Indicators of Justice: Measuring the Performance of Prosecution, Defense, and Court Agencies Involved in Felony Proceedings

A GUIDE TO PRACTITIONERS

National Institute of Law Enforcement and Criminal Justice
Law Enforcement Assistance Administration
U. S. Department of Justice

Indicators of Justice: Measuring the Performance of Prosecution, Defense, and Court Agencies Involved in Felony Proceedings

A GUIDE TO PRACTITIONERS

SORREL WILDHORN
MARVIN LAVIN
ANTHONY PASCAL

This project was supported by Grant Number 75-NI-99-0003 awarded to The Rand Corporation by the National Institute of Law Enforcement and Criminal Justice, Law Enforcement Assistance Administration, U.S. Department of Justice, under the Omnibus Crime Control and Safe Streets Act of 1968, as amended. Points of view or opinions stated in this document are those of the authors and do not necessarily represent the official position or policies of the U.S. Department of Justice. The Law Enforcement Assistance Administration reserves the right to reproduce, publish, translate or otherwise use, and to authorize others to publish and use all or any part of the copyrighted material contained in this publication.

MAY 1977

National Institute of Law Enforcement and Criminal Justice
Law Enforcement Assistance Administration
U. S. Department of Justice

PREFACE

This report, the first in a series of two volumes, presents the results of an eighteen-month study of the use of statistical performance measures in the context of felony proceedings. The study, supported by a grant from the National Institute of Law Enforcement and Criminal Justice of the Law Enforcement Assistance Administration, U.S. Department of Justice, had two broad objectives:

- To identify, screen, and evaluate sets of statistical performance measures as indices of progress.
- To demonstrate the applicability of these performance measures in two selected (county) jurisdictions.

This report summarizes and synthesizes the approach, the methods used, and the overall findings of the study, and draws general implications for jurisdictions interested in applying the approach. It is intended as a guide to officials and practitioners in the criminal courts—judges, prosecutors, public defenders, private defense counsel, and court administrators. It should also be of interest to other criminal justice practitioners, such as police and probation officials, whose work brings them in contact with the courts.

The companion volume (R-1918-DOJ), *Indicators of Justice: Measuring the Performance of Prosecution, Defense, and Court Agencies Involved in Felony Proceedings: Analysis and Demonstration,* is a comprehensive and detailed description of all aspects of the work of this study. It includes a background discussion and literature review; professional views on performance measures; a discussion of a theoretical basis for selecting performance measures; a description of the data collection efforts; how the demonstration jurisdictions were selected; the application of selected performance measures in the two demonstration jurisdictions; the role of criminal case auditing in performance measurement; the methods, procedures, and results of surveying lay participant attitudes; and the general findings and implications of the study. R-1918-DOJ is directed primarily to the criminal justice research community —analysts, academicians, and survey researchers—but it should also be of interest to practitioners who wish to examine the details of the analysis supporting the findings of the first report.

EXECUTIVE SUMMARY

FOCUS AND PURPOSES

The primary focus of the study is on the selection, estimation, and analysis of performance measures as *statistical devices* that aid in the interpretation of data drawn from *court system operations* (i.e., from case files and other records in court, prosecution, and public defender agencies). Performance measures may be viewed either as (1) quantitative descriptors of *what* is being done in felony proceedings or (2) progress indices of *how well* these functions are being performed. (Examples of the former are the number and proportion of all felony filings that are disposed of by dismissal, plea of guilty, conviction at trial, and trial acquittal or dismissal; whereas an example of the latter is the proportion of felony trials exceeding the speedy trial standard.

This study emphasizes the latter role of performance measures. Its emphasis is to be contrasted with, for example, the application of standards and goals as articulated in the series of volumes issued since 1968 by the American Bar Association on Standards for Criminal Justice,[1] or those that resulted in 1973 from the work of the National Advisory Commission on Criminal Justice Standards and Goals. Few of these hundreds of individual goals and standards relating to criminal proceedings are couched in quantitative terms or lend themselves to quantitative interpretation, the primary exception being those concerning the "speediness" of the proceeding.

The secondary focus of this study is on performance measures of the court system as viewed through the eyes of *lay participants* in the felony proceeding—victims, other witnesses, jurors, and defendants. That is, their attitudes toward the court system are performance measures of interest, which can be elicited through survey techniques. Moreover, with proper statistical analysis, their attitudes can be related to their individual experiences with, and treatment by, the court system.

The objectives of the study were:

- To identify, screen, and evaluate sets of performance measures (to be estimated from agency records and surveys of lay participants) as *indices of progress.*
- To demonstrate the applicability of these performance measures in two jurisdictions, Multnomah County, Oregon, and Dade County, Florida.

SCOPE

The scope of the study was confined to *adult felony proceedings;* thus, we considered neither misdemeanor nor civil proceedings. We addressed only the *primary* activities of the court system, excluding supporting activities and functions performed by court clerks, court reporters, bailiffs, or paralegals in the prosecutor's office or the public defender's office. Of the several potential roles and uses of

[1] See footnote 1, page 1, for a listing of these volumes.

statistical performance measures, two broad types of applications were made: *retrospective* comparisons of the full court system and of its component agencies (rather than of individual practitioners) *within* the jurisdiction at different times;[2] and *retrospective interjurisdictional* comparisons of court systems and component agencies at the same time.

The study focused on a limited set of persistently important issue areas (described below) and sought to select and apply sets of performance measures that would clarify them.

THE ISSUE AREAS AND THE POLICY INTERESTS OI PRACTITIONERS AND AGENCIES

We selected issue areas that involve major aspects of the performance of the court system. For some issue areas—such as delay, efficiency, evenhandedness, charging accuracy, and attitudes of lay participants—practitioners and observers all agree on the direction of improvement to be sought, even though they might not agree on a structure of goals (and their relative importance) for the felony proceeding. For other issue areas—such as the charging threshold, the effect of plea bargaining, and sentence variation—they can agree only that further clarification is desirable. Also, certain agencies and types of practitioners find particular issues to be of greater interest than others, either because of relevance to their own performance or because of concerns about current policy.

Charging Standards

Prosecutors' offices in most jurisdictions need objective evidence of the standards being implemented to discern whether they conform to policy and what the effects on the system are if policy changes. For example, if the charging threshold is lowered, how are court workload, delay, and plea bargaining affected? Measures of the operation of the charging threshold over time also can reveal trends in police performance as a by-product.

Charging Accuracy

Is the nature of the disposition of cases being unduly affected by the accuracy with which charges are filed against defendants? Is the court workload being magnified by the consequences of inaccurate charging? Such questions concern not only prosecutors, but also judges, defense counsel, and court administrators in assessing charging policy and practices in their jurisdiction.

Plea Bargaining

What are the nature and frequency of the practice of plea bargaining in the jurisdiction? Quantitative evidence available to practitioners is often scant on this

[2] Comparisons were made in two contexts: when no major change (i.e., routine monitoring) is introduced, and when a major policy change or innovation (procedural, legislative, administrative) is to be evaluated.

question. And the public rarely has seen even the rudiments of an objective picture. What is the court system gaining or losing from plea negotiation? How are delay and efficiency of resource use being affected? Is punishment significantly lighter than in the absence of such negotiation? All practitioners, and the public as well, have a vital stake in these questions, even though plea bargaining policy is primarily a prosecutorial responsibility.

Sentence Variation

Judges and other practitioners want to know the degree of consistency in sentencing in their court system as compared with others, how sentencing practices change over time, and how they vary among judges within a court system. If, in addition, quantitative evidence were available to explain how much of the observed variation was accounted for by various legitimate and illegitimate factors, such information can help to reduce disparities or enhance the effectiveness of specific devices (e.g., sentencing panels or appellate review of sentences) aimed at reducing sentence disparity.

Evenhandedness

All practitioners and lay participants in the system, as well as the general public, are concerned that the courts be evenhanded in the delivery of justice, although the bench has the primary responsibility in ensuring that it occurs. If "illegitimate" factors (e.g., the defendant's ethnicity, pretrial custody status, or type of defense counsel) that should not significantly affect how cases are disposed or sentences imposed, have done so, then steps can be taken to guard against such occurrences in the future.

Delay

Although it is universally recognized that justice should be speedy, few jurisdictions have comprehensive objective evidence on the duration of their cases. They are even less prepared to isolate the effects of the separate factors (e.g., nature of offense, type of disposition, type of defense counsel, backlog problems) that may tend to delay individual cases. Although some factors that might cause delay are not readily controllable (e.g., court caseload), there are others that are (e.g., continuance policy). Better measurement of delay and its determinants can help the court administration, for example, to improve the allocation of resources.

Efficiency

Court system managers—presiding judge, court administrator, district attorney, chief of the public defender office—have responsibility for efficiency of operations. They must seek to make resources, including manpower, appropriate to the workload, given some standard for individual productivity.

Attitudes of Lay Participants

Because jurors, victims, and other witnesses represent a bridge between the court system and the general public, practitioners will be concerned that these lay

participants come away from their criminal justice experience with favorable attitudes, other things being equal. The courts can institute policies designed to enhance favorable attitudes; therefore, knowledge of the relationship between characteristics of the treatment of lay participants and their attitudes becomes critical.

Thus, each set of performance measures that we discuss in this report has an "audience" among practitioners (and often in the general public, too). Access to various sets of statistical indicators can assist in the assessment of performance, and in the design and evaluation of new policies and innovations in the myriad aspects of criminal prosecution, defense, adjudication, and sentencing.

STUDY METHODS AND SOURCES OF INFORMATION

The information used in this study was obtained from literature relevant to performance measurement; interviews with practitioners and defendants; case files in various agencies; and mail surveys of victims, witnesses, and jurors.

Practitioner Interviews

Structured interviews with 33 experienced criminal justice practitioners in 13 jurisdictions were conducted to elicit their views on the value of performance measures, the selection of issue areas and the relevant performance measures, and the choice of the two demonstration jurisdictions.

Data Collected from Agency Records and Case Files

Rand data collection teams obtained data manually on approximately 2000 cases from various records made available by officials in the two demonstration jurisdictions and at the state level. In addition, a pilot-case auditing activity was conducted in which a team of outside practitioner-consultants examined 20 burglary-type cases disposed of by plea of guilty in each jurisdiction for the purpose of making judgments about the appropriateness of decisions that were made by participants at various stages in the felony proceeding. The case-audit activity also included extensive interviews with practitioners in both jurisdictions.

Surveys of Lay Participants

Mail surveys, using questionnaires designed for this study, were administered to 1200 individuals—200 to each group of victims, other witnesses, and jurors in each jurisdiction. Questions covered attitudes, experiences, and background characteristics of these lay participants. In addition to analyzing their responses, we tested the efficacy of such questionnaires as potential tools for jurisdictions interested in measuring attitudes and determining which policy factors affect attitudes.

Personal interviews were conducted with upward of 50 defendants, split about equally between the two jurisdictions. The major purpose was to field test the interview questionnaire and contact procedures, although some data were collected on defendants' experiences and attitudes.

Once performance measures were calculated from the raw data elements, standard and specially developed software packages (i.e., computer programs) were used

to cross-tabulate the performance measures and to analyze (i.e., explain) their variation across cases.[3]

THE PERFORMANCE MEASURES

Because of space limitations, we do not list in this summary the many performance measures that were selected to illuminate each of the issue areas. The reader is referred to Chap. 2 for a discussion of (1) the selection criteria used to screen candidate sets of measures; (2) the selected sets of performance measures and the data elements necessary for computing their value; and (3) the rationale of what they can reveal, as well as conceal, about performance in each issue area.

GENERAL FINDINGS AND IMPLICATIONS

On the Feasibility of Applying Performance Measures

Our study has shown that *it is feasible to apply performance measures* to data already available in court agencies' files, even though incomplete, and *to draw inferences about whether and how performance in specified issue areas changed in a jurisdiction.* To a lesser extent, too, we have shown that it is feasible (within carefully specified limits) to make interjurisdictional comparisons of performance, using the measures specified in this study. The careful collection of specified data elements, the computation, grouping, and cross-tabulation of performance measures, and the analysis (using multivariate statistical techniques) of what factors account for the variation in key performance measures can provide greatly strengthened informational bases for officials in court, prosecution, and public defender agencies to improve criminal proceedings.

We were more successful in applying performance measures to certain policy issues than to others because of inherent differences in the precision or ambiguity of the performance measures (e.g., in those that measure changes in delay compared with those that measure changes in the charging threshold) or because of differences in the availability of data (e.g., the availability of data on sentence agreements in Multnomah County compared with its unavailability in Dade County for measuring plea bargaining effects).

The actions to be taken jointly by the court, prosecution, and public defender agencies in a jurisdiction to strengthen the informational and analytical base for measuring their performance may be visualized as an *integrated performance measurement program* (IPMP). A fairly comprehensive IPMP would consist of:

- An enumeration of required data elements (or categories) and performance measures.
- Standardized data collection and output forms for each policy issue area of

[3] For example, multivariate regression analysis was employed to isolate the independent effect of selected factors that were hypothesized to affect three key performance measures of outcomes in the felony proceeding: probability of conviction, sentence severity (given conviction), and delay (i.e., elapsed time between arraignment and final disposition).

interest (the ones we considered and/or others of interest to particular jurisdictions).

- Flexible, modular software (i.e., computer programs) packages for computing, displaying, and analyzing performance measures within each issue area (e.g., for performing cross-tabulations and for applying multivariate regression models that help to explain conviction probability, delay, and sentence severity imposed).
- Guidelines for conducting case audits at each major decision point (screening, guilty plea, trial, and sentencing) in the proceeding, using either outside practitioner-consultants or in-house supervisory personnel.
- The administration of sampling plans and standard mail survey questionnaires and the analysis of responses of victims, other witnesses, and jurors (using appropriate software packages).
- The administration of sampling plans and standard personal interview questionnaires and the analysis of responses of defendants (using appropriate software packages).

If data collection procedures and software packages were flexible and modular in design, the scale and scope of an IPMP could be tailored to individual jurisdictions. For example, the three agencies in a jurisdiction could decide whether to embrace all elements (e.g., to include case auditing and defendant interviews) and whether to measure performance in all of the listed issue areas (e.g., to include the measurement of case-processing efficiency in the prosecutor's and public defender's office, as well as in the court). What would be vital to proper tailoring is a clear enunciation by agency officials of the management and policy issues on which performance measurement should focus.

This study is a *first step* toward the design of an IPMP. We have enumerated required data elements and performance measures and, with varying degrees of success, have devised and applied statistical models to explain key performance measures. We have also designed and applied mail and personal interview questionnaires to the four classes of lay participants. More work needs to be done, however, and its nature is discussed below.

On Methodology and Data Availability

Case Audits. Our pilot-case auditing exercises (for cases in which there was a plea of guilty) in the demonstration jurisdictions strongly suggest that they provide *complementary* information about qualitative factors that aid in the interpretations of the statistical performance measures. (By their very nature, case audits are much more expensive per case included than the data collection required to develop statistical performance measures. Thus, with limited resources, audit samples are inevitably too small to stand alone as a *substitute* for statistical performance measures.)

