf. our discussion of the New Jersey guidelines in Chapter 4 above. Methods exist for dealing with such situations in the analysis of contingency tables (in which the numbers in cells are counts of cases rather than measures of a response variable such as sentence length): see, e.g., Bishop, Fienberg and Holland, 1975. However, the extraction of overall effects using Tukey's (1977) methods may lead to results where there are structural zeros that are just as misleading as they would be in conventional analysis of variance. Analysis of sub-matrices seems the only solution in such cases.

[66] Or convicted of a crime involving the use of a deadly weapon (where this is not part of the definition of the offense): Pennsylvania Bulletin, vol. 11, no. 4, 24 January 1981 at p. 463. The prior record part of this mandate indeed suggests that the Commission was to start by defining ranges, etc., for the left-hand column of its matrix, and then decide how much increased severity was appropriate for those with prior records. Greater increases for those convicted of the less serious offenses is of course compatible with this directive; but it is not entailed by it, and nothing in law or logic required the Commission to adopt the approach it did. No justification for the ranges that were eventually adopted is presented by the Commission, nor does Martin's (1981) account suggest that the matter was ever explicitly considered. Here, we feel, is yet another example of a situation in which the consequences of a particular structure of sanctioning rules may not be obvious; the analytical methods we describe in this chapter are useful precisely because they may reveal such consequences.

[67] We have no details of the projection methods used; however, it appears that no computer simulation like that used in Minnesota was employed by the Pennsylvania staff.

[68] Cf. Chapter 3 above, where we discuss briefly the evidence on this point (e.g. Banks, 1964); see also the paper by Sparks cited in note 32 above, and Zalman et al. (1979:44).

[69] There is, so far as we are aware, little if any persuasive evidence on the ways in which offenders themselves conceptualize such matters; we suspect -- in part on the basis of the research described in Chapter 5 above -- that there is considerable variance among inmates in this respect.

[70] By this we do not mean that Zalman et al. ever envisaged guidelines in which an empirical summary of past sentencing practice was not to be modified at all, on grounds

of justice, expediency, or whatever. On the contrary, this idea is expressly repudiated at several places in their report (see, e.g., Zalman et al. (1979:25, 37-8, 96, 275)). However, it appears from their report (cf. p. xiv) that the first objective of their research was to produce "an information package detailing current felony sentencing practice to assist the Legislature, the Supreme Court and the Executive Branch in determining the sentencing policy for Michigan." As we show later in this section, the work of Zalman et al. is light-years away from the crude "empiricism" recommended by writers such as Gelman, Kress and Calpin (1977). Nonetheless it seems fair to say that their mandate was to produce guidelines which substantially embodied past practice, at least as a starting-point.

[71] People v. Tanner, 387 Michigan 683, 694-5, 199 N.W. 2nd 202 (1975).

[72] See, for example, Zalman et al. (1979:3-5, 174-85); there is evident confusion here and elsewhere in their report between the concepts of excessive variation in a correctly-principled penalty, differences in morally appropriate principles, and discrimination on the basis of morally inappropriate criteria: see above, Chapter 3.

[73] By "design effect" we mean the ratio of the standard error for a particular sample design (e.g. a multi-stage stratified one) to the standard error asociated with drawing a simple random sample of the size from the same population: see, e.g., Moser and Kalton (1971) for a discussion. For the Michigan sample the design effect appears to us to lie between 2.0 and 2.5; if this is so, then sampling variability probably does not much effect inferences made in the report, especially given the likely magnitude of non-sampling errors common in research of this kind.

[74] See above, Chapter 3; and Sparks (1981a), on this limitation; and compare the views of the Massachusetts judges discussed in Chapter 6.

[75] See Zalman et al. (1979:61). In our experience it is extremely rare for researchers even to report the results of such consistency checks, at least if the percentage of concordance is less than about 99.9. (However, compare our discussion of the Massachusetts guidelines project in Chapter 7.) The Denver guidelines study done by Wilkins et al. (1976), for example, discusses the problem of missing data at great length; but says virtually nothing about the problems of coding reliability; cf. also McCarthy (1978); Gelman, Kress and Calpin (1979).

[76] Curiously, there is little discussion in the Zalman et al. report of the problem of multiple-count cases, which may well need special treatment in this kind of research even if concurrent rather than consecutive terms are imposed; nor, apparently, was there any special consideration given to sentences imposed on parole violators for new offenses, though "status in the criminal justice system" was included as a candidate explanatory variable. Cf. our discussion of the "package" concept employed by some Massachusetts judges: above, Chapters 6-7.

[77] We infer this from the fact that in the summary table of maximum penalties in their report (Zalman et al., 1979:63-71), several offenses have the designation "M", though this is not explained in the note to the table or anywhere else in the report. Presumably these crimes, though technically only misdemeanors, carry penalties which under certain circumstances may be as great as those provided for by some of the felonies with which the project was mainly concerned.

[78] See, for example, the analyses in Chapter 7 of this report.

[79] As might be expected, almost 56 percent of the population cases came from the "metropolitan" stratum. Though no confidence intervals or levels of significance are reported for between-stratum comparisons, these are mostly large enough that they seem very unlikely to have arisen purely by chance.

[80] Other methods of estimating perceived within-category relative seriousness (which is what is in fact at stake here) are possible; e.g. one might carry out an exercise in ranking or rating similar to that done in Minnesota or by Gottfredson et al. (1978). No justification for using statutory maxima as "surrogates" for seriousness is given by Zalman et al.

[81] Contrast Zimmerman and Blumstein (1979); Larntz (1981); and for a general discussion see Berk (1981). No evidence in support of Zalman et al.'s conclusion is presented in their report; however, in our experience the conclusion seems reasonable. See also Gottfredson and Gottfredson (1981).

[82] We describe this as arbitrary because no evidence on the frequency distributions of expected probabilities of incarceration is presented by Zalman et al.; given the possible effects of sampling and measurement errors, among other things, it seems thus merely dogmatic to adopt a 50-50 split without inspecting the shapes of those distributions. Contrast Zimmerman and Blumstein (1979).

[83] The treatment of this finding (which is reported with commendable frankness) may indeed be symptomatic of a deep confusion in the "descriptive" approach ostensibly adopted by Zalman et al. and others working in this tradition. Having noted that their findings suggest that their model "is not doing a very good job of discriminating between those individuals that (sic) are ultimately incarcerated and those that are not" (1979:97), they should surely have junked -- or at least reconsidered -- their model? Yet it is precisely the results from this model that serve as the basis for their conclusion, near the end of their report (1979:295) that, while there are some patterns in sentencing in Michigan, these are only "faint and fuzzy"!

[84] The maximum percentages correct in these predictions are of course constrained by the marginal distributions involved: in particular by the fact that Zalman et al.'s model generally predicted over 80 percent of the cases as "out" dispositions (1979:97). It is not clear from their report how far these errors in prediction may be due to the decision by Zalman et al. to treat "split sentences" (e.g. probation and jail) as "out" decisions, and how far they may owe to the 50-50 split in percentages incarcerated as a criterion of the "in-out" decision (cf. note 82 above).

[85] This variable (no. 41), though labelled "Age of primary victim" in the Zalman et al. report, was apparently in fact coded +1 if the victim was "senior or juvenile", and zero otherwise (1979:E-5). The variable is thus presumably best interpreted as a proxy for "vulnerability". But what evidence is there that judges treat "vulnerability" as a parabolic function of age?

[86] We use the term "reliability" here not to refer to the consistency of the coding carried out by the Zalman et al. research team, but to the initial recording of the information contained in the pre-sentence reports from which they collected their data. The notion of validity only applies, strictly speaking, to the pre-sentence reports (as distinct from records compiled by the research staff); it is not of much importance if one accepts the assumption (made by most of those who have advocated "empirically-based" sentencing guidelines: see, e.g., Gelman, Kress and Calpin (1979), discussed in Chapter 3 above, that the contents of those reports constitute the most important part of the information base actually used by judges in sentencing. See, however, our discussion of some Massachusetts judges' views on the difference between tried and pleaded cases in this respect: Chapter 6 above; and our discussion of the Massachusetts probation records in Chapter 7 above.

[87] Such transformations may be helpful on grounds of mathematical tractability; they may also be theoretically appropriate, e.g. if the relationship of offense seriousness to penalty severity is curvilinear. See, however, our discussion of findings from the Massachusetts data, in Chapters 7-8 above.

[88] For the view that double cross-validation (in which separate samples are used to estimate the form of the statistical model, the coefficients of its included variables, and the stability of both) may be useful in criminal justice and similar social research, see Larntz (1981). A general discussion of "jacknifing" and other statistical validating procedures is contained in Mosteller and Tukey (1977), Chapter 8. See further, Chapter 3 above.

[89] To wit, the carrying out of a z-score transformation. See also p. 154, where the "in-out" scoring sheet for Sex Crimes shows a mean Offense effect of .118 and a mean Offender effect of .09 -- which suggests that Offense effects are about 1.3 times as important as Offender effects for that decision. (It is quite a separate point that those effects may have been incorrectly estimated; the point is that on Zalman et al.'s own assumptions, their conclusion seems mistaken.)

[90] Strictly speaking, of course, Zalman et al. did not test anything at all; they had no model and did not validate their findings. These are separate criticisms from the one we make here -- though it is arguable that this does not justify our loose usage.

[91] E.g. Zalman et al. 1979:170, 270-72, 277-78.

[92] Zalman et al. (1979:126) state that the matrix "allows us to see if judges, as a group, are making decisions in a predictable fashion. In addition, the sentence matrix provides the Steering Committee with an introduction to sentence guidelines -- with all their attendant problems -- so that when (sic) the decision to develop sentencing guidelines is made, the methodological issues underlying their construction will be clearly understood." No doubt, given the ubiquity of the matrix form in which sentencing and parole guidelines have been presented since the earliest work of Gottfredson and Wilkins, the Michigan team's decision to create such matrices was a natural one. In addition, the methods by which they transformed their regression results into matrices with cells containing ranges is (as we discuss in the text) a marked advance over those suggested by others (e.g. Gelman, Kress and Calpin (1977)). However, it should be clear that the matrix form does nothing whatever to show

whether or not "judges, as a group, are making decisions in a predictable fashion". If anything, it obscures this crucial question; cf. our analyses of the expected sentences under the Massachusetts guidelines, and the residuals from the associated regressions, in the preceding chapter. Moreover, it seems clear to us that the regression-like form taken by the Massachusetts guidelines is at least as useful for explaining the central concepts behind "empirically-based" guidelines, as the matrix form which has been so widely used. We return to this point in our concluding chapter.

[93] The cutting points of this "standardization scale" (Zalman et al. 1979:129) were evidently chosen for ease of use; if the distribution of expected probabilities or terms had in fact been normal, they would have resulted in about 16 percent of the cases falling in the two extreme categories (viz. 1 and 5), 24 percent in the next two, and 20 percent in the middle category. Given even approximate normality of expected scores, this is a much more even distribution of cases across categories than appears in other states' guidelines (e.g. Minnesota's). Of course even this method will not work where the cases in the matrix are very homogeneous in some respect, e.g. prior record; for example, in their Homicide in-out matrix, Zalman et al. (1979:156) found all of the cases falling either in column 1 or in column 4 of the matrix. Not much can be done about that, so far as we can see.

[94] This assumption is in fact a little shaky. In their report, Zalman et al. (1979:129) state that correlations between their two vectors for their ten major crime types "are small, supporting the conclusion of orthogonality (independence) between scores". Waiving the fact that small correlation coefficients show nothing of the sort (since a zero correlation may easily result from a perfect but non-linear relationship between the two variables in question), inspection of their table of coefficients (p. 130 of their report) shows that seven of the 20 are between zero and .10; seven between .10 and 20; and six greater than .20 with the largest being about .40.

[95] Disparity, "as a value term", is taken to refer both to "arbitrariness in sentencing" and also to "sentencing that is systematically related to invidious factors." (Zalman et al., 1979:182). Leaving aside the question of what is "invidious" (e.g. gender or violation of a position of trust, neither of which was considered at all), this definition seems deficient, at least if measures such as R square are used (as they are, in this chapter of their report) to signal the imperfections of an unvalidated and possibly mis-specified model.

[96] For one thing, the dependent variable in these analyses is a single composite, in which a range of zero to minus 60 stands for non-incarceration and positive values for numbers of months imposed; no argument for the validity of this scale is presented. This variable is then plotted against the statutory maxima for offenses dealt with by a total of 21 judges; the offenses in question were lumped into "violent" and "non-violent", and offender scores dichotomized at their mean, to calculate average expected scores for the 21 judges. Each judge's observed sentences are then compared with these averages, in (to us) some rather complicated graphs (Zalman et al., 1979:200-227). Why, we wonder, did they not simply calculate individual judges' regression equations, and look at variations in the weights (as was done, e.g. by Rich et al., 1980 in their re-analysis of the Denver data)? Alternatively, each judge's residuals from the Zalman et al. model (which model may admittedly not be correct) could have been examined.

[97] Not least because only the variables in Zalman et al.'s overall model are examined in this analysis; and it may be that none of their judges were following that model, even in the (debased) sense of "following" implied by the model's giving a passable external description of the judges' sentencing behavior.

[98] E.g. because they tend to commit what are perceived as very much more serious crimes, though this perception is not correctly measured by the "seriousness" variable being used -- as may well be the case here.

[99] We are indebted to an unpublished paper by Professors Franklin Fisher and Joseph Kadane, and to discussions and personal communication with them, for much insight into this general problem. The two senior authors of this report intend to examine this issue in more detail in a separate report.

[100] It cannot of course be concluded that variables having no zero-order associations with sentencing outcomes are in fact having no influence on those outcomes; they may be masked by "suppressor" or "distorter" variables (Rosenberg, 1968), for example. Testing all possible interactions of that kind is literally insane (since there is no logical limit to the numbers of things that may be having such effects). This point merely underlines the desirabil ty, to put it mildly, of having some theory of how judges actually do decide what sentences to impose -- a theory of the kind which we had hoped to develop in the never-funded Phase II of our research (cf. Chapter 1 above), and of the kind which neither Zalman et al. nor anybody else who has tried to develop "empirically based" guidelines has done.

[101] It appears from the guidelines test material which we received from Zalman that this pilot study was to have continued for six months, and would at that time be evaluated. At the time of this writing we have no information about the outcome of this study, though we hope to obtain some in the near future.

[102]The category of "negligent homicide" is primarily concerned with motor vehicle offenses.

[103] See above, Chapter 3.

[104] Except for the "negligent homicide" category, in which one of the variables relates to previous driving offenses. This seems to us highly sensible; it raises the question however, of why different kinds of prior record variables should not also have been used for other categories of crimes.

[105] This variable, apparently borrowed from the New Jersey research (McCarthy, 1978; and see Chapter 4 above) is in the Michigan scheme of things a dichotomy, giving credit (through a weight of -1) for such things as getting a job, voluntarily entering a drug treatment program, making restitution, etc. That these things entered several regressions may perhaps show that judges were in fact impressed by them; but should they be? Presumably not, since this variable is not in the guidelines themselves.

[106] Note, however, that it would take only a slight modification of the supposed facts to locate the case in a different cell -- for example, failing to make any "good moves" would (if our arithmetic is correct) move the case to column 4.

[107] The interested reader, if there is indeed one, may consult Zalman et al. (1979:134-35), and the associated matrix of the guidelines.

[108] See Sparks (1981a).

[109] We have no information on the machinery by which such a review would be carried out; the Zalman et al. report (1979) mentions both legislatively-mandated and judicially-sponsored sentencing councils as possibilities. Cf. our discussion of this problem in Massachusetts, in Chapter 7 above.

[110] Although, as noted, the equations in question might have needed to be purged of variables with counter-intuitive coefficients. The weights would presumably also have been

rounded to facilitate use by the courts; but as the Massachusetts example (above, Chapter 7) shows, this is not a problem.

[111] It is important to note, however, that such an even distribution, even if it could be accomplished, might not produce a linear (or otherwise reasonable) progression of terms of probabilities of incarceration; to bring about that result, it may be necessary judgementally to add one or more rows or columns. Cf. Perline and Wainer's (1981) discussion of the U.S. Parole Commission's guidelines on this point.

Chapter 10: An Empirical Comparison of Guidelines Structures in Two States

By Donald M. Barry and Timothy L. Kennedy

Rationale

There are several legitimate approaches to understanding the structure of sentencing guidelines and comparing those in one jurisdiction to those elsewhere. Examining pertinent legislation and comparing provisions from one set of guidelines to another is certainly a worthwhile endeavor, but it cannot provide a complete picture of any set of guidelines, nor of the implications that the guidelines have for practice, unless there is a way to tie the results of that kind of a priori analysis to empirical data reflecting actual populations of offenders.

To cite an admittedly extreme hypothetical example: it may be that State A's sentencing guidelines prescribe three months of probation for anyone convicted of ferret breeding without a license, while State B's guidelines call for a hefty prison term for the same offense; and this difference may be of interest for any number of reasons. However, if it is also the case that no one in either state has been convicted of the offense for the last hundred years, nor is likely to be for the next hundred, then the difference loses some of its significance, at least insofar as judicial and correctional policy-making are concerned.

More realistically, but to the same point, it is assuredly a fact that some types of offenses occur more frequently than others; it follows that those provisions which govern the more common offenses will affect a larger number of cases than will the provisions for less frequent offenses. Similarly, since most guidelines schemes take account of the offender's criminal history, the prescriptions governing offenders with few or no prior convictions will affect more cases than the prescriptions governing long-time repeat offenders, simply because the former substantially outnumber the latter.

A third determinant of sentences under any guidelines scheme is created by the particular combinations of current offense types and prior records that may be exhibited in a given population of offenders. Any statistical association between current offense seriousness and extent of prior record will influence the distribution of sentences prescribed by guidelines that take account of those two factors, and the nature and strength of any such association can be ascertained

only via the analysis of empirical data. For example, a strong positive correlation between offense seriousness and criminal history would result in a U-shaped distribution of punishment severity; i.e., there would be concentrations of cases at the lower and upper extremes of the severity continuum. Conversely, a strong negative correlation would produce a mound-shaped distribution (cases concentrated in the middle), because serious current offenses would tend to have been committed (assuming the negative correlation) by those offenders with the least serious criminal histories, and the less serious current offenses by those with the most serious criminal histories. The two factors would, in a sense, cancel each other out, resulting in sentences of intermediate severity for the bulk of offenders.

These concerns led us to search for a way of linking guidelines provisions with empirically-determined characteristics of offender populations, so that straightforward comparisons among guidelines would be possible -- comparisons which would take into account and control for: (1) differences in the frequency of occurrence of various crime types, (2) differences in the frequency of occurrence of various degrees of prior criminal involvement, and (3) statistical associations between current offense and prior involvement. We were looking, in other words, for a method for characterizing guidelines which would automatically "recognize" that differences in the two states' ferret breeding sanction provisions were not terribly important.

After we had thoroughly explored a number of blind alleys, we came upon an idea which, in theory at least, was quite simple and appealing, and which has guided the analyses reported in this chapter. The idea was this: In order to compare State A's sentencing guidelines to State B's, process the same group of offenders through the two sets of guidelines; then tabulate and compare the two sets of outcomes. Certainly, any differences observed cannot be attributed to differences in offenders or their offenses, and in this sense such comparisons identify "pure" differences between guidelines provisions.

Our original plan called for processing each of the two offender groups in our data base (those from Massachusetts and Minnesota) through both the Massachusetts and Minnesota guidelines. This approach would have enabled us to gauge the extent to which differences in offender populations affect the eventual conclusions produced by the technique, since the two samples, prima facie, represented two different kinds of offender populations. Our sample of Minnesota offenders is almost surely more representative of a "typical" population of all persons convicted of felonies in a "typical" state, since

its sampling frame was, in fact, all persons convicted of felonies in Minnesota during fiscal 1978. The Massachusetts sample, on the other hand, consisted of cases disposed of in the Massachusetts Superior Court where one or more of the offenses charged was a felony offense, and was more likely to represent a population of the more serious cases.

It soon became evident that processing Minnesota offenders through the Massachusetts guidelines was not feasible given the data at hand, and that even the simpler of the two procedures -- processing Massachusetts offenders through Minnesota guidelines -- would be a formidable task indeed. (And it was.)

In order to implement the procedure, it was first necessary to answer, for each offender in the data base, the question, "What if this person had (hypothetically) exhibited the same behavior in State B as he/she (actually) exhibited in State A?" To address the question, one must reconcile the considerable differences which exist between the statutory definitions of various crimes in the two states. Next, one must have available all of the current and prior offense information required for computing guidelines sentences. As it turned out, we did have sufficient information to translate Massachusetts offenses and offenders into their Minnesota equivalents, but the converse was not true. Determining what would have been guidelines sentences for Minnesota offenders, had they committed their crimes in Massachusetts, was next to impossible because of the relative complexity, and detailed information requirements, of the Massachusetts guidelines. Particularly in regard to current and prior offense seriousness, weapon use, and injury to victims as required by the Massachusetts guidelines, the Minnesota data base did not contain enough micro-level information to permit even educated guesses as to the Massachusetts guidelines equivalents. (Of course, there is no reason why it should have. The Minnesota data were gathered with a view toward establishing Minnesota sentencing guidelines, not for facilitating sentencing research in general, or for cross-jurisdictional comparisons.)

Specifically, there were over 200 Massachusetts offense categories, most of which could be changed somewhat in meaning by any of nine "modifiers". Associated with each offense is a seriousness score; these are weighted, summed over all present offenses and all prior offenses, and added to other indices reflecting weapon use and victim injury, to produce the final Massachusetts guidelines sentence, as we described in detail in Chapter 7. The result is a nearly continuous range of possible Massachusetts guidelines sentences. Of course, the Massachusetts data were sufficiently rich to permit these calculations.

In contrast, the Minnesota guidelines required (and the Minnesota data contained) much less refined information. One can compute the Minnesota guidelines sentence by locating the offender within one of only ten current offense seriousness categories, and one of only seven prior record categories. In the case of Minnesota, the result is a much more manageable set of only seventy possible guidelines sentences.

Since the information we had about the characteristics of sentenced offenders in each of the two states was quite different, as were both states' guidelines structures, it would not have been possible for us to compare the effects of both sets of guidelines on both sets of offender populations. Thus, some choices had to be made as to how to proceed. We had more than enough information about Massachusetts offenders to enable us to generate hypothetical guidelines sentences had those offenders been convicted in Minnesota, but not enough detailed information about Minnesota offenders to determine their equivalent Massachusetts guidelines sentences. These considerations led to our decision to examine only the outcomes for Massachusetts offenders processed through the two states' guidelines.

One consequence of this decision is that the results reported here will not necessarily be generalizable to populations of all felony offenders, because of the character of the Massachusetts sample (Superior Court cases only). Results and projections ought to be taken as representative of the more serious felony cases. This is probably not too severe a limitation: the more serious cases are precisely those which warrant the most attention, consuming as they do the bulk of court and correctional resources, and culminating as they do in the widest range of sentencing outcomes, particularly incarceration for terms of one year and up.

A second consequence may be a bit more far-reaching in its implications for the conclusions we draw from this comparison of the structures of the Massachusetts and Minnesota guidelines, and the sentence outcomes that those structures produce. The Minnesota sentencing guidelines do not distinguish among plea and trial dispositions when determining sentence. The Massachusetts guidelines, on the other hand, do make a distinction between those cases disposed of by plea and those disposed of by trial. As we described in detail in Chapter 7, the Massachusetts sentencing guidelines are mandated for use in trial cases and may be voluntarily consulted in cases involving a plea disposition. We have chosen not to emphasize the distinction between trial and plea cases in this chapter -- and thus have analyzed the effect of the Massachusetts and the Minnesota sentencing guidelines on sentence outcomes for both trial and plea cases combined. One

reason that we proceeded in this manner was that, though the distinction between plea and trial cases may be important to an analysis of the impact of the Massachusetts guidelines on Massachusetts sentencing and correctional policies, a comparative analysis of two states' guidelines structures will not be much affected by it. Also, since Minnesota's guidelines do not differentiate between trial and plea cases at all, and since Massachusetts judges may, if they wish, use their guidelines for both types of cases, it appeared simpler to hypothesize that there would be uniformity in the application of both sets of guidelines. The uniform application that we decided to impose in this analysis in the interest of simplicity of comparison was that all cases would be processed through both sets of guidelines. We have taken pains here to warn the reader about this point, however, to emphasize that the results of our analyses merely represent a comparison of the severity of the overall guidelines structures for both states; these results are not intended to make a definitive statement on how severe the Massachusetts guidelines actually are in use in comparison with those of Minnesota.