One benefit results because the average practitioner probably regards case auditing as a natural and nontechnical way of revealing performance. His confidence in the correctness of what is shown by statistical performance measures is undoubtedly increased when the results of (even quite limited) case auditing corroborate the statistical story. Another possible benefit of case auditing is that it may help reveal the explanations for the "behavior" of statistical indices. And, finally, it may considerably strengthen the credibility of *inter*jurisdictional comparisons made by means

of statistical measures. (Our suggestions for broadening case auditing to test its value more fully are discussed below.)

Data Availability. A salient lesson in our attempt to demonstrate the application of performance measures in two selected jurisdictions was that many necessary or desirable data elements normally recorded in various files were missing from the customary records; some were simply not recorded at all. And this is likely to be the situation in other jurisdictions as well.

Among the data elements that had been (at best) *incompletely recorded* and preserved were defendant-related characteristics, such as ethnicity, prior criminal record, occupation and employment, family status, income, and transiency; the number of appearances per victim or other witness in the course of a proceeding; data describing how judges apportion their time among judicial tasks; and attribution of continuances to the responsible movant(s). However, even with incomplete recording of these data, we were successful in applying performance measures to issue areas requiring these data (with the exception of the judicial weighted caseload).

Among the data elements that were *not recorded* at all were the apportionment of time among the principal activities of prosecutors, public defenders, and jurors; background characteristics of suspects whose cases were screened out before arraignment on felony charges; full information on the outcome of plea bargaining, including the nature of any sentence agreement reached; judicial statements of the rationale for sentences in individual cases; detailed reasons for case dismissals in lower court; duration of appearances of victims and other witnesses; and information on the attitudes of lay participants and defendants toward their experiences and toward the performance of the court agencies. The unavailability of these data not only made it impossible to analyze such issue areas as the use of time by prosecutors and public defenders and evenhandedness in screening but also permitted only partial analysis of the plea bargaining balance and charging accuracy in one jurisdiction. With *special* data collection through surveys of lay participants, it was possible to assess such issue areas as the use of their time and their attitudes toward the court system.

DESIRABLE EXTENSIONS

We feel that a fuller foundation for the design of an operational IPMP would be provided by the following extensions in scope and refinement in methodology to our demonstration work:

- Classes of data that were not recorded or were incompletely recorded in Multnomah and Dade counties should be collected and analyzed elsewhere. *Evenhandedness in screening* should be analyzed with a proper body of data containing appropriate defendant-related characteristics. The *allocation of prosecutors' and public defenders' time* to their various activities is another performance area warranting examination and would need a proper body of data.
- The assessment of case auditing should be broadened in the *screening area* (to include rejected cases) and also extended to the *trial area*, so that our

inferences as to the value of case auditing as a complement to statistical performance measurement can be tested more fully.

- Improved statistical models should be constructed to help explain performance outcomes in criminal proceedings. Those we developed for explaining sentence outcomes and delay in proceedings worked fairly well but need further refinement. Because we were unsuccessful in explaining the determinants of conviction probability, we believe much more theoretical and empirical work is necessary. We speculate that data on the seriousness of the crime incident, on mitigating and exacerbating circumstances of the defendant and the crime incident, and on factors describing the strength of the case at the time of screening are relevant for constructing better conviction probability models.

How Potential Capabilities of Planned Information Systems Compare with an IPMP

One major consideration for local agencies that may be interested in moving toward an IPMP is that considerable resources already have been or will be devoted to existing or planned information systems such as CCH/OBTS, SJIS, and PROMIS.[4] It is important to know how their potential performance measurement capabilities (which issue areas can be analyzed in what depth?) compare with an IPMP under two conditions: (1) the basic systems with only those data elements that are already collected, assuming that simple software packages (with a cross-tabulation capability) are available; (2) modest, inexpensive upgrading of the basic systems (by adding a few new data elements[5] to be collected, together with a more sophisticated software package, for example, statistical models and standard multivariate statistical analysis routines for estimating the independent effect of important factors on delay and sentence severity imposed).

Capabilities of the Basic Systems Planned.[6] Given the data elements collected by these information systems and the availability of (at best) simple software packages, all of the systems have a valuable capability for measuring performance in the delay estimation and charging accuracy issue areas, and a partial capability in the areas of plea bargaining, sentence variation, evenhandedness, and determinants of delay. *In addition,* SJIS is capable of very gross estimates of the use of judicial time, and PROMIS has a good capability in the charging threshold area and a partial capability in addressing evenhandedness in screening (whereas CCH/OBTS and SJIS have no capabilities in the screening area).

Capabilities If Planned Systems Are Upgraded. Upgrading any of the basic systems (as noted above) would enable better analysis of the plea bargaining balance, the independent effects of important factors on delay, and the independent effects of legitimate and illegitimate factors on sentence severity imposed for all the systems. *In addition,* upgrading of PROMIS would improve the capability to analyze evenhandedness in screening.

[4] See Chap. 5 for definitions and descriptions of these information systems.

[5] See the footnotes to Tables 9 and 10, Chap. 5, for the few additional data elements that can be collected inexpensively.

[6] We assess system capabilities in terms of the issue areas addressed in this study; capabilities of these systems to address *other* issue areas are not assessed.

Capabilities of an Improved IPMP. If an IPMP were improved and extended in the ways noted above, its performance measurement capabilities would have greater breadth, because many data elements specified for it are not collected by the basic (or upgraded) existing or planned systems. Although each system could function as a partial IPMP, *none* of the systems are designed to measure (as would an IPMP) performance in the following areas: the effect of legitimate factors on conviction probability; continuances (except for PROMIS); the use of lay participant (victim, witness, juror) time; the use of practitioner (judge, prosecutor, public defender) time; and the attitudes (and their determinants) of lay participants.

The Costs and Utility of Various Information Systems

Careful estimates of the range of incremental costs for implementing and operating a partial or full (improved) IPMP or of upgrading existing or planned systems were beyond the scope of this study. However, based on actual resources used in various activities of this study and on rough guesses of costs of activities not covered in this study, we can *bound* the range of likely costs within, say, a factor of two.

For a jurisdiction *with* one of the existing or planned information systems, *incremental* (i.e., over and above the costs of the basic system) annual costs on the order of $10,000 might be incurred for upgrading the system and assessing performance annually. This assumes that appropriate software packages are made available free and that any practitioner time devoted to additional raw data generation is "free." Given the relatively low marginal costs associated with upgrading an existing or planned system and its major benefits outlined above, it is probably cost-effective for a jurisdiction to pursue this alternative.

For a jurisdiction *without* an existing information system (but with access to a computer) that wishes to implement and operate a fully improved IPMP, on the order of $50,000 per year in operating costs are implied, once it is set up. (First-year costs should be considerably higher because of nonrecurring setup costs.) This rough estimate assumes (as with the previous case) that software and practitioner time are free and that the number and size of case file and survey response samples to be collected and analyzed are similar to those collected and analyzed in this study. Of course, additional samples or larger samples would increase costs, and exclusion of certain issue areas from an IPMP would reduce costs.

Whether implementing a full IPMP has adequate utility—that is, whether incremental benefits sufficiently outweigh incremental costs—is a judgment that can be made by an implementing jurisdiction only after such an approach is installed and operated over several years. At that point, the costs will be much less uncertain and its benefits can be assessed by the policymakers involved.

APPLICATION OF PERFORMANCE MEASURES

In the interests of brevity, we do not summarize the findings obtained from the applications of performance measures in the two demonstration jurisdictions, but refer the reader to Chaps. 3 and 4 of this report and Secs. VI, VII, IX, and X of the accompanying report (R-1918-DOJ).

ACKNOWLEDGMENTS

We are especially indebted to members of our Study Advisory Group, who helped substantially in giving direction to our study efforts and in commenting on earlier drafts of our reports. In addition, they suggested many of the experienced criminal justice practitioners we interviewed during the study. The Advisory Group members were District Attorney Louis Bergna (Santa Clara County, Calif.); Professor Al Blumstein (Carnegie-Mellon University); The Honorable Winslow Christian (Justice, California Court of Appeals); District Attorney Harry Connick (New Orleans, La.); John Flynn (Attorney, Phoenix, Ariz.); Edward McConnell (Director, National Center for State Courts); Public Defender Sheldon Portman (Santa Clara County, Calif.); Harvey Solomon (Executive Director, Institute for Court Management); Professor Harry Subin (New York University Law School, formerly Office of U.S. Attorney); District Attorney Preston Trimble (Cleveland County, Okla.); The Honorable Ernst Watts (Dean, National Center for State Judiciary); and Jerry Wilson (Chief, Metropolitan Police Department (retired), Washington, D.C.).

The names of the many individuals who often granted us considerable time in which to interview them about their experiences and views as criminal justice practitioners as these related to our study are listed in App. B of the companion report. We are deeply appreciative of their cooperation.

Our gratitude extends to the many officials of the court systems and other agencies in Multnomah County, Oregon, and Dade County, Florida, who made possible our demonstration work in these jurisdictions, aided in collecting and interpreting data from their files, and commented on earlier drafts of our reports. In Multnomah County, they included: Harl Haas, District Attorney; J. G. McClain, Chief Deputy; Senior Deputies M. D. Shrunk, K. C. Frankel, and J. E. Grayson; Deputies J. C. Ray and Kelly Bacon of the District Attorney's Office; The Honorable C. B. Olsen, Presiding Judge; Judges John Beatty and John Jones of the Circuit Court; Michael Hall, Court Administrator, and James Murchison, Assistant Court Administrator of the Circuit Court; Charles Barnard, Chief Criminal Clerk, and Gwen Biehre of the Office of the Chief Criminal Court; James Hennings, Metropolitan Public Defender, Robert Lockwood, Assistant Public Defender, and Samuel Imperati, Executive Officer; O. R. Chambers, Executive Assistant, and Louis Lewandowski, Manager, ADP Services of the Oregon Division of Corrections; Commander Kerner, Rocky Butte Jail, Multnomah County; and Sheriff Lee Brown, Multnomah County.

In Dade County they included: Richard Gerstein, State Attorney, Janet Reno, First Assistant State Attorney, and James Regan, Administrative Manager, State Attorney's Office; The Honorable Grady Crawford, Chief Judge, The Honorable Gene Williams, Administrative Judge of the Criminal Division, Edward Coward, Judge of the Criminal Division, and Seymour Gelber, Judge of the Juvenile Division, Circuit Court, Eleventh Judicial Circuit of Florida; Wilbur McDuff, Court Executive Officer, Richard Brinker, Clerk of the Circuit and County Courts, James Sykes, Operations Officer of the Magistrate's Division, Montague Eadie (formerly Operations Officer of the Criminal Division), Circuit Court, and Cheryl Johnson of the Pretrial Release Program; Phillip Hubbart, Public Defender, Richard Gillen,

Business Manager, and David Weed, Office of the Public Defender; Captain Gallagher, Director, Dade County Jail, and G. R. Worley, Administrator, Research and Statistics Section, Florida Department of Offender Rehabilitation.

Invaluable efforts were contributed by the members of our data collection teams. In Multnomah, they were either students from the Northwest School of Law (Lewis and Clark College) or students in Administration of Justice at Portland State University. The team included: M. Craven, C. Francis III, C. Henry, M. Ketelle, J. Sokol, and K. St. Amant. In Dade, they were students from the University of Miami School of Law (or, in one instance, a graduate who was already a member of the Florida Bar). The team included: A. Cohen, J. Kaiser, K. Rizzo, J. Sloto, D. Weinberger, and J. Wulchak.

Interviews with defendants in Multnomah County were conducted by K. Bobzein, M. Craven, and C. Henry, all students in Administration of Justice at Portland State University. In Dade County, they were conducted by T. McDonald, J. Taylor, and D. Weinberger, students or recent graduates from the University of Miami School of Law.

We are obligated to William Hutchins, California Bureau of Criminal Statistics, for materials that are a foundation for the Sentence Severity Indices devised in our study.

We are indebted to the respondents who gave us their time to complete the mail surveys and personal interviews described in both reports.

Very helpful criticisms of the survey instruments were given by Dr. Jonathan Casper and Arthur Morganstern. Many items in the questionnaires were derived from earlier questionnaires by Dr. Casper and Dr. Mary Queeley.

Especially deserving of mention for their contributions are Professor Harry Subin, New York University Law School, who reviewed an earlier draft; and a Rand consultant, Professor Peter Haynes, Criminal Justice Department, Arizona State University, who assisted during the formulation of the research tasks.

Many of our Rand colleagues aided us in varied and important ways. Bryant Mori and Edward Woo performed the vital computer programming for the analysis of the case file and survey data collected in the demonstration jurisdictions. Stephen J. Carroll, Charles E. Phelps, Daniel Relles, and John E. Ware provided professional guidance in resolving statistical problems that arose in the course of our work. Gene H. Fisher, Robert M. Gay, Peter W. Greenwood, D. Michael Landi, and Harriet B. Stambul served as technical reviewers of drafts of these reports. William R. Ahern and Robert J. Gladstone assisted during the formulation of the research tasks. Deborah Hensler contributed to the overall selection of strategies and design of procedures for the lay participant surveys; Marilyn Fisher managed the distribution of the mail surveys and ensured a satisfactorily high rate of responses; and Donald P. Trees, of Survey Data Processing, provided valuable advice and strong support in the cleaning and processing of survey data.

Finally, we are especially indebted to Cheryl Martorana and Voncile Gowdy of the National Institute of Law Enforcement and Criminal Justice of LEAA for reviewing previous drafts of the reports and for suggesting a variety of improvements.

CONTENTS

PREFACE iii

EXECUTIVE SUMMARY v

ACKNOWLEDGMENTS xv

TABLES xix

Chapter

1. INTRODUCTION 1
   - Overview 1
   - Roles and Uses of Performance Measures 2
   - Purposes and Scope of the Study 3
   - Relationships between Issue Areas and the Policy Interests of Practitioners and Agencies 4
   - Study Methods and Sources of Information 6
   - A Guide to the Study Reports 8

2. THE PERFORMANCE MEASURES AND REQUIRED DATA ELEMENTS 10
   - General Selection Criteria 10
   - Case Screening: Charging Standards or Threshold 11
   - Case Screening: The Accuracy of Charging 11
   - Performance Measures in Plea Bargaining 14
   - Measuring and Explaining Sentence Variation 17
   - Performance Measures of Evenhandedness 17
   - Performance Measures of Delay and Use of Lay Participant Time 19
   - Performance Measures of the Use of Practitioner Time 21
   - Measuring Lay Participant Attitudes and Their Determinants 24
   - Gauging Overall Performance 25

3. APPLICATION OF PERFORMANCE MEASURES IN ONE JURISDICTION 27
   - Did Case Quality and Charging Standards Change for Impact Offenses? 28
   - Did Charging Accuracy Increase for Impact Offenses? 28
   - Plea Bargaining Changes: Were the Experiment's Objectives Achieved and Did the Balance Shift? 29
   - How Did Sentencing Variation Change? 30
   - Were Dispositional and Sentence Outcomes More Evenhanded? 30
   - How Were "Habitual Offenders" Treated? 31
   - Was Justice Swifter? 31
   - The Use of Jurors and Witnesses 32
   - Lay Participant Attitudes 32
   - Gauging Overall Performance 33

4. COMPARING PERFORMANCE IN TWO JURISDICTIONS 35
Case Screening 35
Plea Bargaining 36
Sentence Variation 36
Evenhandedness 36
The Role of Prior Record 37
Delay 37
Use of Victims, Other Witnesses, and Jurors 38
Lay Participant Attitudes 38
Overall Performance 39

5. GENERAL FINDINGS AND IMPLICATIONS OF THE STUDY 40
The Desirability of an Integrated Performance Measurement Program (IPMP) 40
Data Availability and Methodology 42
Desirable Extensions 43
Potential Capabilities of Other Information Systems for Performance Measurement 44

TABLES

1. Performance Measures in Case Screening: Charging Standards or Threshold 12
2. Performance Measures in Case Screening: Charging Accuracy 13
3. Performance Measures in Plea Bargaining: The Frequency of Plea Bargaining 15
4. Performance Measures in Plea Bargaining: Balance of Systemwide Effects and Systemwide Concessions to Defendants 16
5. Measuring Evenhandedness in Dispositional and Sentence Outcomes 19
6. Peformance Measures of Delay and the Use of Lay Participant Time 22
7. Classifications of Judicial Activities in Criminal Proceedings 23
8. A Comparison of Data Elements and Categories Collected That Are Relevant to Performance Measurement 46
9. Potential Performance Measurement Capabilities of CCH/OBTS .. 48
10. Potential Performance Measurement Capabilities of PROMIS (with or without CCH/OBTS, SJIS) 50
11. Rough Cost Estimates for Performance Measurement 54

Chapter 1

INTRODUCTION

OVERVIEW

This report, intended as a guide to practitioners, summarizes and synthesizes the overall findings of a broad study of performance measurement of criminal justice agencies involved in the felony proceeding—post-arrest through disposition (and sentencing if it occurs). The primary focus of the study is on the selection, estimation, and analysis of performance measures as *statistical devices* that aid in the interpretation of data drawn from *court system operations* (i.e., from case files and other records in court, prosecution, and public defender agencies). Performance measures may be viewed either as (1) quantitative descriptors of *what* is being done in felony proceedings or (2) as progress indices of *how well* these functions are being performed. Examples of the former are the number and proportion of all felony filings that are disposed of by dismissal, plea of guilty, conviction at trial, and trial acquittal or dismissal; whereas an example of the latter is the proportion of felony trials exceeding the speedy trial standard.

This study emphasizes the latter role of performance measures. Its emphasis is to be contrasted with, for example, the application of standards and goals as articulated in the series of volumes issued since 1968 by the American Bar Association on Standards for Criminal Justice,[1] or those that resulted in 1973 from the work of the National Advisory Commission on Criminal Justice Standards and Goals.[2] Few of these hundreds of individual goals and standards relating to criminal proceedings are couched in quantitative terms or lend themselves to quantitative interpretation, the primary exception being those concerning the "speediness" of the proceeding.

The secondary focus of this study is on performance measures of the court system as viewed through the eyes of *lay participants* in the felony proceeding—victims, other witnesses, jurors, and defendants. That is, their attitudes toward the court system are performance measures of interest, which can be elicited through survey techniques. Moreover, with proper statistical analysis, their attitudes can be related to their individual experiences with, and treatment by, the court system.

In this report we describe the roles or uses of performance measures; the broad purposes and scope of the study; the relationships between the issue areas considered in this study and the interests of agencies and practitioners responsible for making policy; general criteria for selecting performance measures; the performance measures themselves and how they illuminate the issue areas; the data

[1] The individual volumes (and their dates of approval) include *ABA Standards Relating to Pretrial Release* (1968); *Providing Defense Services* (1966); *Fair Trial and Free Press* (1968); *Pleas of Guilty* (1968); *Speedy Trial* (1968); *Joinder and Severance* (1968); *Trial by Jury* (1968); *Sentencing Alternatives and Procedure* (1968); *Appellate Review of Sentences* (1968); *Post-Conviction Remedies* (1968); *Discovery and Procedure before Trial* (1970); *Probation* (1970); *Criminal Appeals* (1970); *Electronic Surveillance* (1971); *The Prosecution Function and the Defense Function* (1971); *The Function of the Trial Judge* (1972); *The Urban Police Function* (1973), *Court Organization* (1974); and *Trial Courts* (tentative draft, 1975).

[2] The six individual reports of the National Advisory Commission are entitled *A National Strategy To Reduce Crime, Criminal Justice System, Courts, Police, Corrections,* and *Community Crime Prevention* (1973).

elements necessary to estimate them; the techniques used to estimate, display, and explain changes in performance measures; the inferences (as well as the qualifications or ambiguities inherent in these inferences) that can be drawn from illustrative applications made in two (county) jurisdictions; the general implications of, and lessons learned from, the study; and the need for certain extensions to this study.

Finally, following a "tri-level approach," we illustrate for jurisdictions that may be interested in applying our performance measurement approach:

1. Which issue areas can be analyzed using information systems *currently installed or planned in some jurisdictions.*
2. To what extent a modest and inexpensive *extension in data elements collected by these systems* could improve the scope and depth of performance measurement capabilities.
3. Those performance measurement applications to issue areas that would require *new* (and more costly) data collection and analysis efforts, because current or planned information systems do not suffice for such applications.

It must be emphasized that certain software packages (such as those applied in this study) would be required for analyzing performance, *whichever* of the three alternatives above is pursued.

ROLES AND USES OF PERFORMANCE MEASURES

The potential roles and uses of statistical performance measures in the felony proceeding can be categorized as follows:

Within a Jurisdiction

1. Routine tracking or administrative monitoring of *pending* cases in the prosecution, defense, and court agencies.
2. *Retrospectively* comparing the performance of the full court system, its component agencies, and its individual practitioners at different times:
   a. When no major policy change or innovation is introduced.
   b. When a major policy change or innovation (procedural, legislative, administrative) is to be evaluated.
3. *Prospectively* estimating the performance effects of a major policy change or innovation.

Among Jurisdictions

1. *Retrospectively* comparing the full court systems and component agencies in different jurisdictions:
   a. Comparing different jurisdictions within a state (e.g., of interest to a state-level judicial council or supreme court).
   b. Evaluating whether a particular policy or innovation in one jurisdiction has similar effects in another.
2. Helping to show the condition of the criminal justice process at the *state* and *national* levels.

Other

1. Prompting and guiding research on ways to enhance the administration of justice in felony proceedings.

Even though the collection and use of statistical data describing the operations of court, prosecution, and defense agencies have increased over the last decade, the major use of these data has been to provide (often rudimentary) assessments of how speedily and efficiently a jurisdiction disposes of its pending caseload. There has been less inclination to employ statistical indicators to measure other aspects of felony proceedings relating to the quality of justice, notwithstanding the upsurge in production of (largely qualitative) standards and goals bearing on these proceedings, as mentioned above. Basically, statistical descriptors have been widely used to depict *what* is going on in felony proceedings. Thus, there is a marked gap between the articulation of (largely qualitative) goals and standards on one hand and the measurement of progress toward goals on the other. Given this observation, the broad purpose of our study was to reduce that gap.

PURPOSES AND SCOPE OF THE STUDY

The specific purposes of the study were:

- To identify, screen, and evaluate sets of performance measures as *indices of progress.*
- To demonstrate the applicability of these performance measures in two selected (county) jurisdictions.

As "outside" analysts, we could, at best, aim to demonstrate the *feasibility of applying* performance measures. Whether this application is "practical" and whether the benefits of applying performance measures would outweigh the incremental costs can thereafter be assessed by having officials in one or more jurisdictions adopt the approach, use it over some period of time, and then make the necessary cost/benefit judgments. Incremental costs of applying the full range of performance measures would vary from jurisdiction to jurisdiction, depending on the type of information system that was planned or already installed and the extent to which the jurisdiction would desire to measure performance in issue areas *outside* of the capabilities of existing information systems. Careful estimates of the range of such incremental costs were beyond the scope of this study. However, we provide (in Chap. 5) very rough incremental cost estimates for two bounding cases: a gross estimate of costs for implementing a relatively comprehensive performance measurement system from scratch, assuming that a jurisdiction has access to a computer but has not installed (or does not plan to install) an information system; and a gross estimate of the incremental cost associated with "upgrading" (i.e., adding a few data elements that can be collected inexpensively and special software packages) existing or planned information systems such as CCH/OBTS, SJIS, or PROMIS.[3]

Resource limitations necessarily limited the scope of this study. Attention was confined to *adult felony proceedings;* thus, we did not consider misdemeanor proceed-

[3] See Chap. 5 for definitions and descriptions of these systems.

ings or civil proceedings. We addressed only the *primary* activities of the court system, excluding supporting activities and functions performed by court clerks, court reporters, bailiffs, or paralegals in the prosecutor's office or the public defender's office. Two broad types of performance applications were made: *retrospective* comparisons of the full court system and of its component agencies (rather than of individual practitioners) *within* the jurisdiction at different times; and *retrospective interjurisdictional* comparisons of court systems and component agencies at the same time. To the extent that our statistical modeling was successful in predicting and explaining certain performance measures of outcomes, we developed a very limited capability to do *prospective* performance analysis.

The study focused on a set of persistently important issue areas. Our view of their importance was confirmed by the results of personal interviews with 33 practitioners (judges, prosecutors, defense counsel, court administrators, and legal scholars) in 13 large urban court systems throughout the United States and by the views of an advisory panel of distinguished practitioners. These issue areas were:

- Prosecutorial (or other) case screening: limited to the subissues of adherence to charging standards and charging accuracy.
- Plea bargaining: viewed as a balance between gains to, and other "operational effects" (some possibly harmful) on, the court system on one hand and system "concessions" to defendants on the other.
- Sentencing variation: how much variation; to what extent "legitimate" as opposed to "illegitimate" factors[4] explain the variation.
- Evenhandedness or consistency of disposition and sentencing: to what extent such outcomes are affected by illegitimate factors.
- Delay (or measures of elapsed time) between major events in the felony proceeding.
- Case processing efficiency (as reflected in the use of judicial, prosecutorial, and defense counsel time) and the use of lay participant (jurors, victims, and other witnesses) time.
- Attitudes of lay participants toward the court system and its practitioners.

RELATIONSHIPS BETWEEN ISSUE AREAS AND THE POLICY INTERESTS OF PRACTITIONERS AND AGENCIES

The issue areas considered in this study are important in the administration of criminal justice because they involve significant aspects of the performance of the court system. For some of these issue areas—delay, efficiency, evenhandedness, charging accuracy, attitudes of lay participants—practitioners and observers all agree on the direction of improvement to be sought. For others—the charging threshold, the effects of plea bargaining, sentence variation—they can agree only that further clarification is desirable. Also, certain agencies and types of practi-

[4] There is considerable controversy over whether *certain* factors are legitimate or illegitimate in sentencing decisions; for others, there is general agreement. For purposes of this study, we have assumed that the defendant's age, prior criminal record, community ties, and the nature of the original and convicted charges and counts are legitimate factors; and that ethnicity, pretrial custody status, type of defense attorney, type of disposition (trial or guilty plea), and correctional facilities crowding are illegitimate factors. See the discussion in Chap. 2.

tioners find particular issues to be of greater interest than others, either because of relevance to their own performance or because of concerns about current policy and how current policy effects compare with effects of past policy.

Charging Standards

Prosecutors' offices in most jurisdictions need objective evidence of the standards being implemented to discern whether they conform to policy and what the effects on the system are if policy changes. For example, if the charging threshold is lowered, how are court workload, delay, and plea bargaining affected? Measures of the operation of the charging threshold over time also can reveal trends in police performance.

Charging Accuracy

Is the nature of the disposition of cases being unduly affected by the accuracy with which charges are filed against defendants? Is the court workload being magnified by the consequences of inaccurate charging? Is inaccuracy in the charging process abusing defendants' rights? Such questions concern not only prosecutors, but also judges, defense counsel, court administrators, and others in addressing charging policy and practices in their jurisdiction. .

Plea Bargaining

What are the nature and frequency of the practice of plea bargaining in the jurisdiction? Quantitative evidence available to practitioners is often scant on this question. And the public rarely has seen even the rudiments of an objective picture. What is the court system gaining from plea negotiation? How are delay and efficiency of resource use being affected? Is punishment significantly lighter than in the absence of such negotiation? All practitioners, and the public as well, have a vital stake in these questions, even though plea bargaining policy is primarily a prosecutorial responsibility.

Sentence Variation

Judges and other practitioners want to know the degree of consistency in sentencing in one court as compared with others, how sentencing practices change over time, and how they vary among judges within a court system. If, in addition, quantitative evidence were available to explain how much of the observed variation was accounted for by various legitimate and illegitimate factors, such clarification may enhance the effectiveness of various devices (e.g., sentencing panels or appellate review of sentences) aimed at reducing sentence disparity.

Evenhandedness

All practitioners and lay participants in the system, as well as the general public, are concerned that the courts be evenhanded in the delivery of justice, although the bench has the primary responsibility in ensuring that it occurs. If illegitimate factors (e.g., the defendant's ethnicity, pretrial custody status, or type

of defense counsel) have significantly affected how cases were disposed or the sentences imposed, steps can be taken to guard against such occurrences in the future.

Delay

Although it is universally recognized that justice should be speedy, few jurisdictions have comprehensive objective evidence on the duration of their cases. They are even less prepared to isolate the effects of the separate factors (e.g., nature of offense, type of disposition, type of defense counsel, backlog problems) that tend to delay individual cases. Although some factors that might cause delay are not readily controllable (e.g., court caseload), there are others that are (e.g., continuance policy). Better measurement of delay and its determinants can help the court administration, for example, to improve the allocation of resources.

Efficiency

Court system managers—presiding judge, court administrator, district attorney, chief of the public defender office—have responsibility for efficiency of operations. They must seek to make resources, including manpower, appropriate to the workload, given some standard for individual productivity.

Attitudes of Lay Participants

Because jurors, victims, and other witnesses represent a bridge between the court system and the general public, practitioners will be concerned that these lay participants come away from their criminal justice participation with favorable attitudes, other things being equal. Lay participants *ought* to feel that justice was done, that the system performs effectively, that they themselves were treated satisfactorily, and they ought to be willing to cooperate again in the future. Court agencies can institute policies designed to enhance favorable attitudes; therefore, knowledge of the relationship between characteristics of the treatment of lay participants and their attitudes becomes critical.

Thus, each set of performance measures we discuss in this report has an "audience" among practitioners and often in the general public. Access to various sets of statistical indicators can assist in the assessment of performance, the design of new policies, and the evaluation of innovations in the myriad aspects of criminal prosecution, defense, adjudication, and sentencing.

STUDY METHODS AND SOURCES OF INFORMATION

In the initial period of the study, we constructed a hierarchy of recognized goals of the criminal justice system as a whole, and of felony proceedings as a whole. This goal structure was intended to be the framework to which the individual performance measures would be related. They would then be assembled into sets that would assess how closely the proceedings in a jurisdiction approached the goals.[5] In an early review by our Advisory Group (made up of distinguished jurists, prosecu-

[5] For a discussion of goals and their links to individual performance measures, see Sec. II, R-1918-DOJ.

tors, defense attorneys, law professors, analysts, and experts in court administration),[6] it was clear that no general consensus among practitioners on any specific goal structure was forthcoming. Thus, a "fail-safe" quality was lacking: If any specific goal structure was not accepted, the credibility of this study would suffer. In response, we reshaped the study's scope by focusing on several persistently important areas (noted above) to which we could apply performance measures to assess at least the desired direction of movement.