All of the analyses to follow will treat differences between the Minnesota guidelines and the more recent version of the Massachusetts guidelines (Version 2); most of these will also examine, for comparative purposes, Version 1 of the Massachusetts guidelines, and the actual sentences handed down by Massachusetts judges for the cases in our sample.

Method of Offense Translation: Current Offense

By far the most difficult part of this analysis involved converting Massachusetts offense descriptions into their Minnesota equivalents, to determine current offense seriousness as required by the Minnesota guidelines. In a great many instances, it was necessary to supplement the information conveyed by the Massachusetts offense description and modifier with other case data (e.g., extent of injury to victim(s), victim's age, type of weapon involved, value of property stolen, type and street value of controlled substances seized) in order to arrive at a unique Minnesota counterpart. The procedures we used, and the resultant cross-classification of offenses, are given in the Appendix to this chapter.

Method of Criminal History Index Construction

The methods of calculating criminal history scores in Massachusetts and Minnesota were described in earlier chapters; we summarize them again here, however, to facilitate our discussion of the translation of one state's scores to the other.

Under the Minnesota guidelines, a criminal history index ranging from zero to six points is calculated on the basis of four types of information: (1) prior felony record; (2) custody status at the time of the present offense; (3) prior misdemeanor and gross misdemeanor record; and (4) prior juvenile record for young adult felons. In general terms, an offender is assigned one point for each prior felony conviction, one-half point for each prior gross misdemeanor, and one quarter point for each prior misdemeanor; but there are several qualifications which somewhat complicate the calculation of the criminal history index, making it more than just a matter of tallying past offenses.

A single course of conduct can contribute no more than one point (or two, in case of multiple victims) to the index, even if numerous separate offenses of conviction were associated with the single course of conduct. Prior felony convictions which resulted in a misdemeanor or gross misdemeanor sentence (up to ninety days, and between ninety-one days and one year, respectively), count as prior misdemeanors or gross misdemeanors rather than as felonies. A felony conviction which results in a stay of imposition (i.e., probation) is counted as a misdemeanor after five years from the date of discharge, if the stay of imposition has been successfully served. Similarly, a misdemeanor conviction is not counted if a five-year, conviction-free period has followed the date of discharge or expiration of the sentence.

If the offender's record is free of all convictions for a ten-year period beginning at the date of discharge or expiration of the sentence, a prior felony conviction is not used in computing the index. An offender can receive a maximum of one point for all prior misdemeanor or gross misdemeanor convictions. One point is assigned if the offender was incarcerated, on probation or parole, or released pending sentencing at the time of the present offense. And finally, for persons being sentenced for current offenses committed before their twenty-first birthday, one-half point is assigned for each juvenile adjudication for crimes which would have been felonies if committed by an adult, up to a maximum of one point for prior juvenile adjudications.

In some cases we could not unequivocally assign a Minnesota criminal history score to the Massachusetts offenders. We adhered as closely as possible to the rules presented above, but when ambiguities did arise, they were handled as follows.

First, our data on Massachusetts offenders' prior convictions were not nearly as detailed as those describing current offenses, so, in differentiating among prior felonies,

gross misdemeanors and misdemeanors, we applied the Minnesota rule (noted above) which uses length of sentence as the criterion for seriousness of prior offenses. That is, an offense for which the maximum sentence actually imposed was greater than one year, or for which the period of probation was greater than one year, was considered a prior felony offense. A maximum sentence, or term of probation, of four months up to one year was considered indicative of a prior gross misdemeanor conviction, and a one-to-three month maximum sentence or term of probation was assumed to reflect a prior misdemeanor conviction.

Second, the Minnesota provision that multiple convictions resulting from a single course of conduct be counted as a single conviction was somewhat problematic; this information was not present in the Massachusetts prior convictions data. As a proxy for it, we used date of prior disposition; that is, when more than one prior disposition was observed for a given offender on the same day, it was assumed that the dispositions reflected convictions arising from a single course of conduct.

Following the other Minnesota rules noted above was fairly easy, as the Massachusetts data base did contain information on each offender's custody status at the time of the offense, number of prior serious juvenile adjudications, and the dates necessary to exclude certain prior offenses.

Finally, our convention for handling guidelines "out" decisions should be noted. Minnesota presumptive "outs" are only "out" of the state prison system -- they may be sentenced to the county jail for terms up to one year. And, Massachusetts "outs", as we have tallied them, consist of those persons who, under Massachusetts guidelines, were not presumptive "ins", i.e., whose guidelines sentences were less than seven months (Version 1), or less than six months (Version 2). Stating that these persons are presumptive "outs" would be inaccurate: under Massachusetts guidelines, the in-out decision in these cases is at the discretion of the sentencing judge.

We have then, under both states' guidelines, some unknown proportion of those labelled "outs" who might in fact serve jail terms up to one year (Minnesota), or up to six or seven months (Massachusetts), in full compliance with the guidelines in the respective states. In the absence of data telling us what judges actually operating under guidelines would do in these discretionary situations in each state, we have simply counted them as "outs". This convention should be borne in mind as the two states' guidelines are compared in the sections that follow. Our major conclusions, however, ought to be unaffected by it.

The In-Out Decision

Tables 10.1 through 10.6 present the six possible pairwise crosstabulations among the four sets of in-out classifications, and sizeable discrepancies between the members of each pair are evident. The entries in the two main diagonal cells of each table represent, of course, agreement on the in-out decision, while off-diagonal cells represent disagreement.

Tables 10.1 through 10.3 make it clear that the Minnesota guidelines prescribe imprisonment far less frequently than either version of the Massachusetts guidelines, and also far less frequently than was exhibited in the actual decisions made by Massachusetts judges.

Insert Tables 10.1-10.3 here

Table 10.4 demonstrates that Version 2 of the Massachusetts guidelines operates as if to apply Version 1 first, and then to incarcerate an additional 42 percent (131 of 310) of those who would have been "out" under Version 1, with no accompanying decrease among the "ins" under Version 1.

Insert Table 10.4 here

Tables 10.5 and 10.6 suggest that actual judicial practice in Massachusetts is considerably more lenient (at least in "in-out" terms) than either version of the guidelines. For Versions 1 and 2, respectively, 28 percent and 34 percent of the entire sample would have been incarcerated under the guidelines, but were not incarcerated by sentencing judges. There are still, however, a few persons (7 percent and 4 percent for Versions 1 and 2) who would have been "out" under the guidelines but were in fact incarcerated.

Insert Tables 10.5 and 10.6 here

The fact that both versions of the Massachusetts sentencing guidelines appear to be more severe than the judges' actual sentences in that state should be qualified, however, by our earlier discussion. We voiced some concern that an analysis of the Massachusetts guidelines that did not distinguish plea from trial cases might overestimate the severity of those guidelines. The artifact produced by not

Table 10.1: The In-Out Decision
Minnesota Guidelines Vs. Massachusetts Guidelines, Version 1

Minnesota Guidelines	Massachusetts Guidelines Version 1		
	Out	In	Total
Out	279	484	763
In	31	497	528
Total	310	981	1291

Table 10.2: The In-Out Decision
Minnesota Guidelines Vs. Massachusetts Guidelines, Version 2

Minnesota Guidelines	Massachusetts Guidelines Version 2		
	Out	In	Total
Out	168	595	763
In	11	517	528
Total	179	1112	1291

Table 10.3: The In-Out Decision

Minnesota Guidelines Vs. Massachusetts Actual Disposition

Minnesota Guidelines	Massachusetts Actual		
	Out	In	Total
Out	430	333	763
In	146	382	528
Total	576	715	1291

Table 10.4: The In-Out Decision

Massachusetts Guidelines Version 1 Vs. Massachusetts Guidelines Version 2

Massachusetts Guidelines Version 1	Massachusetts Guidelines Version 2		
	Out	In	Total
Out	179	131	310
In	0	981	981
Total	179	1112	1291

Table 10.5 The In-Out Decision

Massachusetts Guidelines Version 1 Vs. Massachusetts Actual Disposition

Massachusetts Guidelines Version 1	Massachusetts Actual		
	Out	In	Total
Out	217	93	310
In	359	622	981
Total	576	715	1291

Table 10.6: The In-Out Decision

Massachusetts Guidelines Version 2 Vs. Massachusetts Actual Disposition

Massachusetts Guidelines Version 2	Massachusetts Actual		
	Out	In	Total
Out	132	47	179
In	444	668	1112
Total	576	715	1291

differentiating among disposition types may be the cause of what appears here to be judicial leniency in comparison with the actual guidelines.

Distributions of Time Incarcerated

Figure 10.1 shows the distributions of expected length of incarceration for the entire sample under each of the four schemes. "Expected length" means the amount of time that would be served if release occurred on the presumptive parole date (Massachusetts) or if the guidelines sentence were imposed and not increased as a result of the loss of good time for institutional infractions (Minnesota). That is, the four frequency distributions have been made as comparable as we could make them, in view of the differences between the states' methods for setting actual incarceration time.

Insert Figure 10.1 here

It is clear from Figure 10.1 that the major differences in outcomes produced by the three sets of guidelines occur among the least severe incarceration times -- from zero through three years. There is a striking similarity among the Minnesota and Massachusetts (both versions) outcomes for times greater than three years. Actual sentences in Massachusetts in the three-years-plus range, particularly in the three-through-six year range, occurred much less frequently than they would have under any of the three sets of guidelines.

Of course, the bulk of cases fall into the 0-3 year interval, where the distributions differed considerably. The two versions of Massachusetts guidelines produced relatively flat distributions in this range, while the Minnesota guidelines and Massachusetts actual sentences were both heavily skewed, with "outs" and up-to-one-year terms much more frequent than 1-3 year terms.

In summary, examination of the distributions of time served under the four schemes reinforces the impression given earlier by the "in-out" proportions: extent of confinement ranges from lowest under the Minnesota guidelines, through actual practice in Massachusetts, through Version 1 and, most extensive, Version 2 of the Massachusetts guidelines. The implications these differences have for projected prison populations are examined later in this chapter.

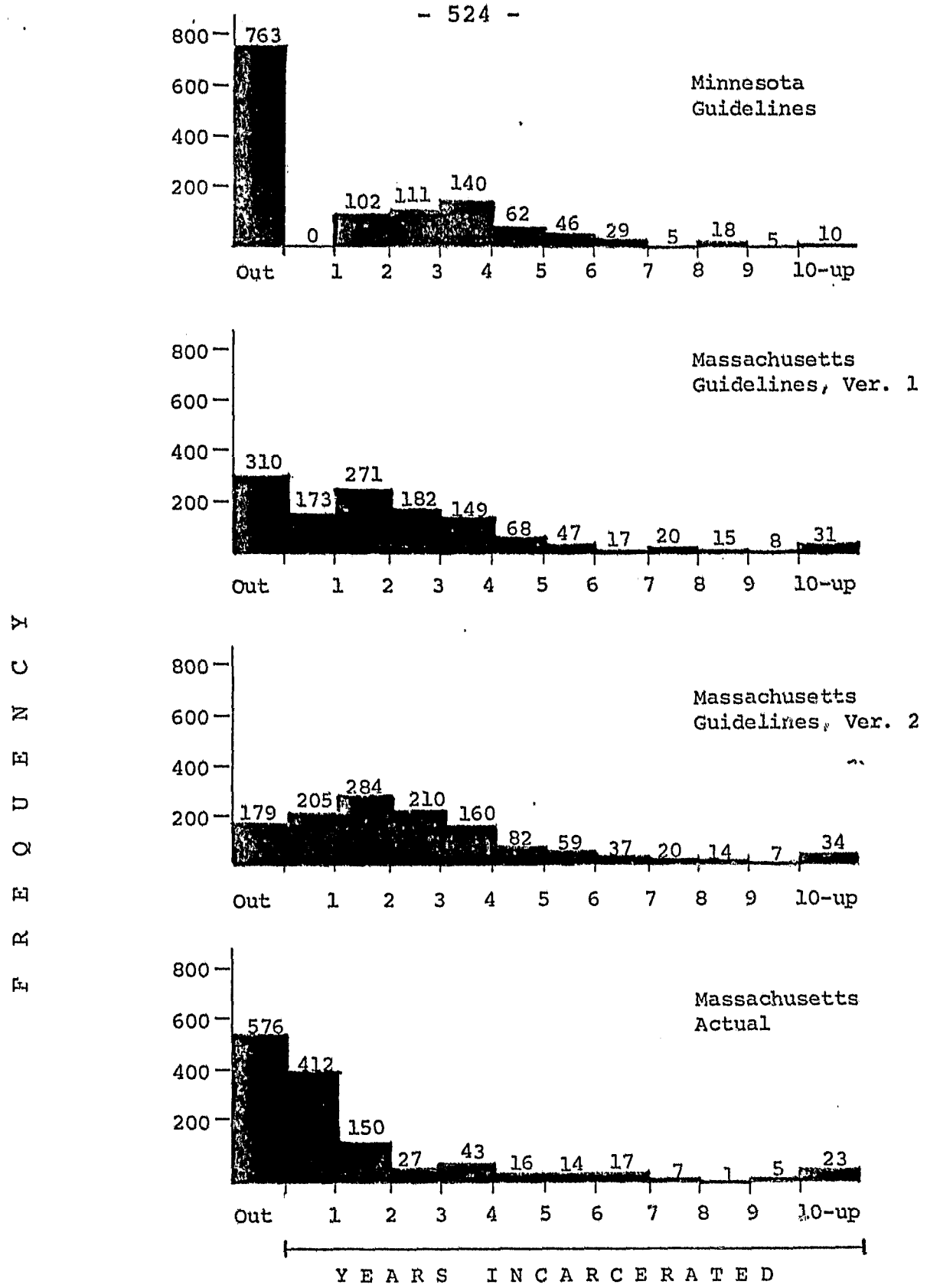


Figure 10.1: Distributions of Expected Lengths of Incarceration

(N=1291 in each distribution)

Sentence Characteristics According to Conviction Offense Type

Table 10.7 presents summaries of the sentencing outcomes for each of the three guidelines schemes and the actual Massachusetts sentences, broken down by Uniform Crime Reports (UCR) Type I offense categories. Given the results reported thus far, i.e., that there do exist major, overall differences among the sentences generated under the four schemes, we now address the question of whether some offense types contribute disproportionately to these differences.

Insert Table 10.7 here

First of all, it simplifies things somewhat to note that in Table 10.7 the two versions of the Massachusetts guidelines are, statistic-for-statistic, quite similar to each other, so we will refer to them together as "Massachusetts guidelines" in this discussion. While their prescriptions do differ more than slightly for three categories (Burglary, Larceny and "Other"), the differences occur only with respect to the proportion of persons incarcerated. There are no exceptions to the rule that Version 2 incarcerates more people than Version 1 (as was also shown earlier in Table 10.4), and only a couple of unimpressive exceptions to the rule that Version 2 incarcerates people for slightly longer periods of time than Version 1.

Differences in incarceration rates between the Massachusetts and Minnesota guidelines vary considerably as a function of current offense types. For the two most serious Type I offenses (murder and rape), the incarceration rates are uniformly high, but for all the other offense categories, moderate-to-extreme differences are evident in Table 10.7. The robbery, larceny and "other" categories exhibit intermediate incarceration rate differences, but the truly striking contrasts occur in the aggravated assault (20 percent vs. 94 percent) and burglary (14 percent vs. 74 percent) groups. In all instances, the Massachusetts guidelines produce the higher incarceration rates.

The results presented up to this point suggest that Massachusetts' guidelines' structure prescribes harsher punishments than does Minnesota's, and generally this is true. However, there is an additional important qualification which must be noted. The mean and median "months served" figures in Table 10.7 exclude those persons for whom an "out" decision was prescribed by the respective guidelines. (And it should be recalled than an "out" sentence for Minnesota can still involve up to one year in jail, while many of those same cases

Table 10.7: Sentence Characteristics by Sentencing Scheme and UCR Offense Type

ROBBERY (N=297)	Minn. Guidelines	Mass. Ver. 1	Mass. Ver. 2	Mass. Actual
Percent Incarcerated	77%	92%	99%	75%
Mean Months Served*	44.3	48.4	46.9	27.0
Median*	40.2	37.0	38.0	18.0
Range*	24-97	7-591	6-591	2-192
Standard Dev.*	21.6	47.4	44.8	28.0

RAPE (N=63)	Minn. Guidelines	Mass. Vers. 1	Mass. Ver. 2	Mass. Actual
Percent Incarcerated	92%	92%	98%	71%
Mean Months Served*	61.8	38.7	44.1	63.3
Median*	52.2	21.4	26.4	30.0
Range*	30-132	8-245	8-255	1-180
Standard Dev.*	26.8	43.0	46.3	63.5

AGGRAVATED ASSAULT (N=270)	Minn. Guidelines	Mass. Ver. 1	Mass. Ver. 2	Mass. Actual
Percent Incarcerated	20%	94%	96%	50%
Mean Months Served*	48.5	42.2	44.3	17.7
Median*	42.8	32.3	34.4	8.2
Range*	22-132	6-387	6-374	1-224
Standard Dev.*	20.1	43.0	42.4	33.2

BURGLARY (N=132)	Minn. Guidelines	Mass. Ver. 1	Mass. Ver. 2	Mass. Actual
Percent Incarcerated	14%	74%	96%	55%
Mean Months Served*	35.0	21.8	23.9	13.1
Median*	33.8	16.2	19.1	6.4
Range*	26-65	7-96	6-135	1-80
Standard*	10.2	16.5	19.9	14.8

*All summary statistics except "Percent Incarcerated" are based only on those cases for which an "In" decision was prescribed (for the three sets of guidelines) or actually imposed (actual Massachusetts dispositions).

Table 10.7 Continued: Sentence Characteristics by Sentencing Scheme and UCR Offense Type

LARCENY (N=52)**	Minn. Guidelines	Mass. Ver. 1	Mass. Ver. 2	Mass. Actual
Percent Incarcerated	15%	50%	60%	44%
Mean Months Served*	28.9	24.7	27.6	8.3
Median*	27.0	19.5	22.3	5.6
Range*	22-41	6-62	6-75	1-64
Standard Dev.*	6.0	15.5	20.4	13.0

MURDER(N=17)	Minn. Guidelines	Mass. Ver. 1	Mass. Ver. 2	Mass. Actual
Percent Incarcerated	100%	100%	100%	88%
Mean Months Served*	78.8	65.4	66.8	75.5
Median*	44.3	51.5	51.5	72.0
Range*	43-243	46-158	51-159	6-180
Standard Dev.*	55.2	33.6	29.6	62.1

ALL OTHER OFFENSES (N=478)	Minn. Guidelines	Mass. Ver. 1	Mass. Ver. 2	Mass. Actual
Percent Incarcerated	25%	56%	71%	45%
Mean Months Served*	41.0	24.4	25.6	17.3
Median*	34.3	16.2	16.0	9.1
Range*	22-132	6-143	6-144	1-180
Standard Dev.*	19.2	23.3	24.4	26.9

*All summary statistics except "Percent Incarcerated" are based only on those cases for which an "In" decision was prescribed (for the three sets of guidelines) or actually imposed (actual Massachusetts dispositions).

**Includes auto larceny.

are considered "in" dispositions under the Massachusetts guidelines.) That is, given that a person would be incarcerated, the statistics in Table 10.7 summarize the length of the incarceration. It can be seen that for all of the offense categories except robbery and murder, those persons who do get incarcerated under the Minnesota guidelines can expect to spend more time in confinement than those incarcerated under the Massachusetts guidelines. In particular, the median times served under Minnesota guidelines in the rape, burglary and "other" categories are roughly twice those under Massachusetts' guidelines, within the same categories.

To summarize, the two states' guidelines are fairly similar in their prescriptions for murder and robbery convictions. For rape, the incarceration rates are similar, but Minnesota's guidelines call for substantially longer prison terms. For most other offenses, Minnesota's guidelines tend to incarcerate fewer numbers of persons, but for longer periods of time.

Finally, we turn to the actual Massachusetts dispositions. There seem to be two fairly consistent patterns exhibited: first, for all offense types except murder and rape, actual incarceration rates fell somewhere between the low and high extremes exhibited, respectively, by the Minnesota and Massachusetts guidelines. Second, again excepting murder and rape, the sentences given to those who were in fact incarcerated were considerably lower than those prescribed by either state's guidelines. At the extreme, median times served under Minnesota's guidelines for aggravated assault, burglary and larceny were all approximately five times greater than the corresponding medians for actual Massachusetts sentences. While the discrepancies between Massachusetts actual practice and Massachusetts guidelines were not as great, they were still considerable, particularly in the three aforementioned categories. (Of course, our earlier qualification about the application, in this analysis, of the guidelines to plea cases should be kept in mind as it may have led to over-estimation of the severity of the guidelines.)

Correlates of Grossly Discrepant Sentences

Since each offender in our sample was, via simulation, processed through both Minnesota and Massachusetts guidelines, it was possible to identify those offenders whose sentences would have differed substantially under the alternative sentencing schemes. There are, of course, two ways this might happen: a severe sentence in Minnesota could translate to a lenient one in Massachusetts; or vice-versa. In this section,

we approach the task of comparing guidelines in the two states by sketching the characteristics of those cases that exhibited the most pronounced differences. In view of the similarity of outcomes for Versions 1 and 2 of the Massachusetts guidelines, only the more recent Version 2 was used.

Two subsamples, each comprising approximately 5 percent of the total sample (N=63 and N=65) were identified. The criterion for their selection was that they exhibit the most extreme discrepancies between Minnesota and Massachusetts guidelines sentences. The major characteristics of the two groups are contrasted in Table 10.8. For the present purpose, "out" sentences were coded as zero months of incarceration, although as things turned out, both groups consisted entirely of persons who would have been incarcerated under either set of guidelines.

Insert Table 10.8 about here

The two subsamples differed considerably in several respects. Of course, by virtue of the way they were selected, they differed considerably in their Minnesota and Massachusetts guidelines sentences: the median differences in expected incarceration time were over three years in the "Minnesota More Severe" group, and over seven years in the "Massachusetts More Severe" group. Comparisons of the remaining characteristics in Table 10.8 reveal some of the correlates of these discrepancies.

First of all, there are three indicators of criminal history (Minnesota priors category, Massachusetts priors index, and number of prior felony convictions), and since the three are in total agreement with one another, it is easiest to begin our discussion with the role of prior offenses. All three of the indicators demonstrate that the offenders who would have been punished more severely in Minnesota tended to exhibit much lengthier criminal histories than those punished more severely in Massachusetts; sixty percent of the latter group, in fact, had no prior felony convictions at all.

Moreover, within the Minnesota-more-severe group, the median values of number of prior felonies (2.5) and Minnesota prior offenses category (4.4) suggest that a sizeable contribution to the prior offense rating was made by criminal history considerations over and above prior adult felonies. Recall that the Minnesota prior offense rating scheme assigns one point for each prior adult felony, but that additional points may be assigned for serious juvenile adjudications, prior misdemeanors, and custody status at the time the present offense was committed (e.g., while on probation).

Table 10.8: Comparison of Characteristics in the Two Subsamples Exhibiting the Greatest Discrepancies Between Minnesota and Massachusetts Guidelines Sentences

Characteristic	Subsample: Minnesota More Severe	Subsample: Massachusetts More Severe
Subsample Size	63	65
Median Minnesota Sentence (Months)	75.6	26.3
Median Massachusetts Sentence	27.0	104.9
Median Sentence Difference	38.6	84.3
Median Minnesota Current Offense Seriousness Category	7.4	6.4
Median Massachusetts Current Offense Seriousness Index	4.4	8.4
Median Minnesota Prior Offense Category	4.4	1.0
Median Massachusetts Prior Offense Index	5.9	1.2
Median No. Current Convictions	1.6	3.4
Median No. Prior Felonies	2.5	0.3
Three most Frequent Conviction Offenses (N)	Rape (26) Armed Robbery (22) Distribution of Heroin (3)	Assault & Battery with Dangerous Weapon (24) Armed Robbery (15) Assault & Battery (5)

Next, seriousness of current offense is summarized by three numerical indicators (Minnesota current offense seriousness category, Massachusetts current offense seriousness index, and number of current convictions), and since these three do not agree completely, their interpretation is a bit more involved; the disagreement they exhibit, however, appears to be resolvable.

The Minnesota index would suggest that the two subsamples differ only slightly with respect to current offense seriousness (medians of 7.4 and 6.4), with the more serious current offenses appearing in the Minnesota-more-severe subsample, while the Massachusetts index reflects a substantial difference in current offense seriousness (medians of 4.4 and 8.4), with the more serious offenses appearing in the Massachusetts-more-severe subsample. Resolution of the apparent disagreement can be attained by referring to the third indicator (number of current convictions), along with the difference in the ways of two states' seriousness indices are calculated.