A series of interviews with experienced criminal justice practitioners and the analysis of their responses comprised the second phase of this study. There were six interviews with judges, seven with court administrators, five with defense counsel, six with prosecutors, and two with academicians—the interviews being distributed in 13 jurisdictions across the country.[7] The interviews, which ranged from several hours to a full day in duration, informed us about the views of experienced practitioners toward the use and value of performance measures, the selection of issue areas and the relevant performance measures, and focused our consideration on the choice of the two demonstration jurisdictions. The results of these interviews were as follows:

- There was some controversy as to the value and acceptability of statistical performance measures standing by themselves; the more experience a practitioner had had with statistical data, the less he distrusted them and the more realistic he was about their use. Many suggested—and we implemented—a complementary (to the mainstream effort of statistical performance measurement) pilot approach we call "case auditing." We asked a team of experienced consultant-practitioners (in this instance, prosecutors) to audit 20 burglary-type cases disposed of by pleas of guilty in each jurisdiction. They made judgments about the appropriateness of decisions and actions by practitioners at various stages in the felony proceeding, given the information available to them at the time.
- The issue areas that we selected for analysis were demonstrated to indeed be important, and there was no consensus on *additional* ones to be analyzed.
- Our choice of performance measures was enriched, and we expanded the check list of "comparability features" that govern the validity of *interjurisdictional* comparisons by means of performance measures.[8]
- We were led to choose Multnomah County (City of Portland), Oregon, and Dade County (City of Miami), Florida, as cooperating jurisdictions in the demonstration phase.

In the demonstration phase, we collected operational data from records in the prosecution, public defender, and court agencies and survey data (through mail questionnaires and personal interviews) from lay participants in both jurisdictions. The information collected is given below.

[6] See App. B, R-1918-DOJ, for a roster of the Advisory Group members.

[7] See App. B, R-1918-DOJ, for identification of the interviewees. The results of the interviews are given in Sec. III.

[8] See Sec. IX, R-1918-DOJ.

From Agency Records and Case Files

Rand data collection teams obtained data directly from a variety of agency records and case files made available by officials in the demonstration jurisdictions and at the state level.[9] These included 1200 cases filed in felony court (samples of 100 each of burglary or breaking and entering, robbery, and all felonies in each jurisdiction in each of two years) containing data on the nature and number of original and convicted charges, plea bargaining information, dates of major events in the case, disposition, sentence, and a variety of defendant-related characteristics. In addition, separate samples were collected for continuances and for victim and other witness appearances, where necessary, as were samples for screening actions (100 each of police-booked burglary and robbery cases in two years in both jurisdictions) and for rejection reasons (samples or census of burglary and robbery "rejections" or "no-information" in two years in both jurisdictions). In addition, a pilot-case audit activity was conducted as noted above, including extensive interviews with practitioners in both jurisdictions.

From Surveys of Lay Participants

Mail Surveys. Using questionnaires designed for this study, 1200 individual mail surveys were administered—200 to each group of victims, other witnesses, and jurors in each jurisdiction. Questions covered attitudes, experiences, and background characteristics of these lay participants. In addition to analyzing their responses, we wanted to test the efficacy of such questionnaires as potential tools for jurisdictions interested in measuring attitudes and determining which policy factors affected attitudes.[10]

Personal Interviews. We interviewed somewhat less than 50 defendants split about equally among the two jurisdictions. The major purpose was to field test the interview questionnaire and contact procedures, although some data were collected on defendants' experiences and attitudes.[11]

Once performance measures were calculated from the raw data elements, standard and specially developed software packages (i.e., computer programs) were used to cross-tabulate the performance measures, to estimate whether observed changes were statistically significant, and to analyze the performance measure variation across cases. For example, multivariate regression analysis was employed to uncover the independent effect of selected factors that were hypothesized to affect performance measures of outcomes in the felony proceeding. These outcome measures included probability of conviction, sentence severity (given conviction), and delay (i.e., elapsed time between arraignment and final disposition).

A GUIDE TO THE STUDY REPORTS

R-1917-DOJ (A Guide to Practitioners)

Chapter 2 discusses the selection criteria used to screen performance measures,

[9] For a complete description of the methods and sources used to collect these data, see App. D, R-1918-DOJ.

[10] For a complete description of the methods and results of the mail surveys, see App. F, R-1918-DOJ.

[11] For a complete description of the methods and results of the defendant interviews, see App. I, R-1918-DOJ.

lists the sets of performance measures selected for each issue area together with the data elements needed to compute their values, and provides a brief rationale on what the measures reveal, as well as conceal, about performance. Chapter 3 illustrates one of the roles of performance measures—an application in one jurisdiction to illuminate how performance changes from year to year. Chapter 4 illustrates another role of performance measures—an application comparing performance in two jurisdictions in one year. Finally, Chap. 5 presents the general findings and implications that emerged from this study.

R-1918-DOJ (Analysis and Demonstration)

The lengthy companion report provides a comprehensive and detailed description of all aspects of the work of this study. Section I discusses the roles and uses of performance measures, the study purposes, and scope and provides an overview of the methods and sources of information used. Section II provides a background discussion and a limited literature review. Section III is a discussion of professional views on statistical performance measures as gathered from a series of interviews with practitioners. Section IV discusses the selected performance measures that were found useful. It includes a description of the selection criteria used to screen and select candidate sets of performance measures; a brief rationale on what the measures reveal, as well as conceal, about performance in each of the selected issue areas; and a description of the data elements necessary for computing values of the performance measures.

Section V provides our rationale for selecting the two demonstration jurisdictions and a (largely qualitative) description of the component agencies of the court systems in both jurisdictions. Sections VI and VII apply the performance measures to each issue area in Multnomah and Dade counties, respectively, to illuminate how performance changes from year to year. Section VIII discusses the role of criminal case auditing in performance measurement in general terms and then describes the results of a pilot application made in the two jurisdictions. Section IX compares performance in the two jurisdictions in one year. Section X discusses the general procedures and results of the mail surveys of victims, other witnesses, and jurors and of the defendant interviews. Finally, Sec. XI, which is identical with Chap. 5 of the *Guide to Practitioners*, presents the general findings and implications that emerged from this study.

Various appendixes discuss the links between goals and performance measures of criminal proceedings (App. A), various considerations for making interjurisdictional comparisons by means of statistical performance measures (App. C), data collection methods and sources used in the two jurisdictions (App. D), and the mail survey procedures used (App. F). Appendix B lists the criminal justice practitioners interviewed and the Study Advisory Group members. Others display results of the statistical analyses of the case file data (App. E) and of the lay participant survey data (Apps. H and I), and show representative questionnaire instruments developed and used in the study (Apps. G and I).

Chapter 2

THE PERFORMANCE MEASURES AND REQUIRED DATA ELEMENTS

In this chapter we describe the selection criteria used to screen candidate sets of performance measures and we list the selected measures of performance that were found useful. We also provide a brief rationale on what the measures reveal, as well as conceal, about performance in one selected issue area, as an illustration of the complexities involved. The data elements necessary for computing the values of the performance measures[1] are indicated with each set of measures used in the selected issue areas.

Our discussion here is, almost in its entirety, from the *intrajurisdictional* point of view, that is, assessing changes in performance over time within a jurisdiction. Interjurisdictional comparisons involve not only the differences between values of the performance measures in, say, two jurisdictions, but also the disparities in the nature of the two systems. Some of these disparities will be touched on in Chap. 4, where performance comparisons between the two demonstration jurisdictions are illustrated, but a more complete list of "comparability features" that govern the validity of interjurisdictional comparisons is given in the companion report.[2]

GENERAL SELECTION CRITERIA

Selecting the sets of measures to illuminate performance in the selected issue areas was a pivotal activity in this study. This required us to set forth, as has not generally been done before, a group of criteria for screening candidate measures. In our view, a performance measure (and a set of these measures) should be more or less preferred, depending on the degree to which it is:

- *Relevant* to the matter that requires measurement.
- *Proximate* to the events whose occurrences are being measured.
- *Directly linked* (i.e., strongly correlated) with the performance area of interest.
- *Applicable* to the analytical task undertaken.
- *Specific* rather than ambiguous in its informational content.
- *Clear* in meaning to the practitioner.
- *Consistent* with other members of a set of measures in producing inferences.
- *Complementary* to other members of a set of measures in contributing the intended information.
- *Capable of being implemented* in terms of the cost and availability of data required to support the measure.

[1] For a more detailed discussion, see Sec. IV, R-1918-DOJ.

[2] See Sec. IX, R-1918-DOJ.

CASE SCREENING: CHARGING STANDARDS OR THRESHOLD

The set of performance measures we found useful in assessing whether and how the charging threshold is changing over time within a jurisdiction is listed in Table 1, together with the raw data elements needed to estimate values. The gross output of case screening, whether done by the prosecutor's office or the court, results in one of three decisions: to file on the most serious charge booked by the police, to file on charges of lesser gravity (and this may be further broken down by level of seriousness), or to reject unconditionally.[3] If, say, charging standards applied to a certain offense category shift over time from a "probable cause" standard toward a trial conviction standard of "proof beyond a reasonable doubt," and if police arrest policies and booking standards remain unchanged, one would expect to see a rise in the rejection rate and in the filing rate at lesser levels and a fall in filing rate at the booked-charge level. However, if the police change their arrest practices as well, and, say, respond to the change in the prosecutor's charging standards by presenting weaker cases less frequently, then changes in these gross measures of prosecutorial screening output are much more ambiguous and more difficult to interpret. Thus, one must know something about whether and how much police arrest policies have changed in order not to confound police and prosecutorial factors.

Additional insights regarding the possible abuse of discretion to reject and the quality of cases submitted by the police can be gained from an examination of the reasons for unconditional rejections, if the reasons are accurately recorded. First, some rejections may not be unconditional; some cases rejected and returned to the police for need of more investigation may reenter the system if the police complete the investigation. If this class of rejections is large or major changes are observed over time, it is essential to track cases rejected by reason of needing more investigation. Second, if the frequency of rejections "in the interests of justice" rises dramatically, it may signal the presence of abuse of discretion, since this broad undifferentiated reason may be used as camouflage. If there are fewer rejections for inadequate evidence and more for recognized specific nonevidence deficiency reasons, whereas the proportion rejected "in the interests of justice" remains relatively constant, this may be a signal that the quality of cases submitted by the police has improved. If the *overall* rejection rate rises too, it may signal an elevation in charging standards as well. But again, one needs to know how police arrest practices have changed in order to be able to make an assessment about how charging standards have changed.

There is a final performance "measure," although not a statistically derived measure, of charging threshold—the judgments made by case auditors regarding the charging standards actually used, given the observed screening decisions made on the basis of information available to the screener at the time of screening.

CASE SCREENING: THE ACCURACY OF CHARGING

In selecting performance measures for assessing change in charging accuracy, we have employed the more restrictive definition of accuracy: A prosecutor (or court)

[3] Setting aside the possibility that prosecution may be suspended or deferred in favor of pretrial diversion.

Table 1

PERFORMANCE MEASURES IN CASE SCREENING: CHARGING STANDARDS OR THRESHOLD

Performance Measures	Required Data Elements
By category of highest offense at police booking as a percentage of cases screened over a specific time period:	Number of cases screened
Filing rate on most serious booked-charge level	Number of cases filed at most serious booked charge
Filing rate on less than most serious booked-charge level[a]	Number of cases filed at less than most serious booked-charge level
Gross rejection rate	Number of cases rejected
Reasons for rejection (in percentage of cases rejected):[b]	Number of cases rejected
Evidence deficiency (illustrative subcategories: no corpus delicti, no connecting evidence, insufficient evidence, inadmissible evidence, returned to police for need of more investigation)	Number of cases rejected for evidence deficiency
Nonevidence deficiency: specific reasons (illustrative subcategories: victim requests no prosecution, need to grant immunity, suspect currently confined or pending conviction on other charges, contrary to legislative intent)	Number of cases rejected for specific nonevidence deficiency reasons
Nonevidence deficiency: general reasons (illustrative category: interests of justice)	Number of cases rejected for general or interests of justice reasons
Case audit results	Judgment about the charging standards used, given the observed screening decision

[a]May be broken down in greater detail to reflect the seriousness of the charge. For example, if there are three felony levels of seriousness (A, B, and C) and if the most serious charge level is, say, Felony A, this measure may be broken down into three categories: Felony B, Felony C, and misdemeanor.

[b]Depending on whether the screening agency (prosecution, lower court) has a pretrial diversion program, which agency makes the diversion decision, and at what point in the felony proceeding diversion may occur, an additional category of "rejection" reasons may be added.

has charged a defendant accurately if the evidence at the time of charging suffices to convict him of the most serious charges filed. Other observers have employed different definitions of charging accuracy. For example, Abrams[4] views accuracy as being the correctness with which the future "disposition" of the defendant is predicted. And some observers feel that, since the prosecutor meets his legal duty if he employs a probable cause standard, this is the appropriate basis for defining charging accuracy.

Inaccuracy in filing charges tends to affect stages of a criminal proceeding and performance measures of events *subsequent* to charging. However, measures of the outcomes of subsequent events are ambiguous, since they are governed by a variety of factors, only one of which is the accuracy of the charges as originally filed. For example, acquittals may also occur because of prosecutorial shortcomings in case preparation, arguing motions, or conducting the case at trial. Or, superior defense counsel performance may be the explanation. This ambiguity could be largely elimi-

[4] N. Abrams, "Prosecutorial Charge Decision Systems," *UCLA Law Review*, Vol. 23, No. 1, October 1975, pp. 49-55.

nated by accurately recording the true reason for acquittal. Similarly, the outcomes of other events (the preliminary hearing, pretrial hearings on motions of various types, plea negotiations, etc.) may also be affected by both charging accuracy and other factors. So we are led to the use of a *set* of performance measures, as listed in Table 2.

Table 2

PERFORMANCE MEASURES IN CASE SCREENING: CHARGING ACCURACY

Performance Measures	Required Data Elements
By category of highest offense at charging as a percentage of cases charged or disposed over a specified time period:	Number of cases charged or number of dispositions
Overall nonconviction rate	Number of nonconvictions
Pretrial dismissal rate	Number of pretrial dismissals
Nolle prosequi rate (if applicable)	Number of nolle prosequi
Pretrial diversion or intervention rate (if applicable)	Number of pretrial diversions
Trial acquittal, dismissal, and mistrial rate	Number of trial acquittals, dismissals, mistrials
Conviction rate on all charges as filed	Number of convictions on all charges as filed
By plea of guilty	As above, by plea
By trial	As above, by trial
Conviction rate on at least one of the most serious charges, but with charge or count reductions	Number of convictions on at least one of the most serious charges, but with charge or count reductions
By plea of guilty	As above, by plea
By trial	As above, by trial
Conviction rate on lesser charges than the most serious filed	Number of convictions on lesser charges
By plea of guilty	As above, by plea
By trial	As above, by trial
Case audit results	Strength and appropriateness of filed charges in average case

Basically, these measures categorize four types of dispositions subsequent to charging: (1) nonconviction rate (the sum of the pretrial dismissal rate; nolle prosequi rate, if applicable; pretrial diversion or intervention rate, if applicable; and trial acquittal, dismissal, and mistrial rate); (2) conviction rate on all charges as filed (by guilty plea and by trial); (3) conviction rate at the most serious charge level, but with some charge or count reductions (by guilty plea and at trial); and (4) conviction rate at lesser charge levels[5] (by guilty plea and by trial).