Specifically, Minnesota's guidelines provide that only the most serious current conviction offense shall be used in the determination of the guidelines sentence, while Massachusetts takes account of all current conviction offenses, by summing the seriousness values for each, to produce its seriousness index. As shown in Table 10.8, the median number of current convictions for the Minnesota-more-severe subsample was less than half of that for the Massachusetts-more-severe one. It thus appears that the different methods for calculating offense seriousness do make a difference in outcomes, at least for these two extreme subsamples.

The most frequent conviction offenses also differed between the two groups: rape was the modal offense type in the Minnesota-more-severe subsample, while the offense of assault and battery with a dangerous weapon predominated in the Massachusetts-more-severe group. Armed robbery was the second most frequent offense in both groups; none of the remaining offense types in either group occurred with any notable frequency.

In summary of this section, we must first emphasize that the results here, unlike the rest of the present chapter, are based upon only 10 percent of our offender sample: those persons for whom the Minnesota and Massachusetts guidelines would have produced the most disparate sentences. Our purpose here has been to give a flavor of the kinds of discrepancies that can result, and their correlates, rather than to carry out a comprehensive, multivariate statistical analysis. The

same objective might have been achieved by presenting detailed descriptions of some of the individual cases from the two subsamples.

That being said, the offender who would fare worst under Minnesota's guidelines, as compared to Massachusetts', is one who is convicted on a single rape charge, committed while on probation or parole, and who exhibits a lengthy criminal history, probably including one or more serious juvenile adjudications. He could expect to spend something over six years incarcerated in Minnesota, as compared to about two and a quarter years in Massachusetts.

The profile of the offender who would, under guidelines, be treated much more severly in Massachusetts than in Minnesota includes multiple (3 or 4) current convictions for assault and battery with a dangerous weapon, and a relatively "clean" criminal history (zero, or at most one, prior felony convictions). This person would be incarcerated for about eight and two-thirds years in Massachusetts, as compared to a little over two years in Minnesota, according to the two states' guidelines.

Projected Prison Populations Under Alternative Sentencing Schemes

Sentencing policies do, of course, affect prison populations. The greater the proportion of persons incarcerated, and the greater the time for which they are incarcerated, the greater the burden on correctional institutions. In this section we present some projected prison populations based upon our simulated sentencing outcomes under Minnesota and Massachusetts guidelines, as well as upon the actual sentences given in Massachusetts, for the offenders in our sample.

Certainly, there are a great many factors other than sentencing policies which influence the level of prison populations. These include policies governing probation and parole violators (what proportion are incarcerated or reincarcerated, and for how long?); policies (formal or informal) governing prosecutorial decisions and plea negotiation; shifts in the demographic makeup of the general population, which may affect the incidence of crime; and there are no doubt many others. Moreover, it is likely that sentencing (and other) policies themselves would change in response to changes in institutional populations. For example, judges in states with severely overcrowded facilities might reserve incarceration for only the most serious offenders.

Projecting prison populations, in other words, can be an exceedingly complex endeavor, and we do not purport here to have assessed and taken into account the factors mentioned above. Rather, we have adopted a very simple projection model, which portrays the "pure" effects of the four sentencing schemes on institutional populations.

The model assumes that the types of crimes for which persons are convicted, their frequency of occurrence, and the sentences given to those persons, do not change from year to year. That is, we have generated the projections by using our sample of 1,291 offenders repeatedly as the "new", incoming group of felony cases for each year. The results are shown in Figure 10.2, which reports prison population rates (per 1000 felony convictions, per year) at the end of each of ten projection years. Populations of zero have been assumed for the beginning of the first projection year.

Insert Figure 10.2 here

It is evident, and perhaps not too surprising in view of the results already reported, that the two versions of the Massachusetts guidelines yield the highest projected prison populations. It might not have been so evident from earlier results, however, that actual practice in Massachusetts would produce the relatively low population levels shown in Figure 10.2. (We, of course, noted that we might have expected actual practice to be less severe than the guidelines as a consequence of the plea vs. trial distinction, though we did not project <u>how much</u> less severe.) It can also be seen that the Minnesota guidelines yield intermediate population levels. There is certainly a considerable amount of variation among the four schemes.

Under the present assumptions, prison populations necessarily reach a state of equilibrium (i.e., the same number of persons are released as are admitted each year) as soon as the projection period exceeds the incarceration time for the longest-incarcerated offender in the sample. Since only a very few sentences (under any of the four schemes) resulted in more than ten years of incarceration, the ten-year time frame given in Figure 10.2 is sufficiently long to permit meaningful comparisons among the four schemes. The actual asymptotic rates per 1000 felony convictions per year were: 1807 and 2079 (Massachusetts Versions 1 and 2, respectively), 1332 (Minnesota), and 698 (Massachusetts actual).

It is important to remember here that Figure 10.2 is expressed in terms of population <u>rates</u>; direct comparisons

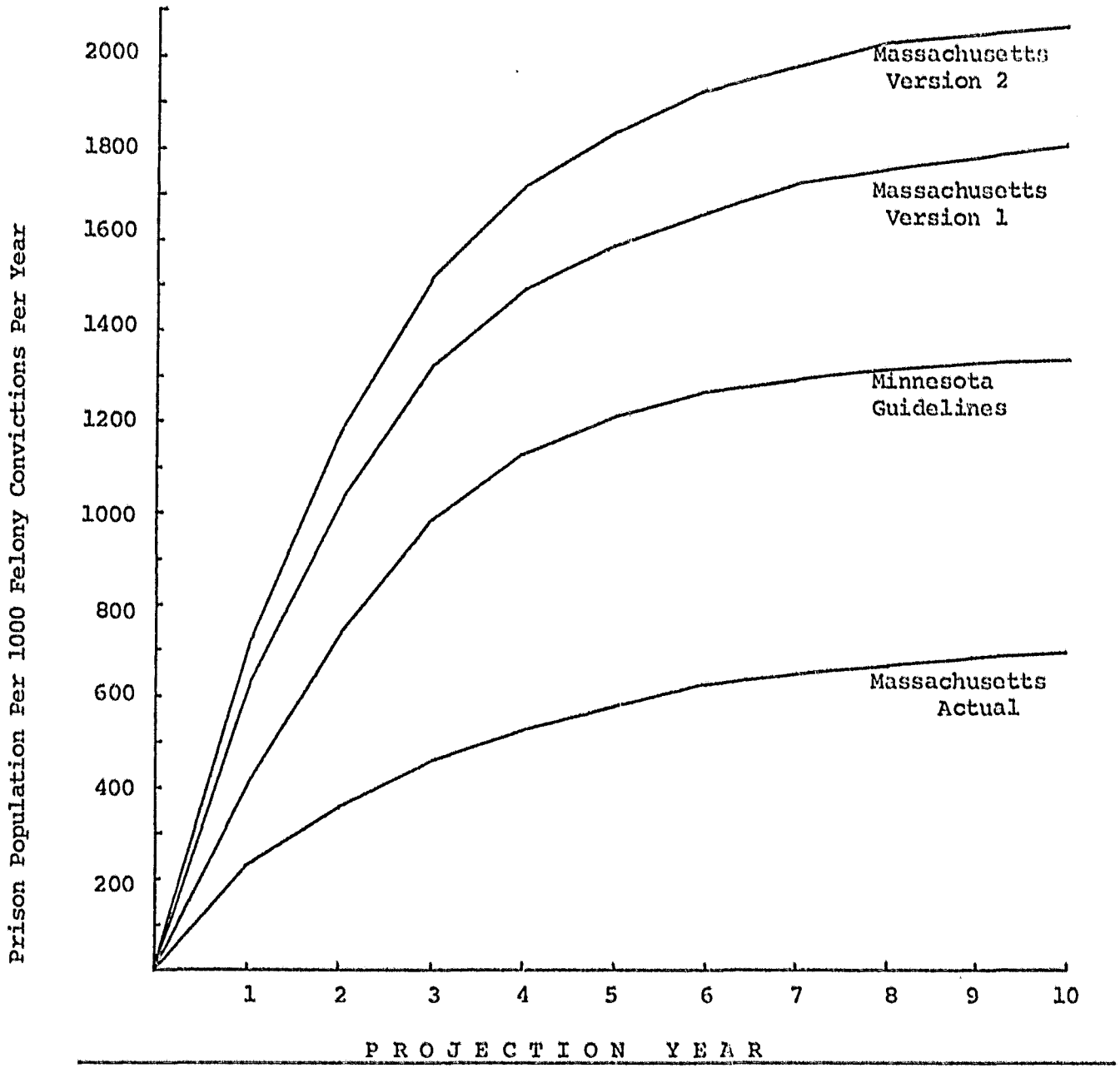


Figure 10.2: Projected Prison Population Rates Under Four Sentencing Schemes

among the four schemes in regard to actual numbers of persons incarcerated in Massachusetts vs. Minnesota are neither intended nor warranted. There are, first of all, obvious differences in size between the general populations, numbers of felony convictions, and prison capacities of the two states, which would have to be taken into account in any comparisons of numbers of incarcerated persons. Second, as suggested earlier, the offenders in our sample, having been selected exclusively from Massachusetts Superior Court cases, probably overrepresent the more serious crime types in both Massachusetts and Minnesota, vis-a-vis the populations of all felony offenders in the two states. In expressing the projections as population rates, we intend to emphasize that if the guidelines under consideration were applicable only to some (known) proportion of all felony offenders, the rates reported here could simply be multiplied by that proportion, yielding a population rate descriptive of only the more serious offenders. Independent estimates of incarceration levels attributable to the less serious offenders could then be combined with the present results, perhaps along with assumptions about the factors we have held constant (shifts over time in crime rates, etc.), to produce projected levels of actual numbers of persons incarcerated.

Summary

Sentences given under the four schemes examined in this chapter were found to vary considerably, with respect to both the in-out decision and the expected length of incarceration among the "in" cases. Differences in overall sentencing patterns were most prominent in the zero-to-three year incarceration range. For most crime categories, the Minnesota guidelines tended to incarcerate far fewer people, but for longer periods of time, then the Massachusetts guidelines. Exceptions to this occurred for the most serious offense types: sentences for murder were comparable under the two states' guidelines, and sentences for rape were considerably more severe under the Minnesota scheme.

The most pronounced differences in the two states' guidelines sentences occurred in cases involving convictions for rape (Minnesota more severe) and for armed assault and battery (Massachusetts more severe). Other factors which seemed to contribute to grossly disparate sentences were prior record (which contributed substantially to the more severe Minnesota sentences) and the difference between the two states' methods for calculating current offense seriousness (multiple current convictions contributed substantially to the more severe Massachusetts sentences).

Finally, projected prison population rates were highest for the two versions of the Massachusetts guidelines, followed by the Minnesota guidelines, and finally, Massachusetts actual sentences. Projected Massachusetts (Version 2) population rates were almost three times as great as those for actual Massachusetts sentences.

Appendix 10A: The Determination of Offense Equivalence Classes in Minnesota and Massachusetts

This appendix describes a scheme which classifies criminal behaviors as defined in Massachusetts into equivalent offense categories, had the offense occured in Minnesota. Since the two states' criminal codes are dissimilar, a set of transpositional rules were needed. Here we outline those rules, the methods that were employed, and the equivalence classes which resulted.

Primary consideration was given to the legal description of an act. For example, both states criminalize the possession of heroin. And under the respective guidelines, one to one correspondence is found. Most offenses were transposed under this rule.

Secondary consideration was given to the definition of the act, i.e., the precise wording of the statute. This criterion was most frequently employed where Massachusetts had several separate offenses which were combined under Minnesota law. For example, Massachusetts separately defines the sale, manufacture, distribution, and possession with intent to distribute various drugs as criminal offenses. Minnesota combines these behaviors under the rubric of drug sales offenses.

The third transposition rule involved the classification of offenses not included in the Minnesota guidelines. Where a Massachusetts crime was not included under the Minnesota guidelines offense list, that crime was treated as a misdemeanor. This rule was found to have a negligible impact on the computation of the Massachusetts current offense score under the Minnesota guidelines structure. The Minnesota sentencing guidelines consider only the most serious current offense for sentencing purposes, and in almost every case where a Massachusetts offense had to be classified as a misdemeanor, another more serious current offense was also present. This rule, however, may have introduced some distortion into the assessment of the Minnesota prior record score in that an offense, which we were forced to treat as a misdemeanor using the above rule, may actually have been a felony in Minnesota, albeit one that was outside of the sentencing guidelines. Thus, the relative weight of such an offense would be underestimated in the criminal history score -- in other words, it would have been counted as a misdemeanor rather than a felony. (For a complete description of the offenses included in the Minnesota sentencing guidelines see the Report of the Minnesota Sentencing Guidelines project, 1979.)

Our fourth criterion examined the severity of the prescribed sanction vis a vis offense categories. For example, Massachusetts defines "burglary", "breaking and entering", and "entering" separately, while Minnesota's sentencing guidelines stipulate specific levels and types of burglary. Thus, Massachusetts' statutes for "breaking and entering" and "entering", with the exception of "breaking and entering with the intent to commit a misdemeanor"

and "entering with the intent to commit a misdemeanor", were classified as lesser burglary offenses in Minnesota.

The final transpositional rule dealt with missing data which could modify offense type. Here we assumed that, in lieu of positive evidence to the contrary, a modifier was not present. For example, a Massachusetts offense of assault and battery could be equated with Minnesota offenses of assault with great bodily harm (a lesser offense). Without positive evidence of great bodily harm, the offense was transposed into assault with substantial bodily harm in Minnesota.

The following table displays, for each of the offense codes present in the Massachusetts data set, the equivalent Minnesota offense code as determined by the procedures described above.

Massachusetts		Minnesota	
Code	Offense	Code	Offense
1000	Abandonment	2	Nonsupport
1040	Accosting	M	Unclassified
1100	Assault	43	Coercion-Threat of Bodily Harm
1101	Assault-Deadly Weapon	72	Assault-Deadly Weapon
1102	Armed Assault in Dwelling		
1103	Assault-Female Under 16	43	Coercion-Threat of Bodily Harm
1104	Indecent Assault-Child Under 14		
1106	Assault-Police Officer		
1108	Assault with Intent	*	Unclassified
1109	Premeditated Assault	43	Coercion-Threat of Bodily Harm
1110	Assault & Battery	91	Assault-Great Bodily Harm
		61	Assault-Substantial Bodily Harm
1111	Assault & Battery-Weapon	91	Assault-Great Bodily Harm
		72	Assault-Dangerous Weapon
1113	Assault & Battery-Female Under 16	91	Assault-Great Bodily Harm
1114	Indecent Assault & Battery-Child Under 14	61	Assault-Substantial Bodily Harm
1115	Assault & Battery-Guard		
1116	Assault & Battery-Police Officer		
1118	Assault & Battery with Intent	*	Unclassified
1130	Confining/Putting in Fear	43	Coercion-Threat of Bodily Harm
		29	False Imprisonment

1140	Defraud/Willful Misapplication	*	Unclassified
1200	Extortion	43	Coercion-Threat of Bodily Harm
		34	Coercion-Loss over $2500
		21	Coercion-Loss $300-$2500
1250	Kidnapping	97	Kidnapping-Great Bodily Harm
		90	Kidnapping-Release Victim Unsafe Plac
		82	Kidnapping-Release Victim Safe Place
1300	Voluntary Manslaughter	98	Manslaughter I-During Commission of a Crime
		92	Manslaughter II-Under Heat of Passion
1301	Involuntary Manslaughter	87	Manslaughter I-Forced to Kill
		85	Manslaughter II-Gross Negligence
		67	Manslaughter II-Hunting Accident
1302	Manslaughter-Motor Vehicle	85	Manslaughter II-Gross Negligence
1350	Mayhem	91	Assault-Great Bodily Harm
		72	Assault-Deadly Weapon
1400	First Degree Murder	101	First Degree Murder
1401	Second Degree Murder	100	Second Degree Murder
1450	Negligence	*	Unclassified
1460	Nonsupport	2	Nonsupport
1600	Rape	96,95	Criminal Sexual Conduct I-With
1601	Rape of Child	94,93	Sexual Penetration
1602	Rape of Child-Second Offense		OR
1603	Rape-Female Under 14	84,66	Criminal Sexual Conduct III-With
1605	Statutory Rape		Sexual Penetration
1700	Armed Robbery	88	Aggravated Robbery

1701	Unarmed Robbery	60	Simple Robbery
1702	Robbery while Masked/Disguised		
1750	Safe Blowing	34	Theft Related Offense over $2500
		23	Theft Related Offense-$150-$2500
3000	Arson	89	Arson I-Dwelling, Known Occupancy
		83	Arson I-Dwelling, Unknown Occupancy
3100	Breaking & Entering	M	Unclassified
3101	Breaking & Entering-Intent to Commit a Felony	56	Burglary
3102	Breaking & Entering-Intent to Commit a Misdemeanor	M	Unclassified
3103	Breaking & Entering-Night Time		
3104	Breaking & Entering-Night Time with Intent to Commit a Felony	56	Burglary
3105	Breaking & Entering-Night Time with Intent to Commit a Misdemeanor	M	Unclassified
3106	Breaking & Entering-Night Time with Larceny	56	Burglary
3107	Breaking & Entering-Night Time with Assault	77	Burglary with Weapon or Assault
3113	Breaking & Entering-Day Time		
3114	Breaking & Entering-Day Time with Intent to Commit a Felony	56	Burglary
3115	Breaking & Entering-Day Time with Intent to Commit a Misdemeanor	M	Unclassified
3116	Breaking & Entering-Day Time with Larceny	56	Burglary

3117	Breaking & Entering-Day Time with Assault	77	Burglary with Weapon or Assault
3119	Burglary	74	Burglary-Occupied Dwelling
3130	Burglary & Assault	77	Burglary with Weapon or Assault
3131	Burglary & Larceny	74 56	Burglary-Occupied Dwelling Burglary
3132	Possession of Burglary Tools	30	Possession of Burglary Tools
3150	Burning to Defraud	37	Arson III-Property
3151	Burning a Building	73	Arson II-Nondwelling
3152	Burning Building Contents	37	Arson III-Property
3153	Burn Dwelling	89 83	Arson I-Knows or Expects Occupancy Arson I-Unknown Occupancy
3155 3156	Burn Property Burn Insured Property	37	Arson III-Property
3200	False Statements via Credit Card	M	Unclassified
3201	Fraudulently Obtain Goods by Credit Card	50 27	Theft-Over $2500 Theft-$150-$2500
3203	Illegal Possession of Credit Cards	36 23	Theft Related-Over $2500 Theft Related-$150-$2500
3204	Unauthorized Use of Credit Cards	50 27	Theft-Over $2500 Theft-$150-$2500
3210	Concealment	M	Unclassified

3220	Damage Property	44	Damage Property-Risk Bodily Harm
3231	Damage Property-Massachusetts Correctional Institution	15	Damage Property
3250	Entering without Breaking		
3251	Entering without Breaking-Night Time	M	Unclassified
3300	False Report of Explosives		
		89	Arson I-Knows or Expects Occupancy
3301	Malicious Explosion	83	Arson I-Unknown Occupancy
3302	Throwing Explosives	73	Arson II-Nondwelling
		37	Arson III-Property
3310	False Pretense		
3320	False Entry	M	Unclassified
3350	Infernal Machine	65	Possession of Incendiary Device
3400	Larceny	50	Theft-Over $2500
3401	Larceny Over $100	27	Theft-$150-$2500
3402	Larceny Under $100	M	Unclassified
3403	Larceny from Person	58	Theft from Person
3420	Loans Violation		
3440	Sell Mortgaged Property	M	Unclassified
3500	Receiving Stolen Property	71	Receiving Stolen Property-Over $2500
		52	Receiving Stolen Property-$150-$2500
3540	Trespass	42	Dangerous Trespass
3560	Vandalism	15	Damage Property-$150-$2500
3580	Wages Violations	M	Unclassified
4000	Violation of Controlled Substance Act	22	Sale of Schedule IV Substance
		6	Possession of Schedule IV Substance

4100	Dispensing Controlled Substance without Registration	22	Sale of Schedule IV Substance
4200	Distribution of Class A Controlled Substance	80	Sale of Heroin
4201	Distribution of Class B Controlled Substance	69	Sale of Remaining Schedule I & II Narcotics
4202	Distribution of Class C Controlled Substance	75	Sale of Hallucinogens or PCP
4203	Distribution of Class D Controlled Substance	46	Sale of Remaining Schedule I, II, III Non-narcotics
4204	Distribution of Class E Controlled Substance	22	Sale of Schedule IV Substance
4205	Distribution of Controlled Substance		
4300	False Prescription for Controlled Substance		
4301	Utter False Prescription		
4302	False Prescription-Class A		
4303	False Prescription-Class B		
4304	False Prescription-Class C	5	Illegal Procurement of Controlled Substance
4305	False Prescription-Class D		
4306	False Prescription-Class E		
4307	False Prescription-Controlled Substance		
4350	Fictitious Registration Number to Obtain Controlled Substance		
4400	Give Prisoner Class A Controlled Substance	49	Bring Contraband into a Prison
4405	Give Prisoner Controlled Substance		
4450	Heroin Being Present	35	Possession of Heroin
4451	Controlled Substance Being Present	6	Possession of Controlled Substance
4460	Possession of Hypodermic Needle		
4470	Inhale Toxic Substance	M	Unclassified
4480	Adaptation of Instrument for Drug Use		

4500	Manufacture of Class A Controlled Substance	80	Sale of Heroin
4503	Manufacture of Class D Controlled Substance	46	Sale of Remaining Schedule I, II, III Non-narcotics
4505	Manufacture of Controlled Substance	22	Sale of Schedule IV Substance
4550	Possession of Class A Controlled Substance	35	Possession of Heroin
4551	Possession of Class B Controlled Substance	16	Possession of Remaining Schedule I, II, III Narcotics
4552	Possession of Class C Controlled Substance	18	Possession of Hallucinogens or PCP
4553	Possession of Class D Controlled Substance	10	Possession of Remaining Schedule I, II, III Non-narcotics
4554	Possession of Class E Controlled Substance	6	Possession of Schedule IV Substance
4555	Possession of Controlled Substance		
4600	Possession with Intent to Distribute Class A Controlled Substance	80	Sale of Heroin
4601	Possession with Intent to Distribute Class B Controlled Substance	69	Sale of Remaining Schedule I, II, III Narcotics
4602	Possession with Intent to Distribute Class C Controlled Substance	75	Sale of Hallucinogens or PCP
4603	Possession with Intent to Distribute Class D Controlled Substance	46	Sale of Remaining Schedule I, II, III Non-narcotics
4604	Possession with Intent to Distribute Class E Controlled Substance	22	Sale of Schedule IV Substance
4605	Possession with Intent to Distribute Controlled Substance		

4650	Steal Class A Controlled Substance		
4651	Steal Class B Controlled Substance		
4652	Steal Class C Controlled Substance	50	Theft-Over $2500
4653	Steal Class D Controlled Substance	27	Theft-$150-$2500
4654	Steal Class E Controlled Substance		
4655	Steal Controlled Substance		
4700	Unlawful Administration of Controlled Substance	55	Using Drugs to Facilitate a Crime
6063	Adultery		
6045	Affray		
6061	Unlawful Manufacture of Alcohol	M	Unclassified
6063	Alcohol-Fees, License, Violation		
6066	Annoyance		
6080	Unlawful Use of Chemical Substance	55	Using Drugs to Facilitate a Crime
6121	Common Nightwalker	12	Solicitation for Prostitution-Over 18
6130	Conflict of Interest		
6145	Contributing to the Delinquency of a Minor		
6147	Cruelty to Animals	M	Unclassified
6142	Contempt of Court		
6166	Disorderly Person		
6167	Disturbing Peace		
6168	Drinking in Public		
6190	Escape	78	Escape with Violence or from Felony Conviction
6191	Escape from Custody	28	Escape without Violence
6192	Escape from Massachusetts Correctional Institution	78	Escape with Violence or from Felony Conviction
6193	Escape from Massachusetts Correctional Institution-Framingham		
6222	Failure to Appear		
6200	False Alarm	M	Unclassified
6201	False Impersonation		
6220	Fireworks		

6225	Obstructing Fire	42	Dangerous Trespass
6230	Flight to Avoid Prosecution	M	Unclassified
6240	Fornication		Decriminalized
6255	Giving of Gratuity/Bribery		
6256	Offering Bribe	47	Bribery
6257	Receiving Bribe		
6265	Health Code Violation		
6266	Housing Code Violation		
6280	Indecent Exposure		
6320	Lewd & Lascivious Behavior	M	Unclassified
6321	Lewd & Lascivious Cohabitation		
6330	Lottery		
6355	Distribution of Obscene Literature		
6360	Obscenity		
6370	Obstructing Justice	13	Aiding Offender to Avoid Arrest
6400	Perjury	62	Perjury in a Felony Matter
6450	Deriving Support from Prostitution	54	Receiving Money from Prostitution-Female 16-18
		38	Receiving Money from Prostitution-Trick, Fraud, Deceit
6460	Procuring Prostitution	31	Patron of Prostitution-Female Under 18
6465	Prostitution	45	Solicitation of Prostitution by Force
		12	Solicitation of Prostitution-Female over 18
6500	Registering Bets	M	Unclassified
6510	Rescuing Prisoner	78	Escape with Violence or on Felony Conviction
		28	Escape without Violence-Non-felony

6520	Resisting		
6525	Sodomy		
6560	Telephone-graph-vision	M	Unclassified
6625	Unnatural Acts		
6640	Vagrancy		
2000	Counterfeiting		
2050	Passing Counterfeit Bills	39	Aggravated Forgery-Over $2500
2051	Possession of Counterfeit Bills	25	Aggravated Forgery-$150-$2500
2100	Forgery		
2101	Forgery by Check		
2102	Forgery by Credit Slip		
2103	Forgery by Endorsement	39	Aggravated Forgery-Over $2500
2104	Forgery of Invoice	25	Aggravated Forgery-$150-$2500
2105	Forgery of Promisory Note		
2107	Forgery, Uttering and Larceny		
2200	Uttering		
5000	Dangerous Weapon	40	Dangerous Weapon
5010	Firearms Act Violation		
5011	Carry Illegal Firearms		
5012	Possession of Illegal Firearms		
5013	Illegal Sales, Leasing, or Rental of Firearms	M	Unclassified
5020	Illegal Discharge of Firearms		
5030	Defaced Serial Number on Firearm Used in Crime		
7100	Possession of Master Key	30	Possession of Burglary Tools
7120	Operating without License		
7121	Operate after License Revoked	M	Unclassified
7130	Operate Stolen Vehicle after License is Revoked		
7150	Unauthorized Use of Motor Vehicle	11	Unauthorized Use of Motor Vehicle

Appendix Notes

[1]The above table includes only offense transpositions which were necessary for the classification of offenders in the sample. Where the offense did not meet the minimum monetary loss requirement, it was treated as an unclassified misdemeanor.

M Unclassified. The Massachusetts offense was not in the Minnesota guidelines list of offenses and was treated as a misdemeanor.

* Unclassified. The Massachusetts offense title did not correspond to any recognizable Massachusetts criminal code. These cases were dropped from the analysis.

Chapter 11: How to Evaluate Statewide Sentencing Guidelines

By the way of conclusion, we briefly present in this chapter what may be described as a general model for the evaluation of sentencing guidelines, especially those intended for use on a statewide basis. The trendy term "model" is perhaps a little misleading in this context;[1] what we mean by it is a set of questions which must be asked if the operation and impact of such sentencing guidelines are to be understood, and the appropriate research and analytical methods for answering these questions. Though we were not able, in this project, to tackle all of these questions ourselves, our research had addressed most of them (including, arguably, the most important ones); and it is now easy for us -- easier than it was at the beginning of our project, at any rate -- to spell out the ways in which the remaining questions will need to be answered by any future empirical research on this subject.

Sanctioning Rules Revisited

We begin by recalling that sentencing guidelines can be regarded as a species of sanctioning rule, in the sense explained in Chapters 2 and 3: that is, as a set of prescriptions making certain sentences (probation, say, or a prison term) permissible or mandatory if certain antecedent conditions described in the rules are found to be present in a particular case. Thus, for example, in Minnesota an offender is convicted of aggravated robbery (seriousness level VII in the Minnesota matrix); the offender's prior criminal history adds up to a score of 4. The guidelines dictate that the judge may impose a state prison sentence of any duration between 60 and 70 months, though there is a presumption that 65 months will be imposed in a "normal" or "nothing special" case of that kind. Further rules prescribe the kinds of "substantial and compelling" circumstances in which the judge may impose a sentence outside the matrix range; the extent of permitted departure on grounds of aggravation or mitigation may be limited by still further rules, as is the case with the Pennsylvania guidelines (though not those in Minnesota).

We think it important to insist that sentencing guidelines are a kind of sanctioning rules; they are meant to control judicial discretion in sentencing, and not merely to provide some possibly useful information about its use in the past. The claim that guidelines are "descriptive, not prescriptive", though possibly well-intentioned, is generally speaking just plain false; and since the utility of falsehood is questionable,[2] it seems to us that people ought to stop making that claim. We say that the claim is generally false, because it is of course logically possible that judges might be presented with a set of guidelines to which they paid no

attention; for all we know, it may be that judges in New Jersey (and some judges in Massachusetts) are doing just that with their respective states' guidelines. Some of the early writing of Gottfredson and Wilkins, and the subsequent promotion of the guidelines concept by Gelman, Kress, and Calpin (1977) and other writers, may perhaps have encouraged the view that guidelines are intended to be no more than an "informational tool". But if it is true, in some jurisdiction, that guidelines are intended to be no more than that, then the answer is clear: there is no need for anyone to pay any attention to them. We think it is not just a coincidence that the "descriptive, not prescriptive" rhetoric has been strongly associated, during the brief history of sentencing guidelines, with efforts to sell the concept to judges; where the guidelines have had a statutory mandate, as in Minnesota and (perhaps) Pennsylvania,[3] their prescriptive nature is much clearer. Two of the states which we have studied -- Massachusetts and Michigan -- may turn out to furnish interesting intermediate cases, if the guidelines in those states should come to have the legal force of rules of the Superior Court.[4] Briefly, what is at issue is the legal standing or authority of the guidelines; rules of court procedure typically do not have the force of statute law, but they are generally much more compelling than voluntary compliance, even if the latter is based on a (problematic) consensus among the judiciary.[5] At any rate, we will not mince words: the idea that sentencing guidelines are "descriptive, not prescriptive" is merely a sham.

It is, moreover, a sham that in the past few years has cost a fair amount of public money. For it is surely only by virtue of the misbegotten idea that sentencing guidelines must have a "descriptive" component that the empirical research on past sentencing practice which we have considered in this report could ever have been thought to be justified: and it is strongly arguable, in our view, that a good deal of the money spent on that research was simply wasted. We return to the questions surrounding that research in a moment. Before turning to those questions, however, we re-emphasize some points made in earlier chapters of this report, which illustrate the utility of the concept of sanctioning rules in the analysis of discretion in sentencing and its control. There are (at least) six points that are especially important here, where guidelines are concerned:

(1) How is discretion to be controlled? We have argued that sanctioning rules can be represented schematicaly as directives of the form "If conditions X are present, then sanctions Y are permissible". Let us assume that a judge charged with applying such a rule really wants to follow it (rather than regarding it as an "informational tool" or some other similarly amusing artifact found in an obscure corner of

the judicial robing room). Our judge's loyalty is thus not in doubt; but how is he to demonstrate it? What is he not supposed to do? The answer is that his "freedom" of choice is limited in two very different ways, by the sanctioning rules which ex hypothesi he is trying to follow:

(i) Sanctioning rules stipulate a set of antecedent conditions X, which must be realized in order to make some sanction(s) Y permissible. But how "tightly" are these X conditions defined by the rules? Are the types of offenses distinguished by the rules relatively distinct, and their definitions generally unambiguous? How precise are the rules for calculating some sort of "Prior Criminal History Score", if such a score is intrinsic to the guidelines? How many discrete categories -- e.g., matrix cells -- are provided by the sanctioning rules, and how unambiguously are those categories defined? (Contrast a rule which stipulated that anyone committing a "serious crime" should be shot, with a rule that anyone with a given birthdate, birthplace and Social Security number should be shot -- on sight.) Here, guidelines offer a real and important contribution to the control of discretion -- more strictly, to its consistent exercise. For the variable(s) relating to "Offense Seriousness" in most guidelines schemes permit much finer distinctions than are made by most criminal statutes, between offenses which ought to be treated differently by sentencing judges. The ten levels of "Offense Severity" recognized by the Minnesota guidelines, for example, illustrate a grouping-together of types of offenses some of which, in their defining statutes, were -- for historical or other reasons -- treated quite differently; conversely, and perhaps more importantly, the guidelines draw distinctions between types of offenses which, in terms of their statutory maxima, were lumped together. The three-dimensional matrix system devised by the Michigan researchers (Zalman et al., 1980) goes even further in this respect, distinguishing within categories of offenses between crimes which can be isolated on the basis of factors specific to the broad categories of crime in question. One may agree or disagree with the particular distinctions drawn by these or other states' guidelines; that is not the point. Rather, it is that sentencing guidelines -- nesting, as they do, within the broader frameworks of statutory definitions of crimes, permit judges to make finer distinctions, in a consistent manner, than the statutory definitions themselves make. At the same time, guidelines allow a reasonable compromise between adequate distinction, and insanely microscopic detail of the kind exemplified by penal codes such as Ferri's (1921), which we described in Chapter 2.

(ii) The second way in which guidelines may constrain judicial discretion in sentencing is through the width of the range of choice left to judges in the "nothing special" case.

The statewide guidelines studied by us show a wide variation in this respect: in the Minnesota guidelines (and in the original Pennsylvania ones), the range widths were commendably narrow; in Massachusetts, by contrast, the "plus or minus 50 percent" rule seems likely to impose almost no constraint on judges, since relatively few cases are likely to require special justification because of falling outside such a wide range. However, the fact that guidelines typically do provide ranges (e.g. of months of incarceration) for specified X conditions which are defined in relatively simple terms (e.g., combinations of offense and offender scores) is a strength of this kind of sanctioning rule. A permitted range (as distinct from a single point, or a small set of points) allows judges to take into account special but rare factors which, though not justifying a sentence outside the range, may nonetheless justify some variation within a given matrix cell or similar category. Obviously there must in practice be some tension between the objective of flexibility, and the objective of reduced variability. But the concept of guidelines permits a reasonable compromise between these two competing objectives; and this flexibility inherent in the concept of guidelines is one of its great strengths.

(2) The hierarchy of sanctioning rules. Whether they have the force of statute law (as in Minnesota) or of a Rule of Court (as may turn out to be the case in Michigan and Massachusetts), guidelines lie within the outer limits of maximum sentences provided by criminal statutes, and thus narrow the range of permissible discretion provided by those statutes. Even so, within the guidelines there is room for still further narrowing, e.g. by sanctioning rules adopted by a particular bench or by a particular judge. However, the concept of guidelines initially propounded by Gottfredson and Wilkins, and exemplified by some of the systems actually in use, has the advantage that the rules used to narrow discretion are explicitly and publicly stated, rather than being private or personal; this explicitness may help to promote consistency, by minimizing mistaken assessments of the facts of cases as well as by inhibiting downright lying (of the kind that e.g., a racist judge might engage in). In passing, we may note that this is not necessarily an argument in favor of statewide sentencing guidelines as distinct from regional (or "multi-jurisdictional") ones. On the contrary: even if the courts of (say) northern New Jersey were to adopt one set of guidelines, and those of southern New Jersey to adopt another, the fact remains that each set could operate to narrow the discretion provided by the state's Penal Code. Indeed, there seems no contradiction in having both statewide and regional guidelines, except that the latter would presumably have to be narrower than the former in the sense that they did not permit or require sanctions outside the scope of the statewide rules. The important point is that if

sentencing guidelines are seen as sanctioning rules operating within the same legal context as statute or common law, they can readily be brought within the ambit of the most important technique for insuring compliance with them: to wit, appellate review. As was noted in Chapter 9, the Minnesota guidelines provide for appellate review; so may those of Michigan and Massachusetts, though as we noted in Chapters 6-8 above the existing Massachusetts appeal procedure, which applies to sentences to Walpole State Prison, seems not to have operated to constrain the discretion of sentencers as well as it might do. As an example of what can be accomplished here, we may consider the rules devised by the English Court of Criminal Appeal (Thomas, 1970); while the exact effect of these rules on sentences is not known, it seems likely that they have increased both the rationality and the consistency of application of sentencing in England, since they were first introduced in 1907.[6] Moreover, appellate review (or something like it) can go a long way toward meeting one of the objectives which Gottfredson and Wilkins have consistently regarded as one of the most important which guidelines can bring about: this is an explicit statement of general policies, which once stated can be publicly considered, debated and perhaps made more rational (see, e.g., Gottfredson, Wilkins and Hoffman, 1978).

(3) The importance of ancillary rules. We have argued that an important strength of sentencing guidelines, as a species of sanctioning rule, is the combination of flexibility and reasoned constraint which they can provide. It is necessary to emphasize, however, that the concept originally propounded by Gottfredson and Wilkins lacked an important element exemplified by the guidelines later developed in Minnesota and Pennsylvania: this is a further set of rules which explicitly lists the kinds of facts which may justify departures from the "normal" range, and those which may not be so used. The original Gottfredson-Wilkins concept, if taken literally, would allow judges to depart from the guidelines if they could cite any fact, however perverse or nonsensical, that could distinguish the instant case from the "typical" one of its kind. While flagrant abuses of the guidelines are perhaps unlikely to happen often in practice in most jurisdictions -- few judges today, we hope, would be so bold as to proclaim that they were imposing a heavier sentence on the basis of the defendant's race, or a lighter one on the basis of kinship -- it is nonetheless true that, in the original Gottfredson-Wilkins concept, the only element of constraint or control of judicial discretion was the (moral) obligation to cite some kind of reason if a sentence outside the guidelines was imposed. Appellate review of sentences may greatly strengthen the control exercised by guidelines; but so too may explicit rules, like Minnesota's or Pennsylvania's, relating to departures. At this point, we must concede, an

analysis of guidelines in terms of sanctioning rules may become a little cumbersome.[7] But that is because the reality of rule-controlled decision-making can be complex even when (as with sentencing guidelines) it is simple enough to be workable in practice; compare again the Ferri (1921) and Glueck (1928) sentencing schemes described in Chapter 2, which make the Minnesota and Pennsylvania rules look positively abecedarian.

(4) The "in-out" decision. One inherent weakness in the concept of sentencing guidelines concerns the so-called "in-out" decision. As we noted in Chapter 3, such a dichotomous outcome is not easy to control by any kind of sanctioning rules (except mandatory ones), for it is very difficult to follow a rule of the form "Do such-and-such in 33 percent of all cases of type X". This problem was not apparent when the concept of decision-making guidelines was originally developed by Gottfredson and Wilkins, since the parole decisions with which they were concerned were typically not choices between dichotomous outcomes. As we have seen, most of the statewide guidelines which we studied have handled this problem rather badly: in particular, the Michigan and Pennsylvania guidelines, which provide ranges of the form "0 - x" in some cells, provide very little direction, let alone control, where the important decision to incarcerate is concerned. But the Minnesota guidelines effectively hide the problem from view, since they are concerned only with state prison sentences, and completely ignore the fact that what is described as an "out" sentence may in fact involve up to a year in jail, in addition to the contingent liability of a much longer prison term if the offender violates some conditions of probation. Where the "in-out" decision is concerned, the flexibility which seems to us such an advantage of the concept of guidelines appears to vanish: the decision to incarcerate or not must be presumptive at best. Such a presumptive decision may indeed be guided by rules or principles, like those exemplified by the Minnesota and Pennsylvania ancillary rules; but the notion of a range of choice, such as applies to lengths of term in sentencing or parole guidelines, seems to have no application. Ironically, the New Jersey guidelines, which give information from which percentages incarcerated under previous sentencing practice can be calculated, seem to provide the most guidance to judges of any statewide guidelines which we have studied. But this is not very much by way of guidance; and as we argued in Chapter 4, it seems to be the only thing that the New Jersey guidelines researchers managed to get right.

(5) The variety of forms of guidelines. Since the original work of Gottfredson and Wilkins for the U.S. Board of Parole, most decision-making guidelines have been presented in the form of a two-dimensional (or, as in Michigan, a

three-dimensional) matrix. But, as we have seen in earlier chapters of this report, this is by no means necessary; and one advantage of the concept of sanctioning rules, in our opinion, is that it facilitates comparisons between techniques for the control of discretion which may superficially take many different forms. By providing a single conceptual framework within which not only guidelines, but also other forms of control of discretion, may be analyzed, we think that the concept of sanctioning rules can be a clarifying one -- and clarification is what the study of discretionary decision-making seems to us most to need, at this particular moment in its history.

(6) Towards a model of how judges really decide what sentences to impose. Our final claim for the utility of studying discretion in sentencing in terms of sanctioning rules is, at this point in our research, necessarily a modest one; but it is one in which we have some faith. We have noted, at several places in this report, that those who have carried out empirical research on past decision-making practice have not (whatever they may have thought) studied previous decision-making policies; nor did they have any clear or cogent theories about how judges make sentencing decisions. (The work of Wilkins (1976, 1978) may seem to contradict this -- since Wilkins is virtually the only researcher working in this area in the past decade who has tried in any way to explain how or why judges behave as they do. However, though we are much indebted to Wilkins' work in many areas, we feel that his work on theories of decision-making has to date contributed little if anything to his -- or anybody else's -- efforts to develop sentencing guidelines.[8])

What we suggest is that the concept of sanctioning rules provides the basis for a viable theory of judicial decision-making -- of the kind which we hoped to state and test empirically in the never-funded second phase of our research. We believe (with degrees of conviction that probably vary among us) that sentencing decisions typically involve the applications to fact-situations of rules of the form "If X, then Y is permissible"; and we believe that this hypothesis is the most fruitful basis for an empirical study of sentencing. At any rate, we would be prepared to start with this notion as a tentative working hypothesis -- meaning, in the social sciences, the kind of claim that one is prepared to defend until at least 50 years after one's death.

An analysis of sentencing decisions in terms of the application of rules of the form "If X, then Y" does not entail any particular objective of sentencing (such as deterrence, incapacitation or rehabilitation). Rather, it is through the application of such rules that any such objective gets turned into sentencing practice. In order to understand

a sentencing decision, it is necessary to understand the purpose or objective which it is consciously intended to accomplish; but it is also necessary to know the kinds of cases in which that objective is regarded as appropriate; the criteria by which such cases are to be identified (these are the antecedent or X conditions of the rules); and the consequences Y which are to be imposed in such cases in order to accomplish the desired objective or objectives.

Our general view, then, is that sentencing is basically the rational and purposive application of sanctioning rules to particular fact-situations. A detailed explication of this view must wait until another time;[9] but there is one point which should be emphasized, since it has been hopelessly misunderstood by many of those who have done research with a view to developing sentencing guidelines. To say that sentencing is basically "rational" in the sense just sketched is not to say that observed patterns of sentencing cannot have causes, in the sense of factors which may play no part in the decision-maker's beliefs and intentions. On the contrary: a variety of "unconscious" factors -- such as racial or social-class bias, stereotypical misperception, and mistaken beliefs of a variety of kinds -- may be important determinants of sentencing practice. Thus, any attempt to explain past sentencing practice should not a priori exclude from consideration such things as racial discrimination or regional variations in perceived seriousness (whether or not these last are well-founded). Yet, as we have seen in earlier chapters of this report, this is precisely what most of those who have tried to develop guidelines "models" have done. In summary: we believe that what we have called sanctioning rules are a central element in any adequate account of how judges do in fact decide what sentences to impose in particular cases. We believe that this is also likely to be true for prosecutors, defense counsel, police officers, and others in the criminal justice system who are required to make discretionary decisions.

The conceptual framework which we have developed in the course of the research described in this report draws a distinction which we believe to be very important to the study of the control of discretionary decision-making. This is the distinction between (1) the range of permitted variation in the structure of a set of sanctioning rules, and the ways in which that variation is distributed across different types of cases; and (2) the behavior of decision-makers operating within a set of sanctioning rules, and the ways in which those rules constrain (or fail to constrain) their behavior. Both of these topics are extremely important to the evaluation of sentencing guidelines. Though we did not have the opportunity, in this project, to address the second topic, we believe that our work has made a number of contributions to

the understanding of the first topic, e.g., through the analytical techniques which we applied to the Minnesota, Pennsylvania and Michigan guidelines in Chapter 9, and the cross-jurisdictional analysis illustrated in Chapter 10. We also have a set of prescriptions for answering questions falling under the second topic, which we will describe in a moment; though we had no chance to apply these in our own research, we think they are worth spelling out, since to our knowledge nobody else who has so far considered that topic seems to have understood its broader points -- let alone its finer ones.

Empirical Research and Sentencing Guidelines

We hope that it is by now clear that sentencing guidelines, as sanctioning rules, do not necessarily have to have an "empirical basis" of any kind whatsoever: had Pilate sought the answer to the question "What is Justice?", he would not have needed the services of Gelman, Kress and Calpin (1979). There are indeed arguments in favor of the view that an understanding of past sentencing policy can be useful, if changes in that policy are contemplated; one has to start somewhere, and it is no doubt better to have accurate information about past policies than to settle for mere rumor or speculation about them. The fact remains that, in every case known to us, the research done as a preliminary to constructing guidelines was not designed so that it could produce any information whatsoever about past sentencing policies -- and that most of that research has produced only minimal information about past sentencing practice, a good deal of which may have been less than useful.

It might, for example, be useful to study sentencing in the past, in order to estimate and document the extent of disparity, and to gain some idea of its causes. Have sentences in the past really been influenced by morally iniquitous factors such as race or social class (all else being equal)? Has there been inordinate variation in the use of incarceration across different regions of a state -- and, if so, what was the reason for this? Within a single jurisdiction, are there particular judges whose sentences are grossly deviant from those of their colleagues -- for example, judges whose sentences were wildly severe because of impending senility or an excessive faith in the "rehabilitative ideal" (or both)? Being in the business of social research, we are not about to denigrate the use of social research to answer questions such as these, which may inform legislators and other policy-makers in many ways. We think it is clear, however, that as a preliminary to formulating sentencing guidelines, such research is scientifically unnecessary. Such empirical research on past sentencing practice as has been done, in the jurisdictions with which we are most familiar,

seems primarily to have been done for political reasons, of a rather different kind -- in particular, to sell the concept of guidelines to the judiciary and to legislators. That some of this research seems to have been politically counter-productive -- as, for example, in Pennsylvania -- does not, to us, excuse the fact of its having been done at all.

Those who find this assessment too harsh will, we hope, at least agree with us that if empirical research aimed at modelling past sentencing practice is to be done at all, then it ought to be done well. Our review of the research which has so far been done by guidelines developers, however, suggests that this has not always -- to put it mildly -- been the case. We do not refer to the mundane problems of sampling, data collection, coding, missing data, etc., which can be troublesome in many kinds of social research -- especially in the field of criminal justice. We refer instead to the analytical and statistical blunders which characterize so much of this research. To begin with, none of the guidelines projects which we have reviewed began with even the most rudimentary theory of judicial decision-making; explanatory variables were thus chosen, and defined, on the basis of the data that happened to be available, e.g. in pre-sentence reports. Those data, gathered retrospectively, were mostly of doubtful reliability and validity; that apart, they may often have failed to include variables that were in fact important to sentencing decisions (e.g. recommendations of defense counsel, in Massachusetts and perhaps elsewhere). In some cases outcome variables were measured inappropriately, e.g. by scoring non-incarcerated cases as zero, before regressing sentence length across all cases. The models fitted to the data were often strikingly crude in form: is there any reason to think that linear additive functions, with no interactions, are the only kind which should be considered? In some cases (e.g. Zalman et al., 1979) variables were retained in modelling equations because they had attained statistical significance, even though they made no substantive sense at all. Little attention was paid to problems of distributional shape, e.g. the extreme skewness of sentence lengths in most jurisdictions; and the estimation procedures used were typically not theoretically the best.[10] The Massachusetts data (see Chapter 8) show clearly that different models may be needed for cases tried and cases disposed of by guilty pleas; no other researchers seem to have considered this. Having decided that morally inappropriate variables like race should not be included in guidelines, most researchers failed to include these variables in their descriptive modelling equations; yet as we saw in Chapter 8, this can lead to mis-estimation of the effects of other variables, and (at least in theory) to the incorporation into guidelines of the very biases which it has wanted to avoid. Few if any guidelines researchers seem to have paid any

attention at all to the residuals from their modelling equations; yet as we showed in our re-analyses of the Massachusetts data, the identification and exclusion of extreme residual outliers can have drastic effects on the statistical models used to describe the bulk of cases. Finally, few (if any) guidelines researchers carried out a statistical validation of their regression models. This is an especially ludicrous error in the case of New Jersey, where there were more than 11,000 cases at hand; as we showed in Chapter 4, there is good reason to believe that a good many of the "findings" contained in the New Jersey guidelines are the results of nothing but chance variation.