Given the ambiguity of these measures, one would have to observe large changes over time to conclude that charging accuracy has changed. Or if it is known from other sources that *other* factors have remained essentially constant (e.g., resources devoted to case preparation, competence of opposing counsel, and plea bargaining policy), smaller changes over time in these measures would suffice to conclude that

[5] These performance measures can be broken down in greater detail to reflect the seriousness of the convicted charge as illustrated in footnote a, Table 1.

charging accuracy has changed. However, if plea bargaining policy shifts dramatically—as it did in one of the jurisdictions in which we applied performance measures —changes in some of the performance measures shown in Table 2 may reflect mainly the consequences of the new plea bargaining policy rather than a change in charging accuracy. (See Chap. 3.)

There is a final "performance measure," although not a statistically derived measure, of charging accuracy: the judgments made by case auditors regarding the strength of the cases and the appropriateness of the charges filed for the samples of cases audited. The audit samples, of course, should be drawn from the same case populations and time periods as those for which statistical performance measures are estimated.

PERFORMANCE MEASURES IN PLEA BARGAINING

Plea bargaining (the terms *plea negotiation* or *case settlement* are used here synonymously) is currently an issue attracting much attention and public debate. Many jurisdictions have conducted or are conducting experiments to ascertain the effects of curtailing or eliminating some or all forms of plea bargaining.

We view performance measures of plea bargaining from two perspectives: those that indicate *what* happens over time and those that indicate how, on one hand, the systemwide *balance* of gains or other (possibly harmful) operational effects change over time, and, on the other hand, how concessions to defendants change over time.

Table 3 displays performance measures relating to *what* happens over time to the gross plea rate and to the frequency of different types of plea bargains: (1) the straight plea rate (to all charges and counts with no *other* bargain); (2) the straight plea rate with *other* bargains (such as a sentence agreement or an agreement by the prosecutor not to oppose a defense-recommended sentence, or an agreement to drop other pending cases); (3) the plea rate to at least one count of the most serious charge, with other charges and/or counts reduced (with and without *other* bargains); (4) the original charge plea rate (the sum of Items 1, 2, and 3); (5) the plea rate to lesser charges (with and without *other* bargains); and (6) the gross plea rate (the sum of Items 1 though 5).

Table 4 displays measures of the systemwide effects of plea bargaining. On one hand are measures of operational effects, some of which may be viewed as "gains" and some as "losses," depending on the direction of the changes in the performance measures and one's philosophical view. A change in plea bargaining policy (and the resulting change in plea bargaining rates) could affect pretrial dismissal rate (and diversion or nolle prosequi rate, where applicable), trial rate, trial conviction rate, overall conviction rate, measures of sentence severity imposed (e.g., percent of guilty pleaders incarcerated or the average sentence severity score imposed),[6] delay in proceedings (such measures as the period from arrest to final disposition or arraign-

[6] A score may be derived by weighting the different sentence elements in each sentence imposition and summing them into a total score. The relative weights are arbitrary. In our study we used four sets of weights to reflect different views of practitioners and citizens. For example, Sentence Severity Index A (an approximation to the weighting scheme derived by the California Bureau of Criminal Statistics) assumes that one point = 1 month in jail = 6 months of probation = $1000 fine, while the score for prison time is 18 plus one per year of prison sentence imposed. The other three weighting schemes we used successively weight probation and fines relatively *less* compared with jail or prison incarceration. (See Table 4.5 and the accompanying discussion in Sec. IV, R-1918-DOJ.)

Table 3

PERFORMANCE MEASURES IN PLEA BARGAINING: THE FREQUENCY OF PLEA BARGAINING

Performance Measures	Required Data Elements
By category of highest offense charged *as a percentage of dispositions over a specified time period:*	Number of dispositions
1. Straight plea rate (to *all* charges and counts) with no *other* bargain	Number of guilty plea dispositions in each designated category
2. Straight plea rate With sentence agreement With agreement to drop other cases With combination of the above	
3. Plea rate to at least one count of most serious charge with other charges and/or counts reduced With no *other* bargain With sentence agreement With agreement to drop other cases With combined sentence/drop other cases agreements	
4. Original charge plea rate (Sum of Items 1, 2, 3)	
5. Plea rate to lesser charges (charge bargaining rate) With no *other* bargains With sentence agreement With agreement to drop other cases With combined sentence/drop other cases agreements	
6. Gross plea rate (Sum of Items 4 and 5)	Total number of guilty pleas

ment to final disposition),[7] and the use of witnesses' and jurors' time (such measures as the number of victim- and witness-hours consumed per disposition and the fraction of time jurors spend in jury selection and trial). Most observers would agree that if delay were reduced by a change in plea bargaining policy this would constitute a gain to the system, as long as the quality of justice was not adversely affected. However, if a change in plea bargaining policy led to a decrease in average sentence severity imposed, there is considerable controversy as to whether this constitutes a system gain or loss.

In contrast is the case of *concessions* granted to the defendant by the system in return for a plea of guilty. We argue in the companion volume to this study that the preferred way of measuring the amount of concession to the defendant is to use the resulting *change* in imposed punishment (i.e., change in sentence severity score) between that which was *threatened* by the prosecutor or judge in a plea bargain and that which was imposed. But the collection of data on "threatened" punishment presents formidable difficulties, because, to our knowledge, it is never recorded. Thus, we are led to two alternative reference levels of punishment: the sentence severity that would have been imposed in the absence of a plea agreement (1) had

[7] Final disposition includes pretrial or trial dismissal, acquittal, and sentencing after conviction by guilty plea or trial.

Table 4

PERFORMANCE MEASURES IN PLEA BARGAINING: BALANCE OF SYSTEMWIDE EFFECTS AND SYSTEMWIDE CONCESSIONS TO DEFENDANTS

Performance Measures	Required Data Elements
Systemwide Operational Effects	
By category of highest offense charged over specified period of time:	
Pretrial dismissal, diversion, nolle prosequi rates	Number of cases in each category/number of dispositions
Trial rate (total, bench, jury)	Number of trials (total, bench, jury)/number of dispositions in each category
Trial conviction rate (total, bench, jury)	Number of trial convictions (total, bench, jury)/number of trials in each category
Overall conviction rate	Number of convictions/number of dispositions
Sentence severity imposed:	
Percent of guilty pleaders incarcerated	Number of guilty pleaders incarcerated/number of guilty pleaders
Average sentence severity score imposed	Sentence elements by type and amount imposed on guilty pleaders
Delay or elapsed time measures:	
Median arrest to final disposition period Median arraignment to final disposition period	Median elapsed time (days) between these events
Number of victim-hours per disposition	(Number of victim appearances/disposition) X (time per appearance)
Number of other witness-hours per disposition	(Number of witness appearances/disposition) X (time per appearance)
Percentage of juror time in jury selection and trial	Juror time in jury selection and trial ÷ total time (including idleness)
Systemwide Concessions to Defendants	
Sentence concession per guilty pleader	Difference between average sentence severity score of all straight pleaders and of all guilty pleaders
Sentence concession:	Difference between average sentence severity score of all straight pleaders and
Per charge bargainer	Average severity score of all charge bargainers
Per count bargainer	Average severity score of all count bargainers
Per sentence bargainer	Average severity score of all sentence bargainers
Per combination of above	Average severity score of all combinations

the defendant gone to trial, *or* (2) had he made a straight plea to the original charges and received no *other* plea bargain—all other things being equal, such as defendant and case characteristics. However, because it is often impossible to determine certain defendant characteristics from recorded data, such as his propensity for risk-taking in a decision between going to trial or pleading guilty, and because of the statistically small *number* of trials, we have used the straight plea sentence as our reference level of punishment in our application of these performance measures. In addition, we believe it is important for a jurisdiction to determine whether and to what extent imposed sentences differ for like cases that go to trial or involve straight pleas in measuring plea bargaining effects, since the conventional wisdom is that the system exacts a penalty in punishment imposed on those who exercise their rights to a trial.

We suggest that the performance measures of system concessions to defendants include the average concession per *convicted* defendant (by all types of guilty plea, including straight plea) and the average concession granted per *type* of plea bargain (charge bargain, count bargain, sentence bargain), as indicated in Table 4.

MEASURING AND EXPLAINING SENTENCE VARIATION

There are two aspects to sentence variation: identifying criteria that measure *how much* sentence variation occurs and *explaining* how much of the total observed variation is accounted for by various factors that should (i.e., "legitimate") or should not (i.e., "illegitimate") cause the variation. A standard way of measuring overall variation in conventional sentence categories is the *percentage of convictees receiving each sentence element by amount of that element* (e.g., 5 to 10 years of probation, 2 to 4 years of prison). A standard statistical measure of variation using sentence severity scores is the *coefficient of variation*—that is, the ratio of the standard deviation of the sentence severity score to the average sentence severity score. These measures should be estimated for each major category of *convicted* offense, if one believes that the convicted charge most accurately reflects the criminal conduct in these cases. Or it may be estimated for each category of highest *charged* offense, if one believes that the charged offense most accurately reflects the criminal conduct in these cases.

Sentences vary for many reasons. Acceptable (or "legitimate") reasons include the defendant's age, prior record, an index of community ties (including employment, education, family status, transiency), the nature and number of charges against the defendant, and judicial sentencing philosophy. Some reasons are not so acceptable (or are "illegitimate"); that is, most observers would agree that these factors should *not* affect sentencing. They may include the defendant's ethnicity, pretrial custody status, type of defense counsel, method of conviction (guilty plea or trial), and changes over time in the crowding of correctional facilities. By applying standard multivariate statistical techniques, it is possible to estimate how much of the explained variation is accounted for by each of these factors (with the exception of judicial sentencing philosophy). (Data on all but the latter factor are normally recorded or could be easily recorded in agency files. Judicial sentencing philosophy would be very difficult to quantify.) The remaining unexplained variation can then be attributed to factors that could not be identified or measured or to random chance. The performance measures here are the *addition to the explained* variation accounted for by each of these legitimate and illegitimate factors.[8]

PERFORMANCE MEASURES OF EVENHANDEDNESS

Evenhandedness is a quality of court system operations that is characterized by the following question: Do defendants in *similar* circumstances fare comparably in criminal proceedings? It is related to the previous discussion of sentence variation

[8] In the interests of conserving space in this report, the reader is referred to an illustration of the items and format of sentence variation performance measures applied in one jurisdiction as given in Tables 6.15, 6.16, and 6.17 of Sec. VI, R-1918-DOJ.

analysis but it includes the effects of illegitimate factors on *dispositions* as well as on sentences and it includes the *direction* of effect on dispositional and sentence outcomes. The indicators of absence of evenhandedness are the direction and magnitude of effects of each of the illegitimate variables on the set of performance measures of dispositional and sentence outcomes. The effects of the illegitimate factors may be estimated by cross-tabulations or by multivariate statistical techniques. The former can indicate only the presence of gross effects. For example, if the straight plea rate of black burglary defendants is 70 percent and that of white burglary defendants is 50 percent, one cannot be sure that all of the difference is due to ethnicity, because some of the difference may be due to *other* factors. Continuing with this example, if a much larger proportion of the black burglary defendants remains in jail and has public defender representation, the effects of these factors may confound the ethnicity effect. Multivariate statistical techniques, however, permit the identification of the unique or independent effect of each of these factors.

For dispositional evenhandedness, we include three illegitimate factors: ethnicity, pretrial custody status, and type of defense counsel.

For sentencing evenhandedness, we add type of disposition (straight plea of guilty versus trial) to this list of three factors if jurisdictions wish to test the proposition that defendants who exercise their right to a trial receive a penalty in punishment, and crowding in correctional facilities, to estimate to what extent crowding leads to less severe sentences.

Most observers would agree that ethnicity alone should not affect dispositional or sentence outcomes under our system of justice. They would also concur that pretrial custody status per se should not affect whether an arrestee is charged or rejected; dismissed, convicted, or acquitted; or, if convicted, given a more or less severe sentence. In practice, the defendant's pretrial custody status may *in fact* be related to his disposition or sentence. One argument holds that compared with defendants in custody, those who are released on bail or own recognizance (O.R.) are better able to strengthen their defense by locating witnesses. Another contends that defendants held in jail (particularly for offenses in which the probability of receiving a nonincarceration sentence is high) have more incentive to plead guilty since this cuts short the total (pretrial and postconviction) incarceration period. In addition, the failure to obtain pretrial release may be a proxy for the viciousness of the offense or the bad character of the defendant (in ways not fully reflected by the actual original charges, the convicted charges, and the defendant's prior record).

Most observers would also agree that the fact that a defendant has the services of a public defender, a privately retained counsel, or a court-appointed attorney should not influence his disposition or sentence. But in reality, the type of defense attorney may, in fact, be related to disposition or sentence. One argument is that a wealthier defendant who retains private counsel can provide more resources in building a defense (e.g., a private counsel may have more investigative resources than a public defender). Another view is that public defenders (compared with the usually less-specialized court-appointed or retained counsel) may achieve better results for defendants because they know the court system and its practitioners better. For example, the public defender may be more skilled in judge-shopping to avoid the more severe sentencers.

Most observers would also agree that, all other things being equal, the type of conviction, whether by trial or straight plea of guilty, should not affect sentence severity. The conventional wisdom, however, is that conviction by trial leads to more

severe sentences because the system penalizes defendants who exercise their right to a trial. (However, in our application of measures of evenhandedness in the two demonstration jurisdictions, the analysis of a small sample of trial convictions revealed *no* penalty in sentence severity compared with other similar defendants who pled guilty to the original charges.)

Finally, the degree of crowding in correctional facilities should not affect sentences, but in reality, it undoubtedly does. In some jurisdictions where correctional facilities are extremely overcrowded, judges may feel compelled to impose probation or other nonincarceration sentences on the least dangerous defendants who would otherwise have been incarcerated.

Table 5 gives the performance measures together with the comparisons that help assess the effects of the illegitimate variables. (Required data elements are not repeated here since they were indicated in previous tables.)

Table 5

MEASURING EVENHANDEDNESS IN DISPOSITIONAL AND SENTENCE OUTCOMES

Performance Measures	Magnitude and Direction of the Illegitimate Factor Effects
By category of highest offense as booked by police and charged by prosecutor over a specified time period:	*(Each of the illegitimate factors must be identified for each case in the sample)*
Dispositional Outcomes	
Rejection rate (based on number screened)	Black, Spanish, other, minority vs. majority Held in jail vs. released on bail vs. released on O.R. Public defender vs. retained counsel vs. court-appointed counsel
Nonconviction rate (based on number of dispositions)	
Pretrial dismissal rate	
Pretrial diversion rate	
Nolle prosequi rate	
Conviction rate (based on number of dispositions)	
By straight plea	
By any plea bargain (gross plea rate)	
By trial (and trial conviction rate)	
Overall	
Sentence Outcomes	
Sentence severity score	*As above but add:* Trial conviction (all, bench, jury) vs. straight pleas Crowding in correctional facilities

PERFORMANCE MEASURES OF DELAY AND USE OF LAY PARTICIPANT TIME

The speediness of felony proceedings is a highly visible attribute by which criminal justice is necessarily judged, and it is a matter of urgent concern to officials and the public. In many jurisdictions, delay has increased because resources allocated to court system agencies have not kept pace with rising caseloads or because practitioner time is not used efficiently. We adopted three basic gross measures of

delay or elapsed time between specified events: median[9] number of days, minimum time for the lengthiest 10 percent of cases (i.e., the shortest of the longest 10 percent of cases), and the percentage of cases exceeding some standard (set by court rule or statute). The event intervals we selected were: from arrest and from arraignment to dismissal; to plea of guilty; to trial; to sentencing; and to final disposition (the sum of cases dismissed, acquitted, and sentenced). In addition, the time between conviction (by any means) and sentencing is of salient interest. Speedy trial standards usually refer to the arrest-to-trial period, but it is revealing also to estimate the percentage of cases exceeding the standard that are disposed of otherwise. These performance measures should be calculated for *all* felonies and for *each* major offense category, as defined at the point of charging, to determine which offense categories account for more or less delay.

In addition, it is useful to apply multivariate statistical techniques to determine the independent effects of certain factors on delay. We have hypothesized that four factors may influence delay: (1) pretrial custody status (because defendants out on bail or O.R. may have more incentives to delay proceedings); (2) type of defense attorney (because court calendar conflicts or incentives stemming from attorney compensation arrangements may lead to more delay by private attorneys, or because one category of attorney might know the system better than another); (3) type of disposition (cases tried may take longer than guilty plea cases); and (4) heavier or lighter caseloads or more or less court backlog as measured directly or as reflected in a time trend (proxy) variable.

Continuance measures, which are indirect measures (and causes) of delay, provide additional insights into performance. Continuance statistics should be estimated separately for contested and uncontested cases (because of large differences between the two types of cases), as well as for the sum of contested and uncontested cases. Where possible, the "movant" or "requestor" should be identified (e.g., prosecution, defense, court, or joint) to reveal which agencies are most responsible for continuance-induced delays. The continuance measures we adopted were the percentage of cases continued, the number of continuances per case, and the average number of days continued per case.

Measures of the use of victims' and other witnesses' time that we adopted were the number of appearances per case disposition, the time consumed per appearance, and the number of total appearance-hours consumed per disposition (the product of the previous two measures). Each measure was calculated separately for victims and other witnesses. At best, court records may contain information on number of victim and witness appearances per disposition. Information on time per appearance is generally not recorded, at least not in the two jurisdictions we examined. It can be measured directly (the preferred approach) by filling out time sheets for each appearance, or it can be estimated from the response of victims and witnesses to mail surveys (as we did). However, the latter approach is flawed because it relies on the memories of respondents. These performance measures should be calculated for the entire (or a sample thereof) felony caseload to obtain a general picture, but they may also be estimated for particular offense types that are thought to consume much victim or witness time.

Conventional measures of juror usage essentially measure the oversupply of

[9] Median instead of mean or average number, because the mean is too sensitive to the presence of a few cases of *very* long duration.

jurors in an indirect fashion. For example, the Juror Usage Index (JUI) used in federal courts is defined as the available number of jurors per day (summed over a month) divided by the number of juries in trial per day (summed over a month). We adopted a more informative and direct set of measures by simply calculating the average fraction of time a juror spends in idleness (although in the courthouse), in jury selection, and in trial. The latter two measures should be calculated separately for civil and criminal matters in jurisdictions that use the same juries for both types of cases. Information on the use of juror time is normally not recorded, at least in the two jurisdictions we examined. Again, it can be measured directly (the preferred approach), or it can be estimated from juror responses to mail questionnaires (as we did).

Table 6 displays the performance measures adopted for delay and use of lay participant time together with the required data elements.

PERFORMANCE MEASURES OF THE USE OF PRACTITIONER TIME

The "weighted caseload" approach has developed over the past several years to be the most widely used way to assess case processing efficiency (or, alternatively, the efficiency of the use of practitioner time) of practitioners in prosecution, defense, and court agencies. This approach is also used (e.g., in California) to ascertain staffing requirements in these agencies. For illustrative purposes, we discuss here only the use of *judicial* time, but similar approaches are available for measuring prosecutorial and public defender case processing efficiency.[10]

The performance measure of judicial criminal case processing efficiency is *the number of weighted cases processed (counted either by filings or dispositions) per available judge-year.* The ingredients of this measure are:

$$\frac{\text{Available judge-year time}}{\text{Weighted caseload processing time per disposition (or filing)}}.$$

The numerator is the total time per year per judge less the time consumed by civil matters, vacations, sick leave, official traveling, professional meetings, etc. The denominator is the product of the average duration of time per judicial activity and its relative frequency of occurrence, summed over all judicial activities. For example, taking a guilty plea might consume 20 minutes of the court's time on the average and occur six times per ten dispositions. If so, the weighted processing time per disposition would be 12 minutes per guilty plea. Summing over the various activity types (see Table 7 for illustrative classifications of judicial activities used in California, Florida, and Multnomah County, Oregon), we obtain the weighted caseload processing time per criminal disposition for all types of dispositions. To estimate this processing time *per criminal filing,* we multiply the preceding result by the ratio of dispositions to filings.[11]

[10] See, for example, *Staffing Requirements Projection Approach for Professional Prosecution and Defense Services,* Peat, Marwick, Mitchell & Co., Los Angeles, September 1974, a study done for the County of Santa Clara, California, under LEAA (Law Enforcement Assistance Administration) funding.

[11] This simple procedure rests on the assumption that nondisposed cases (i.e., backlog) resemble disposed cases in the consumption of judicial time.

Table 6

PEFORMANCE MEASURES OF DELAY AND THE USE OF LAY PARTICIPANT TIME

Performance Measures	Required Data Elements
Elapsed Time between Events	
For all felonies and specific offense categories over a specified time period:	
1. Median number of days	
2. Minimum number of days for longest 10 percent of cases	
3. Percent of cases exceeding standard[a]	
These measures to be calculated between the following events:	*For each case in each sample, applicable dates of arrest, arraignment, dismissal, guilty plea, trial (commencement and end dates), sentencing*
Arrest and arraignment } and { Dismissal, Guilty plea, Trial, Final disposition	
Conviction and sentencing (Meas. 1 only)	
Continuances	
Separately for all, contested and uncontested, cases in a random sample of all felonies over a specified time period:	
4. Percent of cases continued	Number of cases continued ÷ number of cases
5. Continuances per case	Number of continuances ÷ number of cases
6. Number of days continued per case	Number of continued days (= number of continuances X number of days per continuance) ÷ number of cases
7. Percent of continuances attributed to defense, prosecution, court, joint	Identity of "movant" or "requestor"
Use of Lay Participant Time	
Victims and Other Witnesses	
For all felonies over a specified time period:	
8. Number of appearances per disposition	Separately for victims and other witnesses
9. Time consumed per appearance	Separately for victims and other witnesses
10. Total number of appearance-hours per disposition	(Measure 8) X (Measure 9)
Jurors	
For all trials over a specified time:	
11. Percent of time idle	
12. Percent of time in jury selection	Separately for civil and criminal trials
13. Percent of time in trial	Separately for civil and criminal trials

[a]The speedy trial standard (defined by court rule or statute) usually applies only to the arrest-to-trial period, but it is of interest to compute this measure as if the same standard applied as well to the period between arrest and dismissal, arrest and guilty plea, and arrest and final disposition.

Table 7

CLASSIFICATIONS OF JUDICIAL ACTIVITIES IN CRIMINAL PROCEEDINGS

California Judicial Council Study[a]	Florida Weighted Caseload Statewide Study	Multnomah County (Oregon) Circuit Court[b]
Short matters (plead not guilty, continuance, calendar call, sentencing and probation hearing, diversion hearing, other pretrial motions, trial confirmation conference)	Case-related, with party and counsel present:	Arraignment
Plead guilty	First appearance hearing	Motion hearing
Dismissal transfer	Preliminary hearing	Plea hearing
§995 PC (penal code) motion	Arraignment	Other hearings
§1538.5 PC motion	Motion hearings	Court trial
Court trial (regular, transcript, transcript and testimony)	Plea hearing	Jury trial
Select jury and jury trial	All other hearings	Sentencing hearing
Habeas corpus hearing	Pretrial conference	
	Other conferences	
	Detention hearing	
	Adjudicatory hearing	
	Disposition hearing	
	Nonjury trial	
	Jury selection	
	Jury trial	
	Sentencing/presentence investigation	
	Postdisposition/trial hearing, motions	
	Other case-related	
	Case-related, office work:	
	Predisposition (legal research, drafting, etc.)	
	Postdisposition (legal research, drafting, etc.)	
	Conferences	
	Jury-related:	
	Grand jury	
	Mass jury selection	
	Statewide grand jury	
	Coroner's jury	
	Non-case-related:	
	Correspondence	
	Travel	
	General research and study	
	Conferences	
	Court administration	
	Ex officio	

[a]*Judicial Weighted Caseload System Project*, Final Report Prepared for the Judicial Council of California, Arthur Young & Company, Sacramento, May 1974. This list of activities pertains to criminal (felony) proceedings in superior court. The cited report gives similar lists for other types of proceedings in superior court as well as the various proceedings in municipal courts.

[b]See Sec. VI, R-1918-DOJ.

Currently, jurisdictions that collect the necessary data and calculate weighted criminal caseload for judges do so only for criminal cases as a whole and not by offense class.[12] In jurisdictions where the case mix by offense tends to vary significantly over time or locality, it would be desirable to calculate weighted caseload performance by offense class (as well as for all felonies) and possibly for different geographic divisions of the court as well. Moreover, maintaining such measures by offense class makes it more feasible to compare jurisdictions by adjusting for the effects of different case mixes on the performance measures. (In comparing jurisdictions, it is also necessary to adjust for differences in the relative frequency of activities.)

MEASURING LAY PARTICIPANT ATTITUDES AND THEIR DETERMINANTS

The attitudes of jurors, victims, other witnesses, and defendants toward the court system and its practitioners constitute the last broad issue area we considered. These attitudes are important because they affect the atmosphere in which practitioners function and may even affect their decisions (e.g., a judge's decisions on sentencing can mirror the community's views, and more particularly, the views of lay participants). Then, too, the attitudes of jurors and witnesses can have direct operational consequences for court systems if they affect the willingness to cooperate in felony proceedings. Attitudinal data of this sort are not normally collected, even occasionally, in local jurisdictions. Thus, we were led to the use of mail surveys of victims, witnesses, and jurors and personal interviews, using a structured questionnaire, with defendants. The mail and personal interview questionnaires we developed seem to have worked fairly well and have been published for jurisdictions that wish to consider their use.[13] (As an alternative to mail questionnaires, jurisdictions may use an exit questionnaire for jurors, victims, and other witnesses.)

For the first three classes of lay participants—jurors, victims, and other witnesses—we found that questions dealing with the following topics adequately captured the attitudes of interest:

- Overall opinion of the court system (both before and after lay participants' experiences).
- Whether justice was done in the case(s) in which they were involved.
- Their degree of satisfaction with their experience.
- Their attitudes toward the practitioners they dealt with.
- Their willingness to cooperate again in the future.
- Whether their understanding of the system increased as a result of their experience.

The hypothesized determinants of these attitudes include the lay participants' background characteristics (age, sex, ethnicity, income, employment, etc.), the char-

[12] However, different types of *proceedings* are already being handled separately, as in the superior courts of California (e.g., criminal, juvenile delinquency, probate, personal injury-property damage, eminent domain), and in the municipal courts of California (e.g., felony preliminary-felony reduction, traffic, intoxication, other misdemeanors, civil, small claims).

[13] See App. G, R-1918-DOJ, which reproduces the mail survey questionnaires administered to jurors, victims, and other witnesses, and App. F, which evaluates them.

acteristics and outcome of the cases they were involved in, the problems they encountered in serving, and the treatment they received by the court system. Standard statistical techniques (such as correlational analysis and multivariate regression analysis) should be applied to determine the relationships that do or do not exist between these hypothesized determinants and their professed attitudes. In particular, jurisdictions should focus on analyzing *policy* factors that could be manipulated to improve attitudes. These include reducing the number of problems encountered (such as parking, transportation, idleness, lack of facilities, and concern for their safety in the court building); improving the system of notifying witnesses when to appear, informing them of reschedulings, and transporting them (if applicable); informing them of case outcomes; treating them with courtesy and respect; and instilling a feeling that their participation helped bring about justice.

For defendants, we found that questions dealing with the following topics adequately captured the attitudes of interest:[14]

- Perceived fairness of procedures and outcome in their case.
- Their attitudes toward prosecutors and judges.
- Their attitudes toward and perceptions of defense counsel.
- Their attitudes toward plea bargaining.

Possible determinants of these attitudes include defendant background characteristics (including prior record), type of offense in the instant case, type of disposition and nature of sentence (if any) imposed, and type of defense counsel.

GAUGING OVERALL PERFORMANCE

Although we did not undertake to gauge "overall" performance of a jurisdiction in this study, we feel it is necessary to touch on the question of objective approaches to the *aggregate interpretation* of an entire battery of performance measure sets of the sort we discuss in this study. One approach might be the construction of an *aggregate* performance measure made up of separate indices of performance in each selected issue area. Of course, this approach immediately presents us with a multitude of "apples and oranges" issues, whose resolution tends to be strongly subjective. What activities within the process should be selected for inclusion in the aggregate measure? How may they be made commensurable? And most difficult of all, how much importance should be attributed to one activity relative to another? For example, what weight should case processing efficiency be given relative to charging accuracy in the construction of an aggregate performance measure? Such dilemmas might be resolved by statements of position by responsible policymakers. Or they might be circumvented by the use of a small set of alternatives, spanning reasonable differences of opinion. But, in practice, a single composite measure of performance for a complex process serving a multitude of (sometimes conflicting) objectives would not gain acceptance. The criminal justice community shows a clear preference for retaining the separate identities of different functions and agencies, and for inte-

[14] See App. I, R-1918-DOJ, which reproduces the personal interview questionnaire administered to defendants.

grating their performance (if any integration is attempted at all) subjectively according to the evaluator's view of their relative contributions.

In the rare case where *clear* changes are observed over time in performance measures (for, say, the better) in *all* or *almost all* of the issue areas of interest, it is then possible to draw a *qualitative* inference that "overall" performance changed for the better. Our application of performance measures in one of the two jurisdictions uncovered such a case.

Chapter 3

APPLICATION OF PERFORMANCE MEASURES IN ONE JURISDICTION

To illustrate one of the roles or uses of performance measures, in this chapter we summarize selected findings from the application of these measures *within* one jurisdiction—Multnomah County, Oregon. Multnomah County is a medium-size jurisdiction with no severe court calendar crowding or crowding problems in its correctional facilities. The circuit court has 11 trial courts handling both civil and criminal cases and one chief criminal court that conducts recognizance and bail hearings, hears motions to dismiss, conducts trials on stipulated facts, receives all pleas except those taken in the course of a trial, and generally sentences on guilty pleas taken. The circuit court handles about 3000 felony case filings annually, including about 300 felony trials.