In summary: empirical research aimed at developing guidelines began with the very simple demonstration by Gottfredson and Wilkins that "seriousness of offense" and "prior record" were important determinants of sentencing and parole release decisions (see Chapter 3). Some of those who have subsequently done research with a view to developing guidelines have stopped after achieving approximately the same modest objective; the Minnesota researchers, for example, seem to us to have had little interest in multivariate modelling of past practice. But most other researchers in this area do seem to have hankered after more elaborate descriptions of past practice; in pursuit of this dubious goal, they seem collectively to have made almost every mistake in the book.

Not that it mattered all that much. In every case except New Jersey, the empirical "findings" produced by guidelines researchers were substantially modified, on more or less explicit grounds of policy, when the guidelines themselves were created. There is nothing wrong with this, of course; but it does reinforce doubts about the wisdom of doing the descriptive research in the first place, especially when -- as in Michigan, and apparently in Pennsylvania -- the prescriptions in the guidelines seem so different from what the antecedent description would suggest. The translation of descriptive models to prescriptive guidelines can <u>never</u> be automatic: if data may occasionally "speak for themselves," they never do so in the imperative mood. Suppose that there is evidence that in the past, blacks received heavier sentences than whites, across the board (cf. Chapter 8). Guidelines might remedy this evil; but it must still be decided whether in future all offenders should be punished as whites formerly were, or whether all should receive the heavier sentences formerly given to blacks; and that is a policy question which the data by themselves cannot answer.

We must emphasize, once again, that we are not arguing that empirical research has no role whatsoever in the process of developing sentencing guidelines. The question is, what <u>kind</u> of research should be done? One important example of

what should be done is the estimation, by the Minnesota researchers, of the impacts on prison population of various guidelines options[11]; another is illustrated by the cross-jurisdictional study described in Chapter 10. A further area in which empirical research badly needs to be done concerns the measurement of the (perceived) seriousness of crimes, and severity of penalties. We believe that the research described in Chapter 5 makes a modest contribution in this area; but much work obviously remains to be done, in particular to explore the general assumption that there is a clear social consensus on those matters. Finally, we still believe that there is a need for research on how judges (and others) actually make decisions, under the constraints imposed by guidelines or other sanctioning rules.... Perhaps our Phase II proposal got lost in the mail.

Most of the sentencing guidelines developed to date have taken the form of a matrix -- two-dimensional in Minnesota and Pennsylvania, three-dimensional in Michigan, n -dimensional in New Jersey. But of course this is by no means necessary; it is, one might say, a matter of style rather than of substance. On this score, we feel that there is something to be said for the rather different form illustrated by the Massachusetts guidelines,[12] which yield an "expected" sentence by means of a fairly direct (though policy-modified) translation of the unstandardized regression coefficients on which the guidelines are supposedly based. Whether or not the guidelines are presented in matrix form, some calculation (whether by judges, probation officers, or prosecutors and defense counsel) of offense and offender scores is likely to be necessary; and the Massachusetts-style guidelines seem to us to have the potential for simplifying this calculation somewhat, which is plainly a good thing. (Could anything be more cumbersome than the New Jersey guidelines -- which come in a sort of matrix form?)

The difference in form -- Massachusetts style versus Minnesota, say -- may have other consequences in practice. A Minnesota judge who consults that state's matrix and is told that the defendant before him falls into cell (i,j) of that matrix can literally see, immediately, the relation of that crime and defendant to a variety of others. A Massachusetts judge using that state's guidelines would not have that immediate visual contrast (though in time he would no doubt acquire a grasp of the relationships in question, if he did not already have it[13]). He would, however, have a quite clear indication of the specific elements which accounted, in varying proportions, for the "expected" or guideline sentence in the case before him. (Whether this difference of form makes any difference to sentencing practice is, of course, an empirical question -- on which somebody, some day, may have data.) It seems likely to us that, in theory, the

Massachusetts guidelines permit somewhat finer discriminations between closely similar offenders than do other states' matrices; this is because they may give "expected" or guidelines sentences which differ somewhat, for cases which in most matrices would be lumped into the same cell. Whether such differences, or nuances, are important in principle is a question we cannot answer. Unfortunately,their practical importance in the Massachusetts guidelines seems likely to be minimal, view of the plus-or-minus 50 percent range which those guidelines provide, before a case must be treated as a "departure" requiring some kind of special justification.

Guidelines in their Social and Institutional Context

Sentencing guidelines have generally been regarded as a technique for affecting the behavior of one decision-maker, namely the judge. But it can not be said too often that sentencing is a complex process, which typically takes place within the context of an inter-related set of institutions whose actions may influence the judge's final choice in a number of ways. It is important, therefore, to look closely at the roles and behavior of the other actors in the criminal justice system, in relation to sentencing guidelines.

To begin with, if the guidelines are meant to be "empirically based" in the sense of mirroring past sentencing practice, it will almost certainly be necessary to study the behavior of prosecutors, defense counsel, probation officers and others, in order to find out what that practice is: case-level outcome data are not enough. As we saw in earlier chapters, there was considerable variation in a number of respects, between the four Massachusetts counties which we observed, e.g. in the ways in which district attorneys' policies were formulated and carried out by their assistants; such between-county variation might well be even more marked in other states, since in many respects Massachusetts has a relatively homogeneous criminal justice system.[14] Anyone aiming to develop guidelines ought to get this kind of understanding of the system before beginning to collect data on past practice. While that understanding might not amount to a full-fledged theory of sentencing, it would at least help to insure that data are collected on all relevant variables (e.g. defense counsel's recommendations, which were omitted from the Massachusetts data). Again, there is evidence that probation officers' recommendations may have some influence on sentences in some jurisdictions[15]; this did not seem to us to be the case in Massachusetts, though we might have thought that it was if we had not observed the ways in which sentencing was actually done there and talked to the various participants in the process.

There is a much more important reason to consider the other actors in the sentencing system, however; and this reason applies to those who make up guidelines, as well as those who try to evaluate them. By virtue of their roles in the sentencing system, not only judges but also prosecutors, defenders and probation officers may crucially affect the way in which guidelines are implemented; and unless all of these parties are fully informed about the guidelines, and their support enlisted, the guidelines may not have their desired effects.

We found a striking contrast in this respect between Massachusetts and Minnesota (see Chapters 7 and 9 above). It is, of course, too early to say anything about the day-to-day operation of the guidelines in either state.[16] Our observations of the ways in which the two states' guidelines were introduced to those who would have to use them, however, leave us pessimistic about the situation in Massachusetts. In every respect we can think of, the Minnesota procedures were superior. The Minnesota Commission, of course, had a legislative mandate, and the Commission itself contained representatives of prosecution, defense and corrections; by contrast, the Massachusetts guidelines were devised on behalf of the judiciary, and received a certain amount of legislative opposition. The Minnesota Commission had held numerous public hearings around the state in order to explain the guidelines, before they were introduced; by contrast, there seemed to us to have been only minimal information about the Massachusetts guidelines disseminated outside the Suffolk County Courthouse, before the guidelines themselves were introduced.

The Minnesota Commission's presentation of its guidelines to the judges' in a one-day meeting, was exemplary in the clarity of the description given by staff, and the written materials provided. The Massachusetts guidelines, by contrast, were presented in about two hours, during a meeting devoted mainly to other business; written materials consisted only of the guidelines themselves, and the description of these (by one of the judges on the committee responsible for them) verged on the surreal, and seemed to confuse and antagonize many of those present. The Minnesota Commission made excellent presentations of its product to prosecutors and public defenders; in Massachusetts, by contrast, no effort of any kind to do this was made. There is some reason to think that the Massachusetts judiciary believed, on constitutional grounds, that prosecutors, as members of the executive branch, had no business contributing to the formation of sentencing guidelines. As we saw in Chapter 7, the Massachusetts legislature had some rather different views on that topic, also on constitutional grounds. In any case, such an argument does not excuse (much less justify) the failure to inform prosecutors and defenders in Massachusetts, well in advance,

of the nature and purpose of the guidelines.[17] Finally, the Minnesota Commission continues to monitor the use of that state's guidelines, and may make revisions of them from time to time; at the very least, the Commission constitutes an effective means of making sentencing policy explicit and permitting its rational discussion. We were informed that arrangements were also being made to monitor the use of the Massachusetts guidelines; but we learned very little about these arrangements, and what we did learn did not inspire confidence.[18]

In summary: the process of introducing and implementing guidelines may be crucial to their subsequent operation; and it needs careful consideration in any evaluation of guidelines once they are in use.

How to Measure Compliance with Guidelines

Suppose that we had had the opportunity to see whether sentencing guidelines -- in Massachusetts, Minnesota or elsewhere -- "made a difference" to subsequent sentencing practice. How would we have done this?

Two preliminary points may be noted. First, to the extent that guidelines are meant merely to mirror prior practice, it may be that they will lead to no differences in sentencing; indeed, it would seem that they should not. In the case of policy-derived guidelines like Minnesota's, it is tolerably clear that these are supposed to reduce variability in prison terms (by squeezing together terms within matrix cells); they were also intended, according to the Commission, to bring about some changes in the use of imprisonment, and presumably to reduce such variability in sentencing as may have existed across the state.[19] But it is a curious fact that most of those who have written about "empirically derived" guidelines, from Gottfredson and Wilkins onward, have said very little about exactly what those guidelines were supposed to accomplish. One may guess that they are intended to reduce variability in sentencing, e.g. by restraining the occasional hanging judge; and in the case of statewide guidelines, it may be assumed that regional differences are supposed to be minimized. But there is little evidence available, to confirm or refute this guess.[20]

Second, compliance with guidelines (or other sanctioning rules) is not, strictly speaking, the same thing as behaving in a way that is not inconsistent with the rules. "Complying" is an intention-dependent concept; one cannot comply with an order inadvertently. Thus if we are to talk of compliance, evidence is needed to show that judges consciously prescribed sentences in accordance with the guidelines and any associated rules. Evidence of this kind may be obtained by asking judges

why they have sentenced in a particular way; better yet, it may be obtained by going out and observing them. Documentary evidence relevant to compliance may be available in cases involving departures from the guidelines, if there is a requirement to state written reasons for these; but there will usually not be that kind of evidence for cases which are not treated as departures, and aggregate data on sentencing before and after guidelines can be even more difficult to interpret. In their evaluation of the Denver and Philadelphia guidelines, Rich et al. (1980) make before-and-after comparisons, e.g. of incarceration rates, which they claim show something about compliance. But (as they admit at one point) they use the term "compliance" to mean nothing more than consistency with a predicted outcome; which means that they should not have used the term at all.

What are the best strategies for making before-and-after comparisons of data on sentencing outcomes -- assuming that more "direct" evidence is not available? Let us consider first the issue of lengths of terms for offenders incarcerated before and after the guidelines -- assuming, for the sake of simplicity, that we are looking at offenders that are similar, e.g. are in the same matrix cell. One possible outcome is illustrated in Figure 11.1, which presents two hypothetical frequency distributions of sentence lengths found before and after guidelines are introduced. (In passing, it is important to note that full frequency distributions should ideally be compared here; since changes in distributional shape can take a variety of forms, and since the distributions are likely to be markedly non-Gaussian, summary statistics of location and spread are likely to be very uninformative, if not downright misleading. [21])

Insert Figure 11.1 here

In Figure 11.1, it is evident by eyeball inspection that there has been an aggregate reduction in severity in the post-guideline period (the curve has shifted to the left somewhat); there has also been some reduction in variability. Yet plainly there are some cases in the post-guideline period which fall outside the prescribed normal range (indicated in the figure by the two vertical lines). Do these out-of-range cases show a failure to comply with the guidelines? The answer is "Not necessarily"; for in those cases given terms that are higher or lower than the "normal" range, there may have been one or more aggravating or mitigating circumstances present, which would fully justify the out-of-range sentence imposed. The only way to check on this possibility is to look at the facts of those cases, to see whether there are such mitigating or aggravating factors apparently present. Since

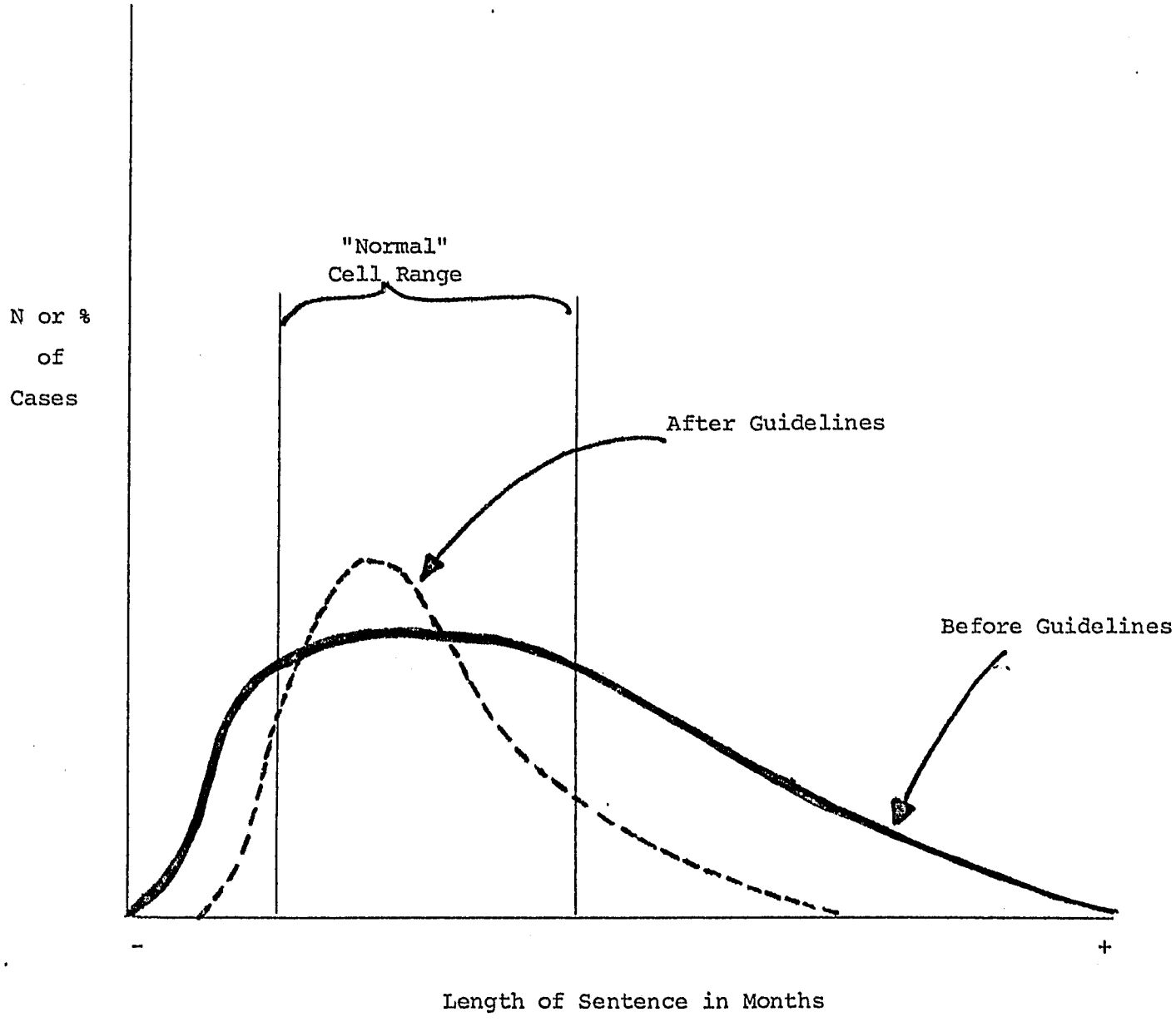


Figure 11.1: Hypothetical Frequency Distributions of Lengths of Prison Terms Before and After Guidelines

these are cases for which special justification is supposed (according to most conceptions of guidelines) to be not merely present but specifically indicated by the sentencing judge, it should not be too difficult in practice to evaluate their status as correct "departures". Of course this task will probably be much easier in those jurisdictions (like Minnesota) in which rules accompanying the guidelines describe, in some specificity and detail, the kinds of factors which may and may not be used to justify out-of-range terms -- even if such lists are not intended to be treated as exhaustive of the possible grounds for departure. Even if there is no such list, and if insufficient information is vouchsafed by the sentencing judge to make clear the justifying facts, these ought to be easy to spot in at least some cases (compare the analyses of outliers in the Massachusetts data in Chapter 8 above). Any case in which no obvious feature justifying a departure is present should surely be classified as a prima facie instance of a failure to comply with the guidelines.

Precisely the same kind of case-by-case analysis can be done, of course, for pre-guideline cases; cf. again our discussion of extreme cases in the Massacusetts data, in Chapter 8. Ex hypothesi, however, there will be no explicit judicial statements concerning the grounds for such extreme deviations from the "normal" sentence imposed in similar cases before the guidelines came into force; so the justification of very high or very low sentences in pre-guideline cases may have to be much more tentative, and it may be difficult indeed, in practice, to show that such extreme terms constituted disparity rather than inadequate evidence concerning the grounds for the judge's decision.

The argument, then, is that the existence of cases falling outside a "normal" guidelines range may not indicate non-compliance with the guidelines; since (because of special features justifying mitigation or aggravation) those cases may be precisely the ones that should receive sentences involving a "departure". Unfortunately, however, this argument cuts both ways: for it may well be that, after the implementation of sentencing guidelines, there are cases which should be treated as "departures", but which for some reason are not dealt with in that way by the sentencing judge; and in practice these cases may be very much more difficult to spot, especially since there will usually be no written statement of reasons for giving a sentence within the "normal" range. The only solution, in our opinion, is to scan those cases falling with the "normal" range after implementation of the guidelines, and to note the numbers in which ostensibly mitigating or aggravating factors are present but apparently did not weigh heavily enough, in the judge's opinion, to justify a higher or lower term.

In the case of guidelines like those of Massachusetts, which do not take the form of a matrix, the analysis of compliance is, at least in principle, somewhat simpler; it can proceed along exactly the lines we illustrated in Chapter 8, in our analysis of the Massachusetts construction data. First, calculate the "expected" or guidelines sentences for cases dealt with under the guidelines; then calculate residuals (observed sentence minus expected), plotting these in various ways -- in particular, against expected values -- and examine the cases showing substantial residuals to see whether the data on them give any reason to believe that a heavier or lighter sentence could reasonably have been felt by the judge to be justified. The "expected" terms can be banded by plus-or-minus 50 percent intervals (or whatever interval may be used in the jurisdiction in question), as in some of the analyses in Chapter 8; for those outside that range, there should presumably be some statements of reasons available, as well as data on the cases themselves. (Again, of course, cases with small residuals should strictly speaking also be examined, to see whether or not there are factors present which should have justified an exceptionally heavy or light sentence but did not.)

Substantially similar logic applies to the "in-out" decision, assuming that a prior probability of incarceration (e.g. percentage incarcerated in a matrix cell, in construction data) can be calculated for post-guidelines cases. A case given an "in" sentence though predicted to be "out" (or the reverse), is not eo ipso evidence of non-compliance with the guidelines; for there may be special features present in the case, which justify a sentence different from what the guidelines would normally prescribe. Again, however, cases differing from what would be expected under the guidelines must be analyzed to see if anything which can reasonably be construed as a factor justifying "departure" can be identified.

In this case, however, a problem arises concerning the predicted status (as "in" or "out") of cases dealt with after the introduction of the guidelines. Some of those who have dealt with this problem have rather arbitrarily assumed that a case with a 51 percent probability of incarceration (in the guidelines, as estimated from construction data) should be counted as predicted "in", and that those with prior probabilities of 50 percent or less should be treated as predicted "out" cases (see, e.g., Rich et al., 1980; cf. Zimmerman and Blumstein, 1979). This seems to us too rigid for a number of reasons, not the least of which has to do with sampling error. Instead, it seems better to plot the predicted probabilities of incarceration of post-guideline "in" and "out" cases separately, as in Figure 11.2. This figure shows the kind of plot which might reasonably be

expected to be found: those cases actually given "in" sentences after the guidelines have probability distributions with a high mode but skewed to the left, perhaps going well beyond the .50 point; those actually given "out" sentences, by contrast, should have generally low probabilities skewed to the right, again crossing the .50 probability point for some cases. The 50 percent cutoff used by Rich et al. (1980) and others corresponds to a situation where the tails of the two distributions are each lopped off exactly at the .50 probability line on the abcissa, with cases across that line (in the left-hand or right-hand direction, as the case may be) being treated as "non-compliance" cases. Yet plainly in the situation envisaged in Figure 11.2, there are a number of what are probably very similar cases (those in the shaded area, under both curves); these may be cases that -- in baseball argot -- are just too close to call, given the relatively crude predictions of the "in-out" decision which most guidelines necessarily make. Again, searching along the whole of the two distributions for the presence and/or absence of possible mitigating or aggravating factors seems the only way to estimate the extent of compliance which has actually occurred since the guidelines were introduced.

Insert Figure 11.2 here

A final problem arises when the two decisions -- "in-out" and length of term if incarcerated -- are combined. Suppose that, following the introduction of the guidelines, there is an increase in the proportion of convicted offenders incarcerated, so that some of those who would (under previous practice) have been placed on probation are now being incarcerated, but are given lighter sentences -- possibly even "departures" in the downward direction, but possibly not -- than the guidelines prescribe for those who would previously have been incarcerated.[22] Is this a case of non-compliance with the guidelines? Not necessarily. Is it a departure from the guidelines -- and if so, in which direction (up or down)?

In order to tackle questions of this kind, it seems that some sort of composite measure is needed, which will reflect both the shift which may take place in probabilities of incarceration, and the variation in sentence lengths which may accompany it, after guidelines (or some other change in the structure of sanctioning rules) takes place.[23] It is not difficult to contrive such composite measures. For example, suppose we define

$$E_g = P_g * S_g$$

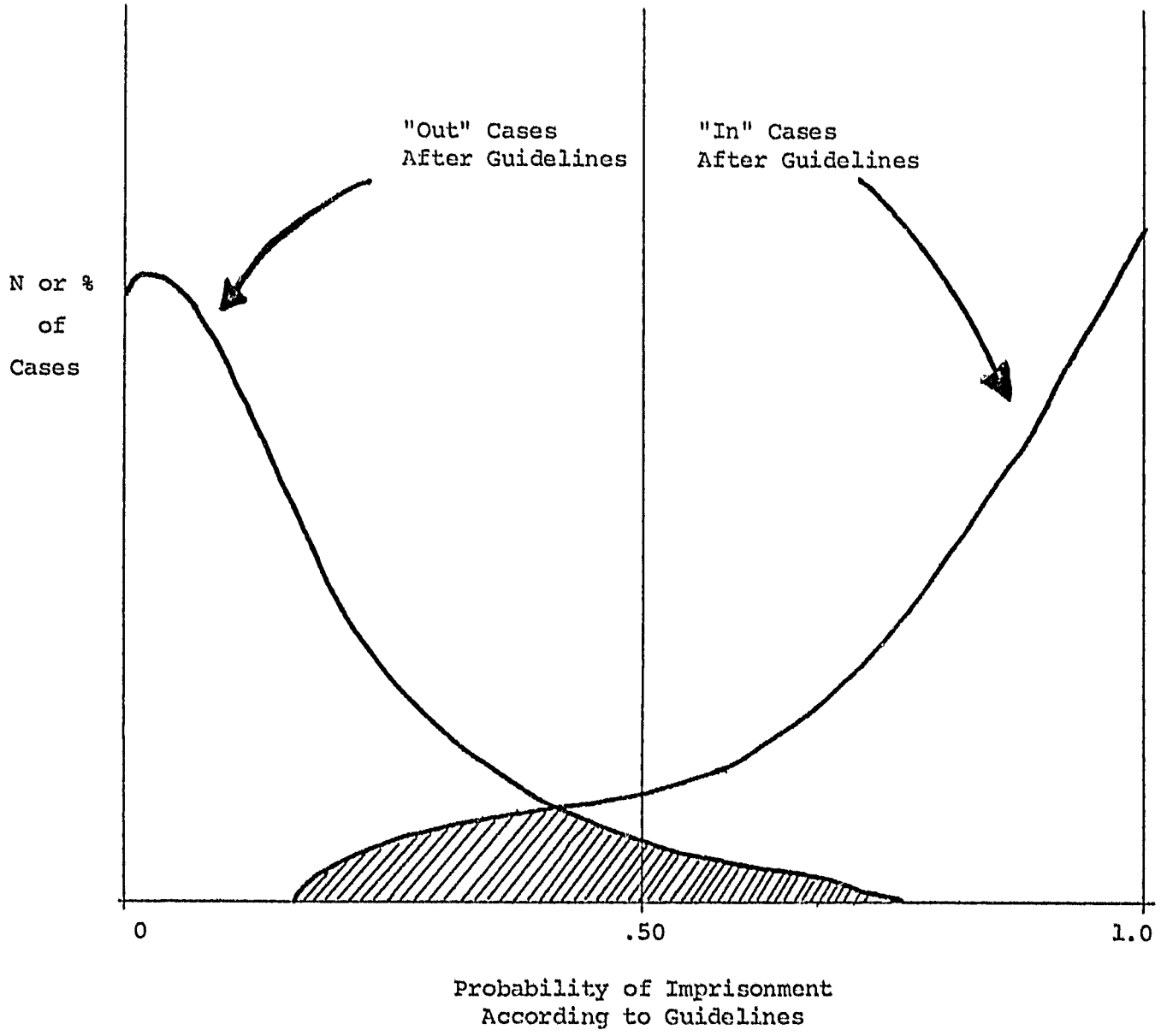


Figure 11.2: Hypothetical Probability Distributions for "In" and "Out" Cases Sentenced After Guidelines

$$R_o = P_o * S_o$$

where Pg is the offender's probability of incarceration according to the guidelines, and Sg is the sentence (mid-range or predicted, as the case may be) which the guidelines prescribe if such an offender is incarcerated -- or was, in the pre-guideline period; Po is 1.0 if the (post-guidelines) offender is in fact incarcerated, and zero if not; and So is the length of incarceration imposed (which will of course be definitionally zero if Po = 0). In words, Eg stands for the offender's expected sentence under the guidelines (which may or may not be what he might have expected in fact, depending on how "empirically based" the guidelines in question are); it equals the probability that he would have been imprisoned, times the sentence he would have been likely to receive if imprisoned.

Ro stands for the sentence actually imposed. To give this term concrete meaning, consider one of the cells in the Minnesota guideline matrix -- say, the cell occupied by those convicted of Simple Robbery (Commission severity level V) with a Criminal History Score of 3 (MSGC, 1980). This cell is just below the presumptive "in-out" line in the Minnesota matrix, and so according to the guidelines is presumptively an "in" cell; yet as we saw in Chapter 9, only about 85 percent of those in this category before the guidelines were actually imprisoned, and we may imagine that about 15 percent might similarly be given "out" dispositions after the guidelines came into force. The prescribed presumptive term in that cell is 30 months; so Eg= .85 * 30 = 25.5 months reflecting an 85 percent chance of a 30-month prison sentence. Suppose an offender falling into that cell after the guidelines come into force is not imprisoned; in his case Ro-Eg) equals -25.5. Suppose that a second offender falling into this cell is imprisoned, but for only 12 months instead of the prescribed 30; in this case Ro = 1.0 * 12 = 12 months, and the "residual" (Ro-Eg) equals only -13.5 months. The magnitudes of the two "residuals" indicate the relative leniency accorded to the two offenders in comparison with what the guidelines prescribe. The first offender, not imprisoned at all, has evidently gotten a better break; but the second offender, though imprisoned, has also had something of a break since he has received a lighter sentence than the guidelines prescribe for the "normal" case. We might reasonably expect more by way of mitigating factors in the first offender's case; if not, there would prima facie be some inconsistency in the practice of "departures" in cases of that kind.

We would emphasize that a situation like the one we have just described is not in the least improbable. Nor must it necessarily be a consequence, in a strict sense, of the introduction of guidelines (or some other new-fangled form of

sanctioning rule). On the contrary, there may be many causes of a composite change in sentencing practice like the one just described (a change in incarceration rates, plus a shift in prison terms for some or all of those incarcerated). It seems not at all unlikely that if offenders formerly placed on probation are now sent to prison, then they are likely to receive lighter prison terms than those who would previously have been sent to prison. Examining separately the two elements of that composite change can thus produce misleading results; the measure we have just suggested may to some extent help to avoid this. We must admit, however, that we have not yet had the opportunity to explore all of the numerical and conceptual consequences of this measure, across the range of cases to which it might be applied; our suggestion of it is therefore a little tentative.

There are, of course, many other issues which would need to be considered in research on compliance with sentencing guidelines. For example, if the guidelines are meant to be applied uniformly across the whole of a state, and if there has previously been evidence of considerable within-state variance in sentencing (controlling for regional variations in crime patterns, etc.), it would be important to see whether or not that variance had been appreciably reduced after the guidelines came into force. There is one topic of particular importance which should be monitored in any case: this is the charging and plea-bargaining practices of prosecutors (and, in some jurisdictions, defense counsel). We saw in Chapter 7 that one of the district attorneys in our four Massachusetts counties professed to be entirely happy with the statewide guidelines: having purloined his own copy of them, he had put an assistant to work calculating guideline sentences for his caseload, and intended (he said) to base his sentencing recommendations on what the guidelines prescribed. Others, however, may not be so content; and they may react by changing their charging practices after the guidelines come into force, say by throwing the book at defendants in the hope of tacking an extra page onto their criminal histories. The precise ways in which this may happen -- indeed, the extent to which it happens at all -- will depend on the forms which plea negotiations take in the particular jurisdiction. In principle, however, any such radical shift in prosecutorial behavior should be detectable by examining the distributions of cases across guideline categories, before and after the introduction of the guidelines. (In Minnesota, for instance, such a change would be suggested by a gradual drifting of cases out of the upper-left-hand corner of the matrix -- where, as we saw in Chapter 9, they were concentrated before the guidelines -- in a rightward and downward direction.)

Concluding Unscientific Postscript

Throughout this report, we have treated guidelines as a technique for controlling sentencing behavior. We did this because that is what they are: they are a way of getting an objective accomplished, and not an end in themselves. Considered as a technique, we are inclined to give them a rating in the vicinity of B plus to A minus -- depending on their structure, range widths, precision of ancillary rules, provision for appellate review, etc.

We are not oblivious, however, of the social and political context in which this technique first appeared -- and in which it continues, albeit uncertainly, to survive. Many, if not most, of those who have called for statewide sentencing guidelines in the past few years have had hidden agendas which they were pursuing: the forestalling of mandatory or presumptive sentencing, the prevention of legislative encroachment on judicial prerogatives, an increase in the severity of punishment, or whatever. All of this has taken place in the midst of a general disillusionment with "rehabilitation" as an objective of sentencing, and a corresponding return to older ideals of distributive justice -- ideals for which sentencing guidelines, as it happens, are well suited. This shift in the aims of sentencing may help to explain why guidelines as a technique have so far survived the peculiar combination of federally-funded hackwork, dismal politics and wastefully inept research to which they have been subjected over the past five years or so. But it may also be that the technique really deserves to have survived that baptism of fire; at any rate, we should like to think that this is so.

Notes to Chapter 11

[1] We use it here only because it was used in the solicitation and in our proposal: see above, Chapter 1. The term has some recognized and fairly clear uses in the evaluation literature, e.g. in relation to "process" as distinct from "outcome" studies. That distinction does not seem useful to us here.

[2] The uttering of falsehood may of course be useful, e.g. to politicians and the adulterous. But statements intended to convey information in virtue of being true are not helped by being false.

[3] As this chapter was being written, and well after the discussion of the earlier Pennsylvania guidelines in Chapter 9 had been printed in final form, we received yet another version of that state's guidelines, which is to be debated by the legislature, probably in November or December 1981. In view of the still-tentative status of these latest guidelines, we have not included an analysis of them in this report; every book has to end somewhere.

[4] Note, however, that (as described in Chapter 8 above) this seems likely to come about in Massachusetts only with the approval of the legislature. The situation in Michigan is not known to us.

[5] And if there is such a clear judicial consensus, who needs (empirically based) guidelines?

[6] So far as we are aware, however, no research on the effect of these rules on sentencing decisions in England has ever been done. As Thomas's book clearly shows, many of the most highly-articulated rules and principles laid down by the Court deal with obsolete measures such as preventive detention for habitual criminals, and for "professional" armed robbery; the generality of those rules and principles is thus unclear.

[7] See above, Chapter 3. It simplifies matters somewhat, as we pointed out there, to think of a set of rules, so to speak nested within each other, which the decision-maker is required to consider in succession in certain cases (e.g. a departure from the "normal" range).

[8] See above, Chapter 3; we remain unpersuaded that the "bifurcated decision" concept is correct.

[9] We hope to address this point in a future project dealing with crime control strategies.

[10] Though the effects of this in practice, e.g. the use of ordinary least-squares with dichotomous outcome variables, may not have been all that harmful. Cf. Zimmerman and Blumstein, 1979; Zalman et al., 1979; and contrast Rhodes, 1980.

[11] Though, as noted in Chapter 9, we have some doubts as to the conclusions reached by the Minnesota researchers, since their projections ran for only five years.

[12] And also by many parole prediction instruments, e.g. the one devised by Mannheim and Wilkins (1965) and the California Base Expectancy Scores.

[13] Whether he would have the same understanding is an empirical question. Recent work on graphical methods of data analysis (for a summary, see Wainer and Thissen, 1981) suggests that this might not be the case.

[14] E.g. judges to some extent travel "on circuit"; and there is a state-organized public defender system. Cf. Chapter 6 above.

[15] See, e.g., Carter and Wilkins (1965).

[16] Data on post-guideline sentences are being collected by the Minnesota Commission, to monitor the guidelines' use; it is unclear how far this monitoring will extend to other issues such as the incorporation of guidelines into plea negotiations. We had no opportunity to study this subject in Minnesota.

[17] We were informed by the Massachusetts researchers that copies of the guidelines had been sent to every District Attorney's office, and to the Massachusetts Defenders Committee. As noted in Chapter 7, however, the dissemination process was uneven to say the least.

[18] Copies of guideline forms are apparently being forwarded to a Boston office where they may be analyzed under the direction of a senior probation officer. So far as we are aware, however, there are no provisions for insuring compliance with this process, and the use to be made of these documents is unclear.

[19] See above, Chapter 9; and MSGC, 1979.

[20] Cf. the analysis by Rich et al. (1980) of between-judge variation in the Denver guidelines data.

[21] For a more detailed discussion of the measurement problems which may be encountered in dealing with sentencing, see Chapter 4 by Sparks in Messinger et al. (1981).

[22] This is, in fact, precisely what seems to have happened in California, since the passage of the Uniform Determinate Sentencing Act of 1976: more people are being sent to prison, but for shorter periods of time, on average. See Messinger et al. (1981), Chapter 19.

[23] Note that such a change need not be caused directly by the introduction of guidelines.

Bibliography

Advisory Council of Judges of the National Council on Crime and Delinquency. Model Sentencing Act. NJ: National Council on Crime and Delinquency (1963).

Advisory Committee on Sentencing and Review. Standards Relating to Appellate Review of Sentences. Peter Low, Reporter. (Draft approved, February, 1968).

Akman, D.D., Figlio, R., and Normandeau, A. Concerning the Measurement of Delinquency - A Rejoinder and Beyond. British Journal of Criminology, 1967, 7, 442-449.

Akman, D.D., and Normandeau, A. The Measurement of Crime and Delinquency in Canada: A Replication Study. British Journal of Criminology, 1967, 1, 129-149.

Akman, D.D., and Normandeau, A. Towards the Measurement of Crime and Delinquency in England. Criminal Law Review, 1968.

Akman, D.D., and Normandeau, A. Towards the Measurement of Criminality in Canada. A Replication Study. Acta Criminologica, 1968, 1, 135-241.

Akman, D.D., Normandeau, A., and Turner, S. Replication of a Delinquency and Crime Index in French Canada. Canadian Journal of Corrections, 1966, 8, 1-19.

Akman, D.D., Normandeau, A., and Turner, S. The Measurement of Delinquency in Canada. Journal of Criminal Law, Criminology, and Police Science, 1967, 58, 330-337.

Allen, F.A. The Borderland of Justice. Chicago: University of Chicago Press, 1964.

Alper, B.S., and Weiss, J.W. The Mandatory Sentence: Recipe for Retribution. Federal Probation, 1977, 41, 15.

Alschuler, A.W. Sentencing Reform and Prosecutorial Power: A Critique of Recent Proposals for Fixed and Presumptive Sentencing. In Determinate Sentencing: Reform or Regression. Washington, D.C.: U.S. Government Printing Office, 1977.

Alschuler, A.W. The Prosecutor's Role in Plea Bargaining. University of Chicago Law Review, 1968, 36, 50.

Annotated Laws of Massachusetts. Rochester, NY: The Lawyers Co-Operative Publishing Company, 1980.

American Bar Association. Comparative Analysis of Standards and Goals of the National Advisory Commission on Criminal Justice of the American Bar Association. Washington, D.C.: U.S. Government Printing Office, 1973.

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

American Bar Association. Section of Criminal Law, Program and Committee Reports. Report of the Committee on Sentencing, Probation, Prisons, and Parole. San Francisco, California: July, 1939.

American Bar Association. Standards relating to Sentencing Alternatives and Procedures. New York: Institute of Judicial Administration, 1968.

American Friends Service Committee. Struggle for Justice. New York: Hill and Wang, Inc., 1971.

American Judicature Society. Sentencing Patterns and Problems: An Annotated Bibliography. 1973.

American Law Institute. Model Penal Code. Philadelphia: American Law Institute, 1962.

And Now For Something Completely Different. American Bar Association Journal, 1977, 63, 907.

An Illustration of the Use of the Illinois Post-Conviction Hearing Act as a Procedural Supplement to Review of Allegedly Irregular Conviction. Georgetown Law Journal, 1953, 41, 282.

Appealability of a Criminal Sentence--Sentence Modified on Appeal. Rutgers Law Review, 1961, 16, 186.

Appellate Modification of Excessive Sentences. Iowa Law Review, 1960, 46, 159.

Appellate Review of Criminal Sentences. Notre Dame Law Review, 1960, 35, 298.

Appellate Review of Primary Sentencing Decisions: A Connecticut Study. Yale Law Journal, 1960, 69, 1453-1478.

Appellate Review of Sentences: A Survey. St. Louis University Law Journal, 1972, 17, 221.

Appellate Review of Sentences and the Need for a Reviewable Record. Duke Law Journal, 1973, 1357-1376.

Appellate Review of Sentencing: A New Dialogue. University of Colorado Law Review, 1973, 45, 209.

Appellate Review of Sentencing: A Symposium. Federal Rules Decisions, 1962, 32, 249.

Appellate Review of Sentencing Procedure. Yale Law Journal, 1964, 74, 379.

A Report by Justice. Criminal Appeals. London: Stevens and Sons, 1964.

As Definite Sentencing Begins, Illinois Still Debates Its Value. Judicature, 1978, 61(8), 384.

Ashe, S. P. A Warden's Views on Inequality in Sentences. Federal Probation, 1941, 5(1), 26-27.

Aspen, M.D. Our Vanishing Species: The Judge Who Judges. Judge's Journal, Spring 1978, 17, 20-24.

Aspen, M.E. Illinois' New Flat Sentencing Law. Corrections Magazine, 1978, 4(2), 72.

Attica-The Official Report of the New York State Special Commission on Attica. New York: Bantam, 1972.

Baab, G.W., and Furgeson, W.R. Texas Sentencing Practices: A Statistical Study. Texas Law Review, February 1967, 45(3), 471-503.

Bagley, J.J. Why Illinois Adopted Determinate Sentencing. Judicature, 1979, 62, 390.

Barry, D.M. The Effects of Compounding Disparities at Successive Stages of Criminal Processing. Paper presented at the Annual Meetings of the American Society of Criminology. Philadelphia, November 1979.

Barry, D.M., and Greer, A. Sentencing vs. Prosecutorial Discretion: The Application of a New Disparity Measure. Journal of Research in Crime and Delinquency, 1981. In press.

Bartolomeo, J. Judicial Reactions to Sentencing Guidelines. Washington, D.C.: U.S. Department of Justice, Federal Justice Research Program, May 1981.

Bartolomeo, J., Clancy, K., Richardson, D., and Berger, P. Sentence Decision Making: The Logic of Sentence Decisions and the Extent and Sources of Sentence Disparity. Washington, D.C.: U.S. Department of Justice, Federal Justice Research Program, May 1981.

Baude, P.L. Grounds for Relief Under 28 U.S.C. 2255: A Suggested Standard. American Criminal Law Quarterly, 1967, 5, 112.

Bazelon, D.L. Missed Opportunities in Sentencing Reform. Hofstra Law Review, Fall 1978, 7(1), 57-69.

Beccaria, C. Essay on Crimes and Punishments. New York: Bobbs-Merrill, 1963.

Beha, J.A., III. And Nobody Can Get You Out: The Impact of a Mandatory Prison Sentence for the Illegal Carrying of a Firearm on the Use of Firearms and on the Administration of Criminal Justice in Boston. Boston University Law Review, 1977, 57, 96-146, 289-333.

Bennet, J.V. The Sentence and Treatment of Offenders. The Annals, 1962, 399, 142.

Bennett, J.V. Countdown for Judicial Sentencing. Federal Probation, 1961, 27, 22.

Bernard, T.J. Individualization vs. Uniformity: The Base for Regulation in Criminal Justice. Federal Probation, December 1976, 40(4), 19-23.

Berk, R.A., and Rossi, P.H. Prison Reform and State Elites. Cambridge, MA: Ballinger Publishing Company, 1977.

Bishop, E.T. Judgement in a Sentence. California State Bar Journal, 1954, 29, 3.

Blalock, H.M., Jr. Social Statistics. Second Edition. New York: McGraw-Hill Book Company, 1972.

Blom-Cooper, L. The Parameters of Sentencing. Howard Journal of Penology and Crime Prevention, 1976, 15(6), 12.

Blumberg, A. Criminal Justice. Chicago: Quadrangle Books, 1967.

Blumberg, A. Law and Order: The Scales of Justice. Chicago: Aldine Publishing Company, 1970.

Blumstein, A. Seriousness Weights in an Index of Crime. American Sociological Review, 1974, 39, 854-864.

Bohnstedt, M. Overview of Sentencing and Parole Guidelines. Paper Presented at Annual Meeting of the Academy of Criminal Justice Sciences. Cincinnati, Ohio, March 15, 1979.

Brewster, L. Appellate Review of Sentences. Federal Rules Decisions, 1966, 40, 79.

Bridges, G.S., and Lisagore, N.S. Scaling Seriousness: An Evaluation of Magnitude and Category Scaling Techniques. Journal of Criminal Law and Criminology, 1975, 66, 215-221.

Brietel, C. Controls in Criminal Law Enforcement. University of Chicago Law Review, 1959-60, 27, 427-435.

Buchner, D. Scale of Sentence Severity. Journal of Criminal Law and Criminology, 1979, 70 (2), 182-187.

Bullock, H. A. Significance of the Racial Factor in the Length of Prison Sentences. Journal of Criminal Law, Criminology, and Police Science, 1961, 52(4), 411-417.

Burr, C.B. Appellate Review as a Means of Controlling Criminal Sentencing Discretion--A Workable Alternative University of Pittsburgh Law Review, 1972, 33, 1.

Byrne, W.M. Federal Sentencing Procedures: Need for Reform. Los Angeles Bar Bulletin, 1966-67, 42, 563-567.

Cain, R.H., and Ours, J.R. PROMIS For the Nonautomated or Semiautomated Office. Washington, D.C.: Institute for Law and Social Research, 1976.

Campbell, W. Developing Systematic Sentencing Procedures. Federal Probation, 1954, 18(3), 3-9.

Cargan, L., and Coates, M. The Indeterminate Sentence and Judicial Bias. Crime and Delinquency, April 1974, 29(2), 144-156.

Carter R.M. The Presentence Report and the Decision-Making Process. Journal of Research in Crime and Delinquency, 1967, 4, 203-211.

Carter, R.M., and Wilkins, L.T. Editors. Probation and Parole: Selected Readings. New York: John Wiley and Sons, Inc., 1970.

Carter, R.M., and Wilkins, L.T. Some Factors in Sentencing Policy. Journal of Criminal Law, Criminology, and Police Science, 1967, 58(4), 503-514.

Casper, J.D. American Criminal Justice: The Defendant's Perspective. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1972.

Cassou, A.K., and Taugher, B. Determinate Sentencing in California: The New Numbers Game. Pacific Law Journal, 1978, 9, 2-106.

Chambliss, W.J. Crime and the Legal Process. New York: McGraw-Hill Book Company, 1969.

Chappell, R. Disparity in Federal Sentencing. Municipal Court Review, 1966, 6, 18.

Clear, T.R. Correctional Policy, Neo-Retributionism and the Determinate Sentence. The Justice System Journal, 1978, 4, 26-48.

Clear, T.R., DiIlio, S., and Lubitz, R. The Middle Third of Guidelines Evaluation: From Multiple Regression to Sentencing Policy. Paper presented at the Annual Meetings of the American Society of Criminology. Philadelphia, November 1979.

Clear, T.R., Hewitt, J.D., and Regoli, R.M. Discretion and the Determinate Sentence: Its Distribution, Control, and Effect on Time Served. Crime and Delinquency, October 1978, 24(4), 428-445.

Cleary, J.J. Plea Negotiation and its Effects on Sentencing. Federal Bar Journal, 1978, 37, 61-75.

Crime in Massachusetts - 1977. Boston: Massachusetts Department of Public Safety, Crime Reporting Unit, 1978.

Coffee, J.C. Repressed Issues of Sentencing: Accountability, Predictability and Equality in the Era of the Sentencing Commission. Georgetown Law Journal, 1978, 66, 975-1107.

Cohen, J. and Cohen, P. Applied Regression/Correlation Analysis for the Behavioral Sciences. New York: John Wiley and Sons, Inc., 1975.

Cohen, L.E. Delinquency Dispositions. An Empirical Analysis of Processing Decisions in Three Juvenile Courts. Washington, D.C.: U.S. Government Printing Office, 1975.

Cohen, L.E. Pre-Adjudicatory Detention in Three Juvenile Courts. An Empirical Analysis of the Factors Related to Detention Decision Outcomes. Washington, D.C.: U.S. Government Printing Office, 1975.

Cohen, L.E. Who Gets Detained? An Empirical Analysis of the Pre-adjudicatory Detention of Juveniles in Denver. Washington, D.C.: U.S. Government Printing Office, 1975.

Cohn, Y. Criteria for the Probation Officer's Recommendations to the Juvenile Court Judge. Crime and Delinquency, 1963, 9, 262-275.

Cole, G.F. Will Definite Sentences Make a Difference? Judicature, 1977, 61, 58.

Comments on the Federal Indeterminate Sentence Bill. Federal Probation, July-September, 1941, 53, 22-42.

Coombs, C.H. Thurstone's Measurement of Social Values Revisited Forty Years Later. Journal of Personality and Social Psychology, 1967, 6, 85-91.

Criminal Procedure--Appealability of a Criminal Sentence--Sentence Modified on Appeal. Rutgers Law Review, 1961, 16, 186-191.

Croce, A. Sentencing Guidelines: A Study of the Sentencing of Indicated Cases in New York State. In Criminal Justice Statistics Association, Proceedings of the Fall Conference in Oklahoma City. Washington, D.C.: Criminal Justice Statistics Association, Inc., August 1977.

Cross, R. The English Sentencing System. London: Butterworths, 1971.

Davis, K.C. Administrative Law Text. St. Paul, MN: West Publishing Company, 1972.

Davis, K.C. Discretionary Justice. Baton Rouge, Louisiana: Louisiana State University Press, 1969.

Dawson, R.O. Sentencing: The Decision as to Type, Length and Conditions of Sentence. Boston: Little, Brown and Company, 1969.

Dawson, R.O. The Decision to Grant or Deny Parole: A Study of Parole Criteria in Law and Practice. Washington University Law Quarterly, 1966, 243-303.

Decision-Making Guidelines and Procedures for Parole Granting, Parole Revocation and Pardons. Boston: Commonwealth of Massachusetts, Parole Board and Advisory Board of Pardons, 1978.

Decore, J.V. Criminal Sentencing: the Role of the Canadian Courts of Appeal and the Concept of Uniformity. Criminal Law Quarterly, 1964, 6, 324.

DeCosta, F.A. Disparity and Inequality of Criminal Sentences: Constitutional and Legislative Approaches to Appellate Review and Reallocation of the Sentencing Function. Howard Law Journal, Winter 1968, 14(1), 29-59.

Definite Sentencing in New Mexico: The 1977 Criminal Sentencing Act. New Mexico Law Review, 1979, 9, 131-165.

D'Esposito, J. Sentencing Disparity: Cause and Cures. Journal of Criminal Law, Criminology, and Police Science, 1969, 60, 182-195.

Devlin, K. Sentencing Offenders in Magistrates' Courts. London: Sweet and Maxwell, 1970.

Diamond, S.S., and Zeisel, H. Sentencing Councils -- Study of Sentence Disparity and Its Reduction. University of Chicago Law Review, 1975, 43(1), 109-149.

Doyle, R.F. A Sentencing Council in Operation. Federal Probation, September 1961, 25(3), 27-30.