The district attorney's office assigns 27 prosecutors to felony cases (or an annual caseload of about 100 cases per prosecutor), and most of the staff is specialized. There are separate units for intake (screening), pretrial matters (grand jury, preliminary hearings, pretrial motions, extradition, etc.), and for trials (specialized by broad offense category). In late 1973 and early 1974, a special Impact Unit was formed (including six additional prosecutors and five support persons) to handle Robbery I, dwelling Burglary I, and "fencing" cases from screening to final disposition. This unit was to implement a No Plea Negotiation Experiment (part of Portland's High Impact Anti-Crime Program funded by LEAA). The experiment's goals were to improve the quality of cases by providing ongoing assistance to the police, to provide swift prosecution of these "Impact" crimes, and to virtually eliminate negotiated pleas (defined to be only *charge* reductions). *Other* forms of plea bargaining, such as count reduction, sentence agreements, and agreements to drop other pending cases, were presumably acceptable.

The Metropolitan Public Defender's Office handles roughly 45 percent of the felony caseload of Multnomah County; eight attorneys handle this caseload (about 175 cases per attorney per year). The felony attorney complement is organized into two-man teams that defend a case from appointment to disposition. A "blind" case assignment system is used.[1]

Our application of performance measures in Multnomah County was largely keyed to a preliminary evaluation of the systemwide effects of the No Plea Negotiation Experiment. Below we summarize some of the more interesting findings (by issue area) through a comparison of changes in performance before (1973) and during (1974) the experiment.

[1] For a statistical overview of the characteristics of the felony caseloads and felony defendants, see Sec. VI, R-1918-DOJ.

DID CASE QUALITY AND CHARGING STANDARDS CHANGE FOR IMPACT OFFENSES?

In applying the performance measures discussed in Chap. 2, we were led to conclude that between 1973 and 1974 *both* case quality and charging standards or threshold were raised for Impact offenses, but the data were insufficient to distinguish how much of the improvement was accounted for by better police investigations and how much by the prosecutor's evaluation of the screening threshold. For example, for Robbery I bookings, no marked year-to-year changes were observed in overall rejection rate (42 to 51 percent) or in the filing rate on the most serious charge (34 to 41 percent), but the filing rate on *lesser* charges notably declined (24 to 8 percent). These measures indicate that there was no apparent attempt by the prosecutor to reduce the booking-charge level of a potential Impact case, which would have made it a non-Impact case on which plea bargaining was not constrained by the experiment. (However, data were not available to gauge whether police arrest and booking practices changed in the direction of eliminating candidate Robbery I Impact cases by booking them as lesser non-Impact robberies, say, Robbery II.)

Two indicators of improved case quality were that the percentage of rejections for evidence deficiencies declined markedly (84 to 55 percent), and within this category the percentage rejected and returned to the police because they needed more investigation was cut in half (31 to 15 percent). However, the percentage of rejections for the broad undifferentiated reason of "interests of justice" rose somewhat (8 to 26 percent), but much less relative to the decline in evidence-deficiency reasons. Also, nonconviction rates (dismissals, acquittals, mistrials) declined for Impact offenses, but not for a comparable non-Impact offense. Given these changes in the measures, we concluded that *both* case quality and charging threshold were raised; had the prosecutor *not* tightened his charging standards, all other things being equal, he would have rejected a *smaller* proportion of Impact cases overall; in fact, no significant change in rejection rate was observed.

DID CHARGING ACCURACY INCREASE FOR IMPACT OFFENSES?

The experiment's ground rules were responsible for changes in certain outcome measures that are normally relevant (albeit ambiguous) to gauging charging accuracy changes (such as a decline in charge bargaining rate and a rise in straight plea rates that were altered deliberately by the experiment). However, an indicator of possible improvement in charging accuracy or in police investigation was the nonconviction rate, which fell strikingly for Impact crimes (53 to 24 percent for Robbery I and 17 to 5 percent for dwelling Burglary I) but not for a comparable non-Impact offense (42 to 44 percent for nondwelling Burglary I). But this indicator alone cannot disclose whether one or both changed. Thus, in this application we can conclude only that charging accuracy *did not lessen.* This conclusion is also supported by the results of the case audit analysis of Impact and non-Impact burglary plea-of-guilty cases: Cases in both years were judged to be fairly strong, with no dramatic year-to-year changes discernible.

PLEA BARGAINING CHANGES: WERE THE EXPERIMENT'S OBJECTIVES ACHIEVED AND DID THE BALANCE SHIFT?

In terms of the changes in the frequency and types of plea bargaining, we can conclude that the experiment's objectives were largely achieved. For example, *of all Robbery I guilty pleas,* charge bargaining decreased dramatically (from 59 to 6 percent); count bargaining fell from 18 to 6 percent), while straight plea rates rose dramatically (from 23 to 88 percent). "Original charge"[2] conviction rate for Robbery I by trial or guilty plea rose from 23 to 71 percent, as a proportion of all dispositions of cases originally charged with at least one count of Robbery I. However, the incidence of *other* types of plea bargains (sentence agreements or agreements to drop other pending cases) in combination with straight pleas or count bargaining *increased* from 31 to 49 percent of all Robbery I guilty pleas. Thus, to some extent at least, *other* types of plea bargaining supplanted charge bargaining during the experiment. There also appeared to be some spillover effects to comparable non-Impact offenses, such as nondwelling burglaries; similar directional changes were observed in these performance measures, but they were of lesser magnitude.

There was also a dramatic shift in the plea bargaining balance. In terms of systemwide gains or other operational effects, Impact offense pretrial dismissal rate fell (44 to 12 percent), and there was a rise in gross plea rates (41 to 61 percent) and overall conviction rates (47 to 77 percent).[3] Trial rate increased for Robbery I (15 to 27 percent) but not for dwelling Burglary I (17 to 13 percent). There was no change in these performance measures for the comparable non-Impact offense of nondwelling Burglary I. A higher proportion of Impact offenders was incarcerated (67 to 87 percent) and their average sentence severity score increased (16.7 to 26.5 for Index A);[4] there were similar spillover effects in one non-Impact offense. The increase may well be associated with the experiment, even though sentence policy was presumably not part of the experiment's ground rules. After adjusting for random sampling errors in case and defendant characteristics between 1973 and 1974 and for correctional facilities crowding in the two years, we observed an escalation in average sentence severity in the latter year. However, the customary rotation of sentencing judges *could* have also contributed to the escalation in sentences. Finally, Impact offenses were moved more expeditiously, although all felony cases were not; for example, median number of days between arraignment and final disposition declined somewhat for Robbery I cases (from 71 to 64), but increased substantially for all felony cases (34 to 63).

Because more straight pleas were taken during the experiment, systemwide sentence severity concessions per *convicted* defendant fell (e.g., from 9.1 to 1.3 in Robbery I cases, as measured by Index A). But concessions per defendant with a *charge bargain* (12.4 to 10.7) and with a *count bargain* (16.8 to 22.2) in Impact cases showed little year-to-year change.

If one believes that stiffer sentences are desirable, it is apparent that the dramatic shift in the plea bargaining balance was a shift for the better, since all of the other performance measures indicate better performance. If one believes that stiffer sentences are undesirable, one would have a mixed view of the new overall balance.

[2] That is, convicted on at least one count of the most serious charge.

[3] Illustrative figures in parentheses are for Robbery I cases in 1973 and 1974 except where noted.

[4] See footnote 6 in Chap. 2 for a description of Sentence Severity Index A.

HOW DID SENTENCING VARIATION CHANGE?

As we indicated above, between 1973 and 1974 average sentence severity imposed (as measured by Index A score) rose substantially for convictees who were charged with Robbery I or Burglary I. But our measure of sentence variation did not show much year-to-year change; it fell from 53 to 41 percent of the average score in Robbery I cases and was constant at 46 percent for Burglary I cases.

Legitimate factors accounted for almost all of the variation explained by the identified factors in Burglary I cases, for example, 28 out of the total 38 percent explained in 1974. In Robbery I cases, legitimate factors accounted for 48 percent out of a total of 61 percent explained by the identified characteristics. In both offense classes, the nature and number of the original and convicted charges accounted for most of the variation explained by legitimate factors (15 to 30 percent) with prior record next in importance (4 to 20 percent).

In Burglary I cases, illegitimate factors accounted for very little of the variation explained in both years (3 to 7 percent), whereas they played a more important role in Robbery I cases (29 to 11 percent). In the latter cases, pretrial custody status had a relatively large though mixed effect (3 to 9 percent) in the two years, and the choice of a trial (compared with a straight plea) accounted for up to 11 percent. The latter statistical result must be qualified because very few trials were included. It must be reiterated that unidentified or unmeasured factors accounted for about 69 percent and 39 percent of the variation in Burglary I and Robbery I sentences, respectively.

WERE DISPOSITIONAL AND SENTENCE OUTCOMES MORE EVENHANDED?

In general, on applying the performance measures, we concluded that dispositional and sentence outcomes were rather evenhanded in 1973 and even more so in 1974. The effects of pretrial custody status on dispositional outcomes were mixed in 1973 (and therefore inconclusive), but being held in jail resulted in more severe sentences (e.g., over 100 percent more in Robbery I cases). However, these effects tended to disappear in 1974.

Although there were scattered indications of some effects, our general conclusion is that, taken as a group, public defenders, retained counsel, and court-appointed counsel are roughly equal in effectiveness. There were weak indications, however, that the public defender seemed to do slightly better in Robbery I dispositions in both years (somewhat higher pretrial dismissal rate and lower conviction rates), and in 1973 Robbery I cases sentence severity imposed was higher for retained (by 37 percent) and court-appointed (by 20 percent) counsel; however, no effects of type of counsel were observed in 1974 Robbery I or Burglary I (1973 or 1974) cases.

Although there were some indications of unevenhandedness in 1973 toward blacks in dispositional outcomes and mixed (i.e., inconclusive) effects of minority status on sentence severity, most of these effects lessened or disappeared in 1974. In 1973, black burglary and robbery defendants were more likely to be dismissed before trial and less likely to plead guilty or be convicted. This suggests either overarrests, overprosecution, or the application of a double standard, but the data could not discern which hypothesis best explained the observed differences. These

differences disappeared in 1974 burglary cases. In 1973, black burglary convictees received sentences that were 24 percent more severe than whites, but for black robbery convictees, the effect was reversed. In 1974 no effects of minority status on sentence severity appeared.

Defendants who went to trial suffered little or no penalty in sentence severity compared with those who were straight pleaders. For example, in Burglary I cases the penalty was 18 percent in 1973 and 8 percent in 1974; in Robbery I cases there was no penalty in 1973, and in 1974 those convicted at trial actually received sentences that were 9 percent *less* severe, on the average.

These results, however, must be qualified: Dispositional differences are based on cross-tabulations, so the effect of each illegitimate factor cannot be isolated; sample sizes were so small, particularly for the trial effect, that inferences cannot be drawn confidently.

HOW WERE "HABITUAL OFFENDERS" TREATED?

In general, offenders with more serious prior records fared no worse than those with no prior record in the adjudication phase, but once convicted, they were sentenced more severely. For example, convicted robbery defendants with prior prison records received 69 and 45 percent more severe sentences in 1973 and 1974, respectively. However, more serious prior records tended to be associated with higher sentence severity scores in *both* years, suggesting that no special effect was associated with the experiment. (This result is not unexpected, since the experiment did not focus special attention on the "habitual" or "career" offender, as programs in other jurisdictions are now doing.)

WAS JUSTICE SWIFTER?

In applying the performance measures, two developments were apparent: For the entire felony caseload, delay increased between 1973 and 1974, whereas for the Impact offenses most performance measures indicated less, or at least no worse, delay in 1974. This indicates that efforts to expedite Impact cases were successful. For example, for all felonies, the median number of days between arrest and final disposition increased from 62 to 77, and between arraignment and final disposition the period almost doubled (from 34 to 63); the percent of trials exceeding the 60-day arrest-to-trial standard increased from 50 to 100 percent. However, for Robbery I cases, for example, the period between arrest and final disposition remained constant at 86 days, and between arraignment and final disposition the period declined slightly from 71 to 64 days. But the percent of robbery trials exceeding the 60-day standard also rose (from 36 to 69 percent), indicating that even Impact cases that went to trial fared less well in meeting the standard.

Of the four factors hypothesized as affecting delay, pretrial custody status and type of disposition had either no effect or mixed (i.e., inconclusive) effects. Court calendar crowding or case backlog increases had a small but significant effect, tending to increase delay from 5 to 9 percent for all felonies, depending on the time period. The largest effect was due to type of defense counsel; compared with public

defenders, defense by retained or court-appointed counsel resulted in a 76-percent increase in delay in 1973 and a 48-percent increase in delay in 1974 for all felonies.

The frequency of continuances showed little year-to-year change except for contested cases (up from 45 to 60 percent of cases); the average duration per case showed only small changes (up from 7 to 9 days for uncontested cases and down from 14 to 13 days for contested cases). The major shift was in the agencies responsible for continuances; for both uncontested and contested cases, the proportion attributable to the prosecution declined from year to year, whereas the proportion attributable to the court rose.

THE USE OF JURORS AND WITNESSES

Based on mail survey responses of victims and other witnesses involved in cases that were closed between March and August 1974, the data show there were 2.5 victim appearances per disposition, each appearance averaging 1.8 hours, with a total number of appearance-hours per disposition of 4.5. Comparable figures for other witnesses were 3.0 appearances per disposition, 1.9 hours per appearance and 5.7 appearance-hours per disposition.

Responses of jurors who served during June 1975 revealed that although their productive time seemed to be equally split between civil and criminal matters, idleness was substantial (over 40 percent of total juror time).

LAY PARTICIPANT ATTITUDES

In both jurisdictions response rates of lay participants to mail surveys were quite high, although jurors were more likely to return questionnaires (85 percent) than victims or other witnesses (62 to 70 percent). It took a number of expensive reminders (postcards, letters, telephone contacts) to obtain the rate of return for the latter groups. Personal interviews with jailed defendants were readily arranged, but there were major difficulties in reaching defendants whose cases were dismissed or acquitted or who received a nonincarceration sentence.

Jurors' attitudes were very favorable toward almost all aspects of the court system. Most felt that justice was done and were satisfied with their experience; almost all (85 percent) said their participation helped bring about justice; all said they understood the court system better as a result of their experience. In contrast, victims and witnesses evinced largely neutral attitudes toward the quality of the court system, but 70 percent felt their participation helped bring about justice and most felt they were treated with courtesy and respect. Ninety percent of all groups said they would be willing to cooperate again. After participation, jurors' attitudes toward the court system improved, but witness and victim attitudes worsened.

All of the indicators of attitudes toward the court system were highly interrelated, indicating the presence of a "halo effect." That is, a positive response regarding satisfaction with one's experience was highly correlated with a positive response regarding one's perception of the quality of the system and its professionals and the feeling that one's participation helped bring about justice. The findings of the correlational analysis between attitudes and their determinants were:

1. Background variables of lay participants were *not* associated with attitudes; attitudes were more associated with the participant's role and experiences.
2. Whether or not something was done to compensate a victim for loss or harm suffered (and 60 percent reported that something was done) was *not* associated with attitudes.
3. Only 60 percent of the witnesses learned about outcomes in their cases through the court process or court officials, although not being told did not seem to affect attitudes.
4. The policy changes associated with improved attitudes were:
    a. Reducing the time spent waiting.
    b. Reducing a number of other problems, such as losing time from work and parking and transportation difficulties (the latter were cited often by victims and witnesses).
    c. Giving victims and witnesses a chance to tell their own stories and having court officials, judges, and attorneys treat them with respect and courtesy.
    d. Reinforcing the feeling that the participants' services helped bring about justice.