Duncan, O.D., Ohlin, L.E., Reiss, A.J., Jr., and Stanton, H.R. Formal Devices for Making Selections Decision. American Journal of Sociology, 1953, 58, 573-584.

Dunsky, G.P. The Constitutionality of Increasing Sentences on Appellate Review. Journal of Criminal Law and Criminology, 1978, 69, 19.

Ekman, G. Measurement of Moral Judgement: A Comparison of Scaling Methods. Perceptual and Motor Skills, 1962, 15, 3.

Emerson, R.M. Judging Delinquents: Context and Process in Juvenile Court. Chicago: Aldine, 1969.

Erickson, M.L., and Gibbs, J.P. on the Perceived Severity of Legal Penalties. Journal of Criminal Law and Criminology, 1979, 70 (2), 102-116.

Espey, L. Towards Certainty in Sentencing: Correctional Alternatives. Criminal Justice Quarterly, 1976, 4, 111-128.

Evans, W., and Gilbert, F.S. The Case for Judicial Discretion in Sentencing. Judicature, 1977, 61, 58.

Evans, W., and Gilbert, F.S., Jr. The Sentencing Process: Better Methods are Available. Federal Probation, December 1975, 39(4), 35-39.

Everson, G. The Human Element in Justice. Journal of Criminal Law and Criminology, May 1919, 10(1), 90-99.

Farrell, R.A., and Swifert, V.L. Prior Offense as a Self-Fulfilling Prophecy. Law and Society Review, 1978, 12, 437.

Federal Court of Appeals Vacates Sentence on Grounds of Severity and Remands for Sentencing. University of Pennsylvania Law Review, 1961, 109, 442.

Ferri, E. Report and Preliminary Project for an Italian Penal Code. Translated by Edgar Betts. 1921.

Ferri, E. The Nomination of a Commission for the Positivist Reform of the Italian Penal Code. Journal of Criminal Law and Criminology, May 1920, 11(1), 67-76.

Ferry, J., and Kravitz, M. Issues in Sentencing: A Selected Bibliography. Washington, D.C.: U.S. Government Printing Office, March 1978.

Figlio, R.M. A Note on End-point Truncation Effect of a Category Scale. Paper presented at the Annual Meeting of the American Society of Criminology, Cincinnati, 1979.

Figlio, R.M. The Seriousness of Offenses: An Evaluation of Offenders and Non-Offenders. Journal of Criminal Law and Criminology, 1975, 66, 189-200.

Finan, T.B. Uniform Post-Conviction Procedure Act: One State's Experience. Harvard Journal of Legislation, 1956, 2, 185.

Flaschner, F.N. The District Courts of Massachusetts. Massachusetts Law Quarterly, 1973, 58(2), 115-129.

Flaxman, K.N. The Hidden Dangers of Sentencing Guidelines. Hofstra Law Review, Winter 1979, 7(2), 259-280.

Flood, G.F. Sentencing Function of the Judge. Journal of Criminal Law, Criminology, and Police Science, 1955, 45(5), 531-540.

Fogel, D. Justice, Not Therapy: A New Mission for Corrections. Judicature, 1979, 62, 372.

Fogel, D. We are the Living Proof...: the Justice Model for Corrections. Cincinnati: W.H. Anderson Company, 1975.

Formula Sentencing: Wave of the Future. American Bar Association Journal, 1978, 64, 815.

Forst, B., Rhodes, W., and Wellford, C. Sentencing and Social Science: Research for the Formulation of Federal Sentencing Guidelines. Hofstra Law Review, Winter 1979, 7(2), 355-378.

Frankel, E. The Offender and the Court: A Statistical Analysis of the Sentencing of Delinquents. Journal of Criminal Law and Criminology, 1940, 31(4), 448-456.

Frankel, M.E. Criminal Sentences: Law Without Order. New York: Hill and Wang, Inc., 1973.

Frankel, M.E. Lawlessness in Sentencing. University of Cincinnati Law Review, 1972, 41(1), 4-24.

Frankel, S. The Sentencing Morass and a Suggestion for Reform. Criminal Law Bulletin, 1967, 3, 365-383.

Gaines, L.P. Post-Conviction Hearing Act. Chicago Bar Record, 1958, 39, 418.

Gaudet, F.J., Harris, G.S., and St. John, C.W. Individual Differences in the Sentencing Tendencies of Judges. Journal of Criminal Law and Criminology, January-February 1933, 23(5), 811-817.

Gaylin, W. Partial Justice: A Study of Bias in Sentencing. New York: Alfred Knopf, 1974.

Gelman, A.M., and Kress, J.M. How Chaotic is Sentencing in Your Jurisdiction? Judge's Journal, Spring 1978, 17, 35-56.

Genego, W.J., Goldberger, P.D., and Jackson, V.C. Parole Release Decision-Making and the Sentencing Process. Yale Law Journal, March 1975, 84, 810.

George, B.J., Jr. Sentencing Methods and Techniques in the United States. Federal Probation, March 1962, 26(2), 33-40.

Gettinger, S. Fixed Sentencing Becomes Law in Three States, Other Legislatures Wary. Corrections Magazine, 1977, 3, 16-36.

Gibson, J.I. Race as a Determinant of Criminal Sentences: A Methodological Critique and a Case Study. Law and Society Review, 1978, 12, 437.

Glick, H.R. Mandatory Sentencing: The Politics of the New Criminal Justice. Federal Probation, 1979, 43(1), 3.

Glueck, S. Further Comments on the Sentencing Problem. Federal Probation, December 1957, 21(4), 47-54.

Glueck, S. Law and Psychiatry: Cold War or Entente Cordiale? Baltimore: Johns Hopkins Press, 1962.

Glueck, S. Principles of a Rational Penal Code. Harvard Law Review, February 1928, 41(4), 453-482.

Glueck, S. The Sentencing Problem. Federal Probation, December 1956, 20(4), 15-25.

Goldstein, H. Administrative Problems in Controlling the Exercise of Police Authority. Journal of Criminal Law, Criminology, and Police Science, 1967, 58, 160.

Goldstein, H. Police Discretion: The Ideal Versus the Real. Public Advocate Review, 1963, 23, 140.

Goldstein, J. Police Discretion Not to Invoke the Criminal Process: Low-Visibility Decisions in the Administration of Justice. Yale Law Review, 1960, 69, 543.

Gottfredson, D.M. Editor. Decision-Making in the Criminal Justice System: Reviews and Essays. Washington, D.C.: U.S. Government Printing Office, 1975.

Gottfredson, D.M. Testimony before the Committee on the Judiciary, Senate of the United States on the Criminal Code Reform Act of 1977, June 8, 1977.

Gottfredson, D.M. and Ballard, K. Differences in Parole Decisions Associated with Decision-Makers. Journal of Research in Crime and Delinquency, 1966, 3, 112-119.

Gottfredson, D.M. and Beverly, R. Development and Operational Use of Predictive Methods in Correctional Work. Proceedings of the American Statistical Association. Washington, D.C.: American Statistical Association, 1962.

Gottfredson D.M., and Bonds, J. A Manual for Intake Base Expectancy Scoring: Form CDC -- BE6IA. Sacramento: Califoria Department of Corrections, Research Division, 1961.

Gottfredson D.M., Cosgrove, C.A., Wilkins, L.T., Wallerstein, J., and Rauh, C. Classification for Parole Decision Policy. Washington, D.C.: U.S. Government Printing Office, 1977.

Gottfredson, D.M., Hoffman, P.B., Sigler, M.H., and Wilkins, L.T. Making Paroling Policy Explicit. Crime and Delinquency, 1975, 21, 34-44.

Gottfredson, D.M., Wilkins, L.T., and Hoffman, P.B. Guidelines for Parole and Sentencing: A Policy Control Method. Lexington, MA: Lexington Books, 1978.

Gottfredson, D.M., Wilkins, L.T., Hoffman, P.B., and Singer, S.M. The Utilization of Experience in Parole Decision-Making, Summary Report. Washington, D.C.: U.S. Government Printing Office, 1974.

Gottfredson, M.R. Parole Board Decision Making: A Study of Disparity Reduction and the Impact of Institutional Behavior. Journal of Criminal Law and Criminology, 1979, 70 (1), 77-88.

Gottfredson, M.R. Parole Guidelines and the Reduction of Sentencing Disparity: A Preliminary Study. Albany, NY: Criminal Justice Research Center, 1979.

Gottfredson, M.R., and Gottfredson, D.M. Decision Making in Criminal Justice: Toward the Rational Exercise of Discretion. Cambridge, MA: Ballinger Publishing Company, 1980.

Gottfredson, S. The Dimensional Scaling of Offense Severity: Applications for the measurement of recidivism. 1979. (Mimeograph).

Governor's Committee on Law Enforcement and Administration of Justice. Revision of the Massachusetts Criminal Code. March 4, 1968. (Mimeograph).

Governor's Task Force on Automobile Theft. Auto Theft in Massachusetts: An Executive Response. Boston, Governor's Task Force on Automobile theft, March 1980.

Green, E. Inter- and Intra-Racial Crime Relative to Sentencing. Journal of Criminal Law, Criminology, and Police Science, 1964, 55, 348-358.

Green, E. Judicial Attitudes to Sentencing. London, 1969.

Greenberg, D.F. A Reply. Prison Journal, Spring-Summer 1972, 52(1), 33-41.

Greenberg, D.F. Socioeconomic Status and Criminal Sentences: Is there an Association (Comment on Chiricos and Waldo). American Sociological Review, February 1977, 42(1), 174-175.

Grindstaff, C.F. Public Attitudes and Court Dispositions: A Comparative Analysis. Sociology and Social Research, 1974, 58, 417-426.

Grosman, B. The Prosecutor: An Inquiry into the Exercise of Discretion. Toronto: University of Toronto Press, 1969.

Hagan, J. Extra-Legal Attributes and Criminal Sentencing: An Assessment of a Sociological Viewpoint. Law and Society Review, 1974, 8, 357-383.

Hagerstrom, A. Inquiries into the Nature of Law and Morals. English translation by C.D. Broad. London: Routledge and Kegan Paul, 1953.

Hall, L. S. 200 - An Act Establishing the Massachusetts Criminal Code. Massachusetts Law Quarterly, 1972, 57 (1), 63-65.

Halperin, D.J. Appellate Review of Sentence in Illinois - Reality or Illusion. Illinois Bar Journal, 1966, 55, 300.

Harris, M.K. Disquisition on the Need for a New Model for Criminal Sanctioning Systems. West Virginia Law Review, February 1975, 77(2), 263-326.

Harris, M.K., and Dunbaugh, F.M. Premise for a Sensible Sentencing Debate: Giving Up Imprisonment. Hofstra Law Review, Winter 1979, 7(2), 417-456.

Hart, H.L.A. Punishment and Responsibility. New York: Oxford University Press, 1968.

Hart, H.L.A. The Concept of Law. Oxford: Clarendon Press, 1961.

Hays, W.L. Statistics for the Social Sciences. Second Edition. New York: Holt, Rinehart and Winston, Inc., 1973.

Heller, N.B., and McEwen, J.T. Applications of Crime Seriousness Information in Police Departments. Journal of Criminal Justice, 1973, 1, 241-253.

Hennessey, E.F. The State of the Judiciary. Annual Report of the Chief Justice of the Supreme Judicial Court. Massachusetts Law Quarterly, 1977, 62 (1), 7-20.

Heumann, M., and Loftin, C. Mandatory Sentencing and the Abolition of Plea-bargaining: The Michigan Felony Firearm Statute. Law and Society Review, 1979, 13(2), 393-430.

Hewitt, J.D., and Little, B. Examining the Research Underlying the Sentencing Guidelines Concept in Denver, Colorado: A Partial Replication of a Reform Effort. Journal of Criminal Justice, 1981, 9, 51-62.

Hindelang, M. The Uniform Crime Reports Revisited. Journal of Criminal Justice, 1974, 2, 1.

Hoffman, P.B. Paroling Policy Feedback. Report Eight. Parole Decision-Making Project. Davis, CA: National Council on Crime and Delinquency Research Center, 1973.

Hoffman, P.B., and Beck, J.L. Parole Decision-Making: A Salient Factor Score. Journal of Criminal Justice, 1974, 2, 195-206.

Hoffman, P.B., and Beck, J.L. Salient Factor Score Validation -- A 1972 Release Cohort. Journal of Criminal Justice, 1976, 4, 69-76.

Hoffman, P.B., and DeGostin, L.K. An Argument for Self-Imposed Explicit Judicial Sentencing Standards. Journal of Criminal Justice, 1975, 13(3), 195-206.

Hoffman, P.B., and DeGostin, L.K. Parole Decision-Making: Structuring Discretion. Report Five. Washington, D.C.: United States Board of Parole Research Unit, June 1974.

Hoffman, P.B., and DeGostin, L.K. Parole Decision-Making: Structuring Discretion. Federal Probation, December 1974, 38, 7-15.

Hoffman, P.B., and Gottfredson, D.M. Paroling Policy Guidelines: A Matter of Equity. Report Nine. Parole Decision-Making Project. Davis, CA: National Council on Crime and Delinquency Research Center, 1973.

Hoffman, P.B., and Stone-Meierhoefer, B. Application of Guidelines to Sentencing. Law and Psychology Review, 1977, 3, 53-70.

Hoffman, P.B., and Stover, M.A. Reform in the Determination of Prison Terms: Equity, Determinacy, and the Parole Release Function. Hofstra Law Review, Fall 1978, 7(1), 89-121.

Hoffman, W.E. A Sentencing Philosophy. Federal Probation, December 1968, 32(4), 3-8.

Hogarth, J. Sentencing as a Human Process. Toronto, Canada: University of Toronto Press, 1971.

Holtzoff, A. Summary of the Indeterminate Sentence Bill. Federal Probation, July-September 1941(a), 53, 20-21.

Holtzoff, A. The Indeterminate Sentence: Its Social and Legal Implications. Federal Probation, January-March 1941, 5(1), 3-7.

Hood, R. Sentencing in Magistrates Courts. London: Stevens and Sons, 1962.

Hopkins, A. Is there a Class Bias in Sentencing? (Comment on Chiricos and Waldo). American Sociological Review, February 1977, 42(1), 176-177.

Hosner, C. Group Procedures in Sentencing: A Decade of Practice. Federal Probation, December 1970, 34, 18-25.

Hrones, S. The Appeal of a Criminal Case to the Appellate Courts in Massachusetts. Boston Bar Journal, 1979, 23 (6), 5-13.

Hruska, R.L. Appellate Review of Sentence. American Criminal Law Quarterly, 1969, 8, 10.

Hsu, M. Cultural and Sexual Differences in the Judgement of Criminal Offenses. Journal of Criminal Law, Criminology, and Police Science, 1973, 64, 348-353.

Illinois Approves Determinate Sentencing; Parole Replaced by New Supervised Release. Criminal Justice Newsletter, January 2, 1978, 1-2.

INSLAW, Inc., and Yankelovick, Skelly and White, Inc. Federal Sentencing: Toward a More Explicit Policy of Criminal Sanctions. Washington, D.C.: U.S. Department of Justice, Federal Justice Research Program, May 1981.

Institute for Law and Social Research. PROMIS: Prosecutor's Management Information Briefing Series: 1) Management Overview of PROMIS, 2) Case Sentencing, 3) Uniform Case Evaluation and Rating. Washington, D.C.: Institute for Law and Social Research, 1964.

Institute of Judicial Administration. Disparity in Sentencing of Convicted Defendants. NY: Institute of Judicial Administration April 7, 1954.

Jacoby, J. The Prosecutor's Charging Decision: A Policy Perspective. Washington, D.C.: U.S. Government Printing Office, 1977.

James, H. Crisis in the Courts. New York: David McKay, Inc., 1968.

Johnson, B. et al. Discretion in Felony Sentencing--A Study of Influencing Factors. Washington Law Review, August 1973, 48(4), 857-889.

Johnson, G.W. The Board of Prison Terms and Paroles: Criteria in Decision Making. Washington Law Review, July 1976, 51(3), 643-651.

Joint Committee on Continuing Legal Education of the American Law Institute and the American Bar Association. The Problem of Sentencing. Philadelphia: American Bar Association, 1962. New York (1964)

Jones, J.B. Prosecutors and the Disposition of Criminal Cases: An Analysis of Plea Bargaining Rates. Journal of Criminal Law and Criminology, 1978, 69, 402.

Judicial Discretion. The Central Law Journal, December 24, 1880, 11(26), 504-507.

Kadish, S.H.D. Legal Norm and Discretion in the Police and Sentencing Processes. Harvard Law Review, 1962, 75, 904.

Kaplan, J. The Prosecutorial Discretion - A Comment. Northwestern Law Review, 1965, 60, 178-193.

Kaufman, H.P. Judicial Discretion. The American Law Review, July-August 1883, 17, 567-573.

Kelsen, H. General Theory of Law and State. New York: Russell and Russell, 1961.

Kennedy, E.M. Criminal Sentencing: A Game of Chance. Judicature, December 1976, 60, 208-215.

Kennedy, E.M. Symposium on Sentencing, Part I: Introduction. Hofstra Law Review, Fall 1978, 7(1), 1-9.

Kerlinger, F.N., and Pedhazer, F. Multiple Regression in Behavorial Research. New York: Holt, Rinehart and Winston, 1973.

Kilpatric, D., and Brummel, J. Sentencing Study. Washington Law Review, 1976, 52, 103-119.

Kress, J.M. Potential Guideline Variables. An Analytical Policy Paper prepared for the Use of the Massachusetts Superior Court Committee on Probation and Parole at their Meeting on March 2, 1979 in Worcester, Massachusetts.

Kress, J.M. Prescription for Justice: The Theory and Practice of Sentencing Guidelines. Cambridge, MA: Ballinger Publishing Company, 1980.

Kress, J.M. Research Problems Encountered in Moving Towards Equity in Judicial Decision-Making. The Justice System Journal, 1978, 4, 71-87.

Kress, J.M. Who Should Sentence: The Judge, The Legislature, Or...? Judge's Journal, Winter 1978, 17, 12-45.

Kress, J.M., Wilkins, L.T., and Gottfredson, D.M. Is the end of Judicial Sentencing in Sight? Judicature, December 1976, 60(5), 216-222.

Krus, D.J., Sherman, J.L., and Krus, P.H. Changing Values over the Last Half-Century: The Story of Thurstone's Crime Scales. Psychological Reports, 1977, 40, 207-211.

Kulig, F.H. Plea Bargaining, Probation, and Other Aspects of Conviction and Sentencing. Creighton Law Review, July 1975, 8, 938-978.

Labbe, R.M. Appellate Review of Sentences: Penology on the Judicial Doorstep. Journal of Criminal Law and Criminology, 1977, 68(1), 122-134.

Lancucki, L., and Tarling, R. The Relationship Between Mean Cost Rating and Kendall's Rank Correlation Coefficient. Social Science Research, Academic Press, 1978.

Lane, H.E. Illogical Variations in Sentences of Felons Committed to Massachusetts State Prison. Journal of Criminal Law, Criminology, and Police Science, 1941, 32(2), 171-190.

Lasker, M.E. Presumption Against Incarceration. Hofstra Law Review, Winter 1979, 7(2), 407-416.

Lax, D.A. An Interim Report of the Monte Carlo Study of Robust Estimators of Width. Technical Report No. 93 (Series 2). Department of Statistics, Princeton Univeristy, 1975.

LeFave, W.R. Arrest: The Decision to take a Suspect into Custody. Boston: Little, Brown and Company, 1965.

LeFrancois, A.G. Examination of a Desert-based Presumptive Sentence Schedule. Journal of Criminal Justice, Spring 1978, 6(1), 35-46.

Leighton, G. Post-Trial Procedures. Illinois Bar Journal, 1958, 47, 211.

Levin, M.A. A Good Sentence is More than a Subject and a Passive Verb. Judge's Journal, 1978, 78, 40.

Levin, T. Sentencing the Criminal Offender. Federal Probation, March 1949, 13(1), 3-6.

Levin, T. Toward a More Enlightened Sentencing Procedure. Nebraska Law Review, 1966, 43(3), 499-512.

Lizotte, A.J. Extra-legal Factors in Chicago's Criminal Courts: Testing the Conflict Model of Criminal Justice. Social Problems, 1978, 25, 564-80.

Lyday, W.M. Sentencing: Judicial Policy in a Political Structure. The Journal of Criminal Law and Criminology, 1977, 68, 458.

Mandatory Sentencing Required by the Bartley-Fox Bill. Massachusetts Law Quarterly, 1976, 61 (3), 157-158.

Mannheim, H., and Wilkins, L.T. Prediction Methods in Relation to Barstal Training. Stationery Office, 1955.

Martinson, R. New Findings, New Views: A Note of Caution Regarding Sentencing Reform. Hofstra Law Review, Winter 1979, 7(2), 243-258.

Marx, T.J. A Study of Racial Disparity in Massachusetts Superior Court Department. Boston: Massachusetts Superior Court Department, April 8, 1980(a). Mimeograph.

Marx, T.J. The Question of Rachial Disparity in Massachusetts Superior Court Sentences. Boston: Massachusetts Superior Court Department, May 16, 1980(g). Mimeograph.

Massachusetts Criminal Law Revision Commission. Proposed Criminal Code of Massachusetts with Revision Commission Notes. Rochester, N.Y.: The Lawyers Co-Operative Publishing Company, 1972.

Massachusetts Superior Court Sentencing Guidelines Project. Massachusetts Sentencing Guidelines. Boston: Massachusetts Superior Court, November 1979. Mimeograph.

Massachusetts Superior Court Sentencing Guidelines Project. Massachusetts Sentencing Guidelines. Boston: Massachusetts Superior Court, May 1980. Mimeograph.

Massachusetts Uniform Crime Report: Preliminary Crime Index 1978. Boston: Massachusetts Department of Public Safety, Crime Reporting Unit, February 1979.

McAnany, P.D., Merritt, F.S., and Tromanhauser. Illinois Reconsiders Flat Time: An Analysis of the Impact of the Justice Model. Chicago-Kent Law Review, 1976, 52(3), 621-662.

McCall, C.C. The Future of Parole - In Rebuttal of S.1437. Federal Probation, 1978, 42(4), 3.

McCarthy, J.P., Jr. Report of the Sentencing Guidelines Project to the Administrative Director of the Courts. October 23, 1978; and Supplemental Report of the Sentencing Guidelines Project to the Administrative Director of the Courts. NJ: Administrative Office of the Courts, March 29, 1979.

McCarthy, J.P., Jr., Sheflin, N., and Barraco, J.J. Report of the Sentencing Guidelines Project to the Administrative Director of the Courts on the Relationship between Race and Sentencing. NJ: Administrative Office of the Courts, September 4, 1979.

McClea, S. Impact Analysis of Mandatory Sentencing Proposals for Pennsylvania. In Criminal Justice Statistics Association, Proceedings of the Fall Conference in Oklahoma City. Washington, D.C.: Criminal Justice Statistics Association, Inc., August, 1977.

McGee, R.A. California's New Determinate Sentencing Act. Federal Probation, 1978, 42(1), 3.

McGillis, D. The Major Offense Bureau, An Exemplary Project --Bronx County District Attorney's Office, New York. Washington, D.C.: U.S. Government Printing Office, 1977.

Miller, F.W. Prosecution: The Decision to Charge a Suspect with a Crime. Boston: Little, Brown and Company, 1969.

Minnesota Sentencing Guidelines Commission. Report to the Legislature. St. Paul, MN: Minnesota Sentencing Guidelines Commission, 1980.

Mistake About Parole Guidelines Does not Permit Judge to Change Sentence. The Criminal Law Reporter: Court Decisions and Proceedings, 1978, 22, 2502.

Morris, F., Kadish, S., and Zimring, F. Sentencing Provisions in the Proposed Federal Code. Federal Rules Decisions, 1978, 80, 151-66.

Morris, N. Sentencing and Parole. Australian Law Journal, 1977, 51, 523-31.

Morris, N. The Future of Imprisonment. Chicago: University of Chicago Press, 1974.

Morris, N. The Future of Imprisonment: Toward a Punitive Philosophy. Michigan Law Review, May 1974, 72(6), 1161-1180.

Morrish, P. and McLean, I. The Crown Court: An Index of Common Penalties and Formalities. Seventh Edition. London: Barry Rose Publishers, 1975.

Morse, W.L., Kennedy, H.M., Ellis, W.J. et al. Comments on the Federal Indeterminate Sentence Bill. Federal Probation, July-September 1941, 5(3), 22-29, 34-42.

Moseley, W.H. Parole: How is it Working? Journal of Criminal Justice, 1977, 5, 185-203.

Mosteller, F., and Tukey, J.W. Data Analysis and Regression: A Second Course in Statistics. Reading, MA: Addison-Wesley, 1977.