The questionnaire instrument we devised for interviewing defendants seemed to work well. Because the very small sample of defendants interviewed was so biased (e.g., it included a disproportionately high fraction of incarcerated defendants), any interpretations about what the obtained data "mean" are not possible. There were certain themes that emerged from this sample, however. Considering defendants in both jurisdictions as one group,[5] convicted defendants felt their sentence was "too harsh" and "harsher" than others received by defendants who had similar prior records and who had committed similar crimes. Most defendants felt it was better to have a retained counsel (87 percent), although most (53 percent) were represented by a public defender. Almost half would want a different lawyer in the future, with this trend appearing somewhat stronger for those who were *not* convicted. (It is interesting to note that the type of defense counsel who actually represented the defendant seemed unrelated to whether he was found guilty.) Almost half felt they were pressured (primarily by their defense attorney) into plea bargaining, and those who plea bargained had neutral attitudes toward the outcome.

GAUGING OVERALL PERFORMANCE

Given the findings discussed above, it seems reasonable to conclude that, for issue areas for which data were gathered in both years (1973 and 1974), performance improved from year to year; the experiment's goals were largely achieved; and the experiment had a substantial effect in improving overall performance (possibly as a result of both greater prosecutorial effort relative to the defense and of a different organization in the prosecution of Impact offenses). Only delay for the entire felony caseload was worse, but performance on all other issues was either better or no worse.

[5] Because of the very small sample size of each jurisdiction.

The additional short-term costs in the district attorney's office amounted to about 25 percent in felony-assigned staff, or about 13 percent in overall staff; no other short-term incremental costs appeared to be imposed on the court or in the public defender's office as a result of the experiment. Over the longer term, if the experiment is continued, there may be additional costs arising in other parts of the system as they adapt and respond, but these potential costs cannot be estimated currently.

Chapter 4

COMPARING PERFORMANCE IN TWO JURISDICTIONS

Interpretation of the observed differences in performance measures between two jurisdictions must necessarily be cautious, for these differences may reflect disparities in the nature of the two systems as well as their relative effectiveness. However, some interjurisdictional comparisons are reasonable and justified no matter how profoundly the systems differ. For example, the proportion of persons in the community on probation and/or the proportion of felony arrestees whose cases were rejected for prosecution are meaningful to the public, irrespective of the differences in the court system that produced them. However, comparing relative performance of two prosecutors' offices in terms of rejection rate would have little meaning if the screening in one jurisdiction is done mainly by magistrates and in the other by prosecutors, as in fact was the case in our two demonstration jurisdictions.

Multnomah and Dade counties do differ profoundly in most of the important "comparability features." These include total case input, caseload per practitioner, certain background characteristics of the defendant population (such as the proportion of ethnic minorities), organization and calendaring in the circuit courts, organization of the prosecutors' offices, case-screening responsibility and procedure, plea bargaining practices, continuance policies and speedy trial standards, the existence and nature of pretrial intervention programs, and so on. We need not recite the details here,[1] but we summarize and illustrate below, in largely qualitative terms, where and why the comparisons *can* be made and *cannot*. The two jurisdictions are compared for the year 1974.

CASE SCREENING

Despite the differences in screening procedure (in Dade screening is done mainly by a lower court magistrate and in Multnomah by the prosecutor), comparisons of the screening threshold may be regarded as meaningful from the *communities'* viewpoints. Comparing Burglary I bookings in Multnomah and Breaking and Entering (B&E) cases in Dade, gross screening output was similar; that is, about one-third are rejected and about two-thirds are filed as felonies. In robbery bookings there were notable differences, with the rejection rate being markedly higher in Multnomah (51 versus 38 percent). But since adequate data on reasons for screening out cases were not available in Dade, adequate comparisons of the operation of the charging threshold in the two jurisdictions were not possible.

Only fragmentary impressions of the relative performance in charging accuracy could be gained. Cases involving (Multnomah) Impact offenses (Robbery I and dwelling Burglary I) seemed to be more accurately charged in Multnomah; for example, Multnomah did better in overall conviction rates for robbery charges as filed (67

[1] See Secs. V, VI, VII, and IX, R-1918-DOJ, for descriptions, statistical overviews, and comparisons of the two jurisdictions.

compared with 33 percent) and dwelling burglary charges as filed (79 compared with 46 percent), although not for nondwelling burglary (25 compared with 54 percent).

PLEA BARGAINING

Only limited comparisons of plea bargaining performance were possible, primarily because *sentence agreements are the predominant type of bargain in Dade and no written record of such agreements was contained in the case files.* In particular, it was not possible to compare adequately the sentence severity concessions embodied in the plea agreements of the two jurisdictions.

A high rate of straight pleas (i.e., to all charges and counts as originally filed, but possibly with tacit sentence agreements) characterized both Multnomah and Dade in the cases involving the Impact offenses in the former and the comparable offenses in the latter—for example, 80 percent or more in robbery cases. But for cases in which the other offense type (nondwelling Burglary I or nondwelling B&E) analyzed in this study was charged, there was a prominent difference evidenced between the two counties; namely, Multnomah engaged in charge and count bargains far more frequently (53 percent compared with 4 percent). This is explained by the policy differences in the two jurisdictions: In Dade, charge bargaining is frowned on; in Multnomah, it is sanctioned for non-Impact cases, but not for Impact cases.

SENTENCE VARIATION

The pattern of sentencing for robbery and burglary convictions was similar in the two jurisdictions—conviction usually being at the highest level and a substantial prison term then being imposed. On the average, the sentence severity score (Index A) was moderately higher in Multnomah for Robbery I convictions (26.8 versus 22.0).

The pattern of sentencing differed between the counties for convictions on Burglary I or B&E charges. In Multnomah, convictions were predominantly at the highest statutory level (felony A), and prison terms were imposed in roughly one-half of the sampled cases. In Dade, convictions were mostly at the second most serious statutory level (second-degree felony), which was generally the highest level applicable, and roughly one-third of the defendants were given prison terms. Again, the average sentence severity score turned out to be higher in Multnomah (19.3 versus 13.4).

Little of the sentence variation in exemplary-offense cases in either jurisdiction was accounted for by the set of "illegitimate" factors we hypothesized, while the "legitimate" factors (i.e., the nature of the defendant's criminal conduct and his prior record) did account for most of the sentence variation that we could statistically explain.

EVENHANDEDNESS

To distinguish the evenhandedness of dispositions between Multnomah and Dade counties, we attempted to relate the nature of a defendant's disposition to his

pretrial custody status, type of defense counsel, and ethnic group but found no clear distinctions on this basis (except that the defendant appeared to gain a small advantage in being represented by the public defender in Multnomah but not in Dade).

Sentence severity scores also manifested no consistent patterns that enabled us to distinguish between the two jurisdictions as to evenhandedness. But there was evidence of a lack of evenhandedness in specific circumstances in both court systems —namely, more severe sentences for B&E convictions in Dade being associated with minority status (22 percent for blacks, 51 percent for other minorities), retained counsel (26 percent), and pretrial jail confinement (26 percent); and pretrial custody status being correlated with the severity of Robbery I sentences in Multnomah.

THE ROLE OF PRIOR RECORD

Judged by comparison of dispositional measures, prior criminal record did not turn out to be a governing factor in the dispositions of the 1974 exemplary-offense cases either in Multnomah or Dade.

More serious prior criminal records increased the sentence severity scores consistently in both Robbery I and Burglary I cases in Multnomah and in B&E cases in Dade, but did not significantly affect these scores in Dade robbery cases. But lacking a persuasive explanation for the latter difference, we cannot infer that it reflects an essential distinction between the two counties as to the importance of prior record.

DELAY

The clearest distinction in the performance of felony proceedings between the two jurisdictions appeared to be in the duration of these proceedings. Much less time generally elapsed between stages of the proceeding in the Multnomah court system. Corresponding events take from 50 to 100 percent longer in Dade to occur; for example, median number of days from arrest to final disposition was 77 in Multnomah compared with 109 in Dade for all felony cases. The comparable figures for the fastest of the slowest 10 percent of cases tells a similar story—151 compared with 270 days.

However, a *smaller* proportion of cases in Dade exceeded the Florida Supreme Court-imposed 180-day standard for the maximum time between arrest and trial than exceeded the statutory 60-day standard for this maximum time in Multnomah; for example, 100 percent did so in Multnomah compared with only 8 percent in Dade, for all felony cases. (This performance relative to the differing standards, that is, 180 days versus 60 days, was also true for the periods from arrest to dismissal, to guilty plea, or to final disposition.)

Statistical performance measures attested to the marked difference in continuance practices between the two jurisdictions (e.g., the average number of continuance days per case was over ten times greater in Dade—97 compared with 9 days in uncontested cases).

There was no significant difference between Multnomah and Dade in the pattern of effects on case processing time attributable to the factors we hypothesized as being

influential. However, these factors accounted for substantially more of the variation in case processing time in Dade than they did in Multnomah (e.g., 52 compared with 33 percent in robbery cases). And certainly, a major factor accounting for the large difference in *average* delay between the two jurisdictions is that of the resources relative to the total caseload; practitioner caseloads are much higher in Dade than in Multnomah.

USE OF VICTIMS, OTHER WITNESSES, AND JURORS

The average number of appearances per witness, the average number of appearances by all witnesses per disposition, and the average time per witness appearance all tended to be higher in Dade—most likely a by-product of the differences observed in the length of proceedings and in the continuance practices in the two counties. On the average, victims devoted about 4.5 hours of their time per disposition in both jurisdictions, but other witnesses spent about 14 hours per disposition in Dade compared with 5.5 hours in Multnomah.

The pattern of juror activities appeared not to differ materially between the two counties. Jurors were idle somewhat less than half of their tour of duty in both systems.

LAY PARTICIPANT ATTITUDES

In the previous chapter, we indicated that there were more similarities than differences in lay participant attitudes and their determinants in the two counties. Here, we mainly summarize the few notable differences that were evident.

The only consistent trends across all three groups (jurors, victims, other witnesses) were: reflecting characteristics of the general population, minority group lay participants were more prevalent in Dade (40 compared with 10 percent), and Dade lay participants said that the problems they encountered were about 20 percent more serious or annoying.

Among jurors the following differences emerged: Overall attitudes were somewhat more favorable in Multnomah, but there was more satisfaction with the performance of judges and attorneys in Dade. Also, jurors in Dade were more realistic in assessing the extent of plea bargaining.

Sixty percent of Multnomah victims compared with only 40 percent of Dade victims reported that something was done to compensate them for the loss or harm they suffered. Other witnesses in Multnomah exhibited more favorable opinions about the court process and their experiences than in Dade—for example, the time it took to settle cases. A higher proportion of the Multnomah group (60 percent) compared with the Dade group (40 percent) learned of the outcome of the case through the court process or from court officials. But it is important to note that *both* victims and other witnesses reported that 40 percent of the time they *never* found out what happened, that 35 percent were not warned in advance of any postponements or reschedulings, and that 40 percent were never told the reason for rescheduling.

OVERALL PERFORMANCE

We find it infeasible to compare "overall" performance between these two court systems for a variety of reasons. Even if it were possible to devise an acceptable composite performance measure that would reflect the importance of one activity (say, screening) relative to another (say, case processing delay), we would be defeated here by the inadequacy of the data available in Dade (e.g., the absence of information on sentence bargaining).

But we may express a "collective" impression of the relative performance in the several issue areas. In some—sentence variation, evenhandedness in disposition, and sentencing—the two jurisdictions performed similarly. In others—charging accuracy, case processing delay—Multnomah seemed to do better than Dade. In still others—applying the charging threshold, the plea bargaining balance—our data do not suffice for a judgment.

To develop a set of *detailed and comprehensive guidelines* for making interjurisdictional performance comparisons would require a very broad and expensive study across many (possibly over 50) jurisdictions. But we need not wait until this is done. By conscientiously examining the relevant differences, as well as similarities, between two jurisdictions, one can spell out the qualifications to be made to comparisons of specified performance measures in various issue areas.

Chapter 5

GENERAL FINDINGS AND IMPLICATIONS OF THE STUDY

A number of *general* findings and implications emerged from this study. They involve the potential benefits of an integrated performance measurement program, the problems of data availability and case sampling, the need for extensions to our demonstration work, and, most important, the relationships between the potential capabilities of existing or planned information systems and the more comprehensive approach outlined in this study for measuring the performance of prosecution, defense, and court agencies. In addition we present very gross cost estimates that bound the range of alternatives discussed in this study.

THE DESIRABILITY OF AN INTEGRATED PERFORMANCE MEASUREMENT PROGRAM (IPMP)

Our study has shown that there is a richness of court system performance information which, if the jurisdictions in which we worked are representative, is largely untapped. (We say this with the full understanding that the data elements recorded in the various agencies' files were by no means complete.) The careful collection of specified data elements, the computation, grouping, and cross-tabulation of performance measures, and the analysis (using multivariate statistical techniques) of what factors account for the variation in key performance measures can provide greatly strengthened informational bases for officials in court, prosecution, and public defender agencies to improve criminal proceedings.

The actions to be taken jointly by the court, prosecution, and public defender agencies in a jurisdiction to strengthen the informational and analytical base for measuring their performance may be visualized as an *integrated performance measurement program* (IPMP). A fairly comprehensive IPMP would consist of:

- An enumeration of required data elements (or categories) and performance measures.
- Standardized data collection and output forms for each policy issue area of interest (charging, plea bargaining, sentence variation, evenhandedness, delay, case processing efficiency (separately for each of the three agencies), or others of interest to particular jurisdictions).
- Flexible, modular, software (i.e., computer programs) packages for computing, displaying, and analyzing performance measures within each issue area—for example, for performing cross-tabulations and for applying multivariate regression models that help to explain conviction probability, delay, and sentence severity imposed.
- Guidelines for conducting case audits at each major decision point (screening, guilty plea, trial, sentencing) in the proceeding—using either outside practitioner-consultants or in-house supervisory personnel.
- The administration of sampling plans and standard mail survey question-

naires and the analysis of responses of victims, other witnesses, and jurors (using appropriate software packages).

- The administration of sampling plans and standard personal interview questionnaires and the analysis of responses of defendants (using appropriate software packages).

An IPMP can be designed for several uses: the routine retrospective monitoring of performance *within* a jurisdiction; the retrospective evaluation of policy, organizational, and procedural innovations *within* a jurisdiction;[1] and (to a lesser extent) the retrospective comparison of performance *between* jurisdictions. To the extent that the statistical models succeed in explaining and predicting conviction probability, delay, and sentence severity imposed, it may be possible to use them in a limited way to do *prospective* performance analysis, that is, to forecast some effects of planned policy or organizational changes or of anticipated changes in agency workloads. As visualized here, an IPMP would *not* be designed to track pending cases; other existing or planned information systems perform this function.

If data collection procedures and software packages were flexible and modular in design, the scale and scope of an IPMP could be tailored to individual jurisdictions. For example, the three agencies in a using jurisdiction could decide