Move to Abolish Parole Termed 'Confusing and Filled with Emotion'. Crime Control Digest, 1977, 11, 5.

Mueller, G.O.W. Oebikigt on Appeal - Appellate Review of Legal but Excessive Sentences. Vanderbilt Law Review, 1962, 15, 671.

Mueller, G.O.W., and Le Poole, F. Appellate Review of Legal but Excessive Sentences: A Comparative Study. Vanderbilt Law Review, 1968, 411.

Nagel, S. Rational Method for Determining Prison Sentences. Judicature, 1978, 61, 371.

National Advisory Commission. A National Strategy to Reduce Crime. A Report by the National Advisory Commission on Criminal Justice Standards and Goals. New York: Avon Books, 1975.

National Conference of State Trial Judges. Appeals from Sentencing. August 1965.

Newman, D. Conviction: The Determination of Guilt or Innocence Without Trial. Boston: Little, Brown and Company, 1966.

Nie, N.H., Hull, C.H., Jenkins, J.G., Steinbrenner, K., and Bent, D.H. Statistical Packages for the Social Sciences. Second Edition. New York: McGraw-Hill Book Company, 1975.

Nordbye, G.H. The Indeterminate Sentence Plan As A Method of Adjusting Inequalities in Sentence. Federal Probation, January-March 1941, 5(1), 23.

Normandeau, A. Crime Indices for Eight Countries. International Criminal Police Review, 1970, 234, 15.

Normandeau, A. The Measurement of Delinquency in Montreal. Journal of Criminal Law, Criminology, and Police Science, 1966, 57, 172-177.

O'Donnell, P., Churgin, M.J., and Curtis, D.C. Toward a Just and Effective Sentencing System - Agenda for Legislative Reform. New York: Praeger Press, 1977.

Ohlin, L.E., and Remington, F.J. Sentencing Structure: Its Effects Upon Systems for the Administration of Criminal Justice. Law and Contemporary Problems, 1958, 23, 495-507.

O'Leary, V., Gottfredson, M., and Gelman, A. Contemporary Sentencing Proposals. Criminal Law Bulletin, 1975, 11(5), 555-586.

Olivecrona, K. Law as Fact. Second Edition. London: Stevens and Sons, 1971.

Orland, L. From Vengeance to Vengeance: Sentencing Reform and the Demise of Rehabilitation. Hofstra Law Review, Fall 1978, 7(1), 29-56.

Orland, L. Is Determinate Sentencing an Illusory Reform? Judicature, 1979, 62, 372.

Orland L., and Tyler, H.R. Justice in Sentencing--Papers and Proceedings for the Sentencing Institute for the First and Second U.S. Judicial Circuits. New York: Foundation Press, 1974.

Orrick, D. Legal Issues in Structuring Sentencing Discretion. New England Journal on Prison Law, 1978, 4, 327-53.

Osborne, H.V. The Establishment of Uniformity in Determining Punishments in Different Courts and Different Jurisdictions. Journal of Criminal Law and Criminology, March 1917, 7(6), 804-809.

Osborough, N. Police Discretion Not to Prosecute Juveniles. Modern Law Review, 1965, 28, 421.

Packer, H.L. The Police and the Community. Stanford Law Review, 1969, 22, 1314-1317.

Packer, H.L. Limits of the Criminal Sanction. Stanford: Stanford University Press, 1968.

Page, L. The Sentence of the Court. London: Fabor and Farbor,948.

Palmer, J., and Zalman, M. People v. Tanner: A Legal and Empirical Study in Sentencing. New England Law Review, 1978, 14, 82.

Palmer, L.I. A Model of Criminal Dispositions: An Alternation to Official Discretion in Sentencing. The Georgetown Law Journal, October 1973, 62(1), 1-59.

Parole Guidelines' Use May Raise Statutory Constitutional Problems. The Criminal Law Reporter: Court Decisions and Proceedings, 1978, 23, 2083.

Partridge, A., and Eldridge, W.B. The Second Circuit Sentencing Study; a Report to the Judges of the Second Circuit. Washington, D.C.: Federal Judicial Center, 1974.

Partridge, A., and Leavitt, M.R. The Feasibility of a National Sentencing Policy: A Critique. Federal Probation, 1979, 44, 69-91.

Patterns of Court Sentences Are Under Examination. Crime Control Digest, 1978, 72, 34.

Pease, K., Ireson, J., and Thorpe, J. Additivity Assumptions in the Measurement of Delinquency. British Journal of Criminology, 1974, 14, 256-263.

Pease, K., Ireson, J., and Thorpe, J. Modified Crime Indices for Eight Countries. Journal of Criminal Law, Criminology, and Police Science, 1975, 66, 209-214.

Pepitone, A., and DiNubile, M. Contrast Effects in Judgements of Crime Severity and the Punishment of Criminal Violators. Journal of Personality and Social Psychology, 1976, 33, 448-459.

Petersilia, J., and Greenwood, P.W. Mandatory Prison Sentences: Their Projected Effects on Crime and Prison Populations. Journal of Criminal Law and Criminology, 1978, 69, 604.

Ploscowe, M. An Examination of Some Dispostions Relating to Motives and Character in Modern European Penal Codes. Journal of Criminal Law and Criminology, May 1930, 21(1), 26-40.

Pope, C.E. The Influence of Social and Legal Factors on Sentence Dispostions: A Preliminary Analysis of Offender-Based Transaction Statistics. Journal of Criminal Justice, Fall 1976, 4(3), 203-221.

Potential Adequacy of the California Uniform Determinate Sentence Law. Southwestern University Law Review, 1978, 10, 149-71.

President's Commission on Law Enforcement and the Administration of Justice. The Challenge of Crime in a Free Society. Washington, D.C.: U.S. Government Printing Office, 1967.

Proceedings of the Special Conference on Determinate Sentencing. Determinate Sentencing: Reform or Regression? Washington, D.C.: U.S. Government Printing Office, 1977.

Progress Report Indicates Great Disparity in Sentencing. Crime Control Digest, 1978, 12, 11.

Pugsley, R.A. Retributivism: A Just Basis for Criminal Sentences. Hofstra Law Review, Winter 1979, 7(2), 379-405.

Radzinowicz, L. and Hood, R. A Dangerous Direction for Sentencing Reform. The Criminal Law Review, 1978, 713.

Radzinowicz, L. and Hood, R. English Attempt to Reshape the Sentencing Structure. Columbia Law Review, 1978, 1145-58.

Reasons, C. E. On Methodology, Theory and Ideology (Comment on Chiricos and Waldo). American Sociological Review, February 1977, 42(1), 177-181.

Rector, M.G. Sentencing Reform and Litigation. Judicature, 1969, 53, 58.

Remington, F.J. A Book Review: Discretionary Justice. University of Chicago Law Review, 1969, 36, 884-897.

Remington, F.J. The Role of Police in a Democratic Society. Journal of Criminal Law, Criminology, and Police Science, 1965, 56, 361.

Remington, F.J., Newman, D., Kimball, E., Melli, M., and Goldstein, H. Criminal Justice Administration. Indianapolis: Bobbs-Merrill Company, Inc., 1969.

Remington, F.J., and Newman, D.J. The Highland Park Institute on Sentence Disparity. Federal Probation, March 1962, 26(1), 3-9.

Resentence--Credit for Time Served Under Erroneous Sentence. Boston University Law Review, 1953, 33, 236.

Rhodes, W.M., and Conly, C. Analysis of Federal Sentencing. Washington, D.C.: U.S. Department of Justice, Federal Justice Research Program, May 1981.

Rich, W.D., Sutton, S.P., Clear, T.R., and Saks, M.J. Sentencing Guidelines: Their Operation and Impact on the Courts. Williamsburg, VA.: National Center for State Courts, 1980. Mimeograph.

Richey, C.R. Appellate Review of Sentencing: Recommendation for a Hybrid Approach. Hofstra Law Review, Fall 1978, 7(1), 71-87.

Riedel, M. Perceived Circumstances, Inferences of Intent and Judgements of Offense Seriousness. Journal of Criminal Law and Criminology, 1975, 66, 201-208.

Right to Credit for Time Served and to Preservation of the Original Sentence. Illinois Law Forum, 1967, 180.

Rose, A.M., and Prell, A.E. Does the Punishment Fit the Crime? A Study in Social Valuation. American Journal of Sociology, 1955, 61, 247.

Rose, G.N. Concerning the Measurement of Delinquency. British Journal of Criminology, 1966, 6, 414.

Rossi, P.H., Freeman, H.E., and Wright, S.R. Evaluation: A Systematic Approach. Beverly Hills, CA: Sage Publications, 1979.

Rossi, P.H., Waite, E., Bose, C.E., and Berk, R.E. The Seriousness of Crimes: Normative Structure and Individual Differences. American Sociological Review, 1974, 39, 224-237.

Roth, J.A. Prosecutor Perceptions of Crime Seriousness. Journal of Criminal Law and Criminology, 1978, 69, 232-242.

Rubin, S. Disparity and Equality of Sentences--A Constitutional Challenge. Federal Rules Decisions, 1965, 40, 553.

Rubin, S. Sentencing Goals: Real and Ideal. Federal Probation, June 1957, 21(2), 51-56.

Ruth, H. Promoting Consistent Policy in the Criminal Justice Process. Virginia Law Review, 1967, 53, 1489-1508.

Saleilles, R. The Individualization of Punishment. Boston: Little, Brown and Company, 1911.

Schiller, S.A. Illinois' New Sentencing Laws - The Effect on Sentencing in Cook County: Some Early Returns. Chicago Bar Record, 1978, 60, 130-2.

Schubert, G. Quantitative Analysis of Judicial Behavior. Glencoe, IL: The Free Press, 1959.

Sebba, L. Some Explorations in the Scaling of Penalties. Journal of Research in Crime and Delinquency, 1978, 15, 247-265.

Sechrest, D.K. Comparison of Inmates' and Staff's Judgements of the Severity of Offenses. Journal of Research in Crime and Delinquency, 1969, 6, 41-55.

Sellin, T. Pioneers in Criminology - Enrico Ferri. Journal of Criminal Law, Criminology, and Police Science, January-February 1958, 48(5), 481-492.

Sellin, T., and Wolfgang, M. The Measurement of Delinquency. New York: John Wiley and Sons, Inc., 1964.

Sentencing: A Symposium. Law and Contemporary Problems, 1958, 23, 399.

Serrill, M.S. Determinate Sentencing: The History, the Theory, the Debate. Corrections Magazine, 1977, 3, 3-13.

Seymour, W.N. 1972 Sentencing Study for the Southern District of New York. New York State Bar Journal, April 1973, 163-171.

Shane-Dubow, S., Smith, W.F., and Haralson, K.B. Felony Sentencing in Wisconsin. Wisconsin: Wisconsin Center for Public Policy, Public Policy Press, 1979.

Shapiro, B.A. Guide to Sentencing in Massachusetts. Massachusetts Law Quarterly, 1976, 61 (2), 87-102.

Sharp, L.J. The Pilot Institute on Sentencing. Federal Probation, December 1959, 23(4), 9-11.

Shelly, P.L., and Sparks, R.F. Crime and Punishment. Paper presented at the Annual Meeting of the American Society of Criminology, San Francisco, 1980.

Signorelli, E.L. A Judicial Analysis and Critique of the New Drug and Sentencing Laws. New York State Bar Journal, January 1974, 46(1), 9-19.

Singer, R. In Favor of Presumptive Sentences set by a Sentencing Commission. Criminal Justice Quarterly, 1977, 5, 88-105.

Singer, R. Just Deserts: Sentencing based on Equality and Desert. Cambridge, MA: Ballinger Publishing Company, 1979.

Singer, R.G., and Hand, R.C. Sentencing Computation: Laws and Practices. Criminal Law Bulletin, 1974, 10, 318-347.

Skolnick, J.H. Justice Without Trial: Law Enforcement in a Democratic Society. New York: John Wiley and Sons, Inc., 1966.

Skrivseth, K. Abolishing Parole: Assuring Fairness and Certainty in Sentencing. Hofstra Law Review, Winter 1979, 7(2), 281-314.

Smith, R. A Probation Officer Looks at Disparities in Sentence. Federal Probation, 1962, 26, 19.

Smith, T. The Sentencing Council and the Problems of Disproportionate Sentences. Federal Probation, June 1973, 27, 5-9.

Smith, W.F. Sentencing Alternatives Available to the Courts. Federal Probation, March 1962, 26(2), 3-5.

Social Service Faculty of the Catholic University of the Sacred Heart in Milan. On the Reform of the Italian Penal Code. Journal of Criminal Law and Criminology, February 1924, 14(4), 524-543.

Somit, A., Tanenhaus, J., and Wilke, W. Aspects of Judicial Sentencing Behavior. University of Pittsburgh Law Review, June 1960, 21(4), 613-620.

Spaeth, E.B., Jr. A Response to Struggle for Justice. Prison Journal, Spring-Summer 1972, 52(1), 4-32.

Spangenberg, R.L., Homans, W.P., Jr., and Flaschner, F.N. Bail Reform in Massachusetts, 1965-1967. Massachusetts Law Quarterly, 1967, 52 (2), 135-152.

Sparks, R.F. Criminal Opportunities and Crime Rates. In Indicators of Crime and Criminal Justice: Quantitative Studies. Stephen E. Fienberg and Albert J. Reiss, Jr., Editors. Washington, D.C.: U.S. Government Printing Office, 1980.

Sparks, R.F. Methodological Problems of Retrospective Social Surveys: With Special Reference to Surveys of Criminal Victimization. Washington, D.C.: U.S. Government Printing Office, 1981b.

Sparks, R.F., Genn, H.G., and Dodd, D.J. Surveying Victims: A Study of the Measurement of Criminal Victimization, Perceptions of Crime and Attitudes to Criminal Justice. New York: John Wiley, and Sons, Inc. 1977.

Speer, B.M. Criminal Jurisdiction of the District Court Under the Court Reorganization Act of 1978. Boston: Massachusetts Defenders Committee, 1979.

Special Joint Committee on Uniform Sentencing and Revision of the Criminal Law Statutes. First Interim Report. Boston, 1978. Mimeograph.

Special Joint Committee on Uniform Sentencing and Revision of the Criminal Law Statutes. Second Interim Report. Boston, 1979. Mimeograph.

Special Joint Committee on Uniform Sentencing and Revision of the Criminal Law Statutes. Third Interim Report. Boston, House No. 6304, January 12, 1981. Mimeograph.

Spica, A. The A to Z of the Presentence Report. Federal Probation, 1978, 42(4), 51.

Stecher, B.A. and Sparks, R.F. Removing the Effects of Discrimination in Sentencing Guidelines. In Sentencing Disparity. Martin Forst, Editor. Beverly Hills, California: Sage Publication, 1981 (in press).

Steinmetz, . Disparity in Sentencing. Trial Judges Journal, January 1968, 7, 21.

Stevens, S.S. A Metric for Social Concensus. Science, 1966, 151, 530-541.

Stevens, S.S. On the Operation Known as Judgement. American Scientist, 1966, 54, 385-401.

Strand, R.G. Sentencing Disparity: How One Jurisdiction Reduced It. Judges' Journal, Spring 1978, 17, 33.

Sutton, L.P. Federal Criminal Sentencing: Perspectives of Analysis and a Design for Research. Washington, D.C.: U.S. Government Printing Office, 1978.

Sutton, L.P. Federal Sentencing Patterns: A Study of Geographical Variations. Washington, D.C.: U.S. Government Printing Office, 1978.

Sutton, L.P. Predicting Sentences in Federal Courts: The Feasibility of a National Sentencing Policy. Analytic Report 19. Washington, D.C.: U.S. Government Printing Office, 1978.

Sutton, L.P. Variations in Federal Criminal Sentences: A Statistical Assessment at the National Level. Washington, D.C.: U.S. Government Printing Office, 1978.

Taylor, E.L. In Search of Equity: The Oregon Parole Matrix. Federal Probation, 1979, 43(1), 52.

Temin, C.E. The Indeterminate Sentence -- The Muncy Experience. Prison Journal, Spring-Summer 1972, 52(1), 42-49.

Tennant, P.A. Three Views on Sentencing. Federal Bar Journal, Spring 1973, 32(2-3), 201-223.

Teslik, W.R. Prosecutorial Discretion: The Decision to Charge. An Annotated Bibliography. Washington, D.C.: Law Enforcement Assistance Administration, 1975.

Texas Study of Adult Competency. Austin: University of Texas, 1976.

The Question of Mandatory Prison Sentences to Reduce Crime. The Congressional Digest, August-September 1976, 55(8-9), 193-224.

Thomas, C.W., Cage, R.J., and Foster, S.C. Public Opinion on Criminal Law and Legal Sanctions: An Examination of Two Conceptual Models. Journal of Criminal Law and Criminology, 1976, 67, 110-116.

Thomas, D.A. Appellate Review of Sentences and the Development of Sentencing Policy: The English Experience. Alabama Law Review, 1967-1968, 20, 193-226; Criminal Law Review, 1969, 235.

Thomas, D.A. Establishing a Factual Basis for Sentencing. Criminal Law Review, 1970, 80.

Thomas, D.A. Principles of Sentencing. London: Heinemann Publishers, 1970.

Thomas, D.A. Sentencing Co-Defendants--When is Uniform Treatment Necessary? Criminal Law Review, 1964, 22.

Thomas, D.A. Sentencing, The Basic Principles. Criminal Law Review, 1967, 455.

Thomas, D.A. Sentencing - The Case for Reasoned Decisions. Criminal Law Review, 1963, 243-253.

Thomas, D.A. Theories of Punishment in the Court of Criminal Appeal. Modern Law Review, 1964, 27, 9.

Thompson, D.R., and Wollaston, H.W. Court of Appeal, Criminal Division. London: Charles Knight and Company, Ltd., 1969.

Thurstone, L.L. A Law of Comparative Judgement. Psychological Review, 1927, 34, 273-286.

Thurstone, L.L. The Method of Paired Comparisons for Social Values. Journal of Abnormal and Social Psychology, 1927, 21, 384-400.

Tiffany, L.P., Avichai, Y., and Peters, G.W. A Statistical Analysis of Sentencing in Federal Courts: Defendants Convicted after Trial, 1967-1968. Journal of Legal Studies, 1975, 4, 369-390.

Tompkins, D. Sentencing the Offender: A Bibliography. Washington, D.C.: Institute of Government Studies, 1972.

Tonry, M.H. The Sentencing Commission in Sentencing Reform. Hofstra Law Review, Winter 1979, 7(2), 315-353.

Tough Sentencing Bills Gain in the States. Criminal Justice Newsletter, April 25, 1977, 6.

Tshelrow-Bebutow, M.A. Indeterminate Sentence and Soviet Penal Law. Journal of Criminal Law and Criminology, November 1928, 19(3), 408-410.

Tukey, J.W. Exploratory Data Analysis. Reading, MA: Addison-Wesley Publishing Company, 1977.

Turnbladh, W.C. A Critique of the Model Penal Code Sentencing Proposals. Law and Contemporary Problems, 1958, 23, 544-556.

Twentieth Century Fund Task Force on Criminal Sentencing. Fair and Certain Punishment. New York: McGraw-Hill Book Company, 1976.

Tyler, H.R., Jr. Sentencing Guidelines: Control of Discretion in Federal Sentencing. Hofstra Law Review, Fall 1978, 7(1), 11-28.

Ulmer, S.S., and Andrews, D.M. Sentencing Disparities in the Fayette Circuit Court: 1972-1975. Kentucky Bench and Bar, 1979, 43, 34.

Underwood, B.D. Law and the Crystal Ball: Predicting Behavior with Statistical Inference and Individualized Judgement. Yale Law Journal, June 1979, 88(7), 1408-1488.

U.S. Congress House of Representatives. Committee on the Judiciary. Improving the Administration of Justice by Authorizing the Judicial Conference of the United States to Establish Institutes and Joint Councils on Sentencing, to Provide Additional Methods of Sentencing. 84th Congess, 2nd Session, Report No. 1946, June 23, 1958.

U.S. Congress Senate. S.1638 - A Bill to Provide the Imposition of Indeterminate Sentences in Certain Cases. Federal Probation, July-September 1941, 5(3), 19-20.

U.S. Court of Appeals. Appellate Review of Sentences. A Symposium at the Judicial Conference of the U.S. District Court of Appeals for the Second Circuit. Federal Rules Decisions, 1962, 32.

U.S. v. Grayson (98 Sup. Ct. 1610): Questionable Support for Broad Judicial Discretion in Sentencing. The John Marshall Journal of Practice and Procedure, 1979, 12, 407-34.

vanden Haag, E. Punitive Sentences. Hofstra Law Review, Fall 1978, 7(1), 123-138.

Velez-Diaz, A., and Megaree, N. An Investigation of Differences in Value Judgements between Youthful Offenders and Non-Offenders in Puerto Rico. Journal of Criminal Law, Criminology, and Police Science, 1971, 61, 549-553.

Verkoeyen, C.M.R. Demanding More Severe Punishment. British Journal of Criminology, 1978, 18, 348.

Veron, J.M. Due Process and Substantive Accountability: Thoughts Toward a Model of Just Decision-Making. Louisiana Law Review, 1978, 38, 919-54.

von Hirsch, A. Book Review: Standards Without Goals: Toward a Just and Effective Sentencing System: Agenda for Legislative Reform. Hofstra Law Review, Winter 1979, 7(2), 457-470.

von Hirsch, A. Doing Justice: The Choice of Punishments. Report of the Committee for the Study of Incarceration. New York: Hill and Wang, Inc., 1976.

von Hirsch, A. The Aims of Imprisonment. Current History, July-August 1976, 71(418), 1-5.

von Hirsch, A., and Hanrahan, K. The Question of Parole. Cambridge, MA: Ballinger Publishing Company, 1979.

Wagner, H., and Pease, K. On Adding Up Scores of Offense Seriousness. British Journal of Criminology, 1978, 18, 175-188.

Walker, N. Crime and Punishment in Britain. Second revised edition. Edinburgh: Edinburgh University Press, 1970.

Walker, N. Sentencing in a Rational Society. London: Allen Lane Publishers, 1969.

Weiss, C.H. Evaluation Research: Methods of Assessing Program Effectiveness. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1972.

Weiss, C.H. The Politicization of Evaluation Research. Journal of Social Issues, 1970, 26(4), 57-68.

Wellford, C.F., and Wiatrowski, M. On the Measurement of Delinquency. Journal of Criminal Law and Criminology, 1975, 66, 175-188.

Wilkins, L.T. Social Deviance. Englewood, NJ: Prentice Hall, 1965.

Wilkins, L.T., and Chandler, A. Confidence and Competence in Decision-Making. British Journal of Criminology, 1965, 5, 22-35.

Wilkins, L.T., Kress, J.M., Gottfredson, D.M., Calpin, J.C., and Gelman, A.M. Sentencing Guidelines: Structuring Judicial Discretion. Final Report of the Feasibility Study. Albany, NY: Criminal Justice Research Center, October 1976.

Wilson, J.Q. Thinking About Crime. New York: Basic Books, 1975.

Wilson, R. Release: Should Parole Boards Hold the Key? Corrections Magazine, September 1977, 3, 47-55.

Wolfgang, M. On Devising the Crime Index, The Index of Crime: Some Further Studies. Collected Studies in Criminological Research, 1970, 7, 55-77.

Yeager, M.G. Do Mandatory Prison Sentences for Hand Gun Offenders Curb Violent Crime: Technical Report I. Washington, D.C.: U.S. Conference of Mayors, 1976.

Youngdahl, L.W. Development and Accomplishments of Sentencing Institutes in the Federal Judicial System. Nebraska Law Review, 1966, 45, 513-519.

Zalman, M. Commission of Sentencing. Notre Dame Lawyer, 1977, 53, 266-90.

Zalman, M. The Rise and Fall of the Indeterminate Sentence, Part III: The Movement Toward Definite Sentencing. Wayne Law Review, 1978, 24, 857-937.

Zalman, M., Ostrom, C.W., Jr., Guilliams, P., and Peaslee, G. Sentencing in Michigan: Report of the Michigan Felony Sentencing Project. Lansing, MI: Michigan Office of Criminal Justice, 1979.

Zeisel, H. Methodological Problems in Studies of Sentencing. Law and Society Review, 1968-1969, 3, 621-631.

Zeisel, H., and Diamond, S.S. Search for Sentencing Equity: Sentence Review in Massachusetts and Connecticut. American Bar Foundation Research Journal, 1977, 881-940.

Zimring, F. Making the Punishment Fit the Crime. Hastings Center Report, December 1976, 13-17.

Zumwalt, W.J. The Anarchy of Sentencing in the Federal Courts. Judicature, October 1973, 57(3), 96-104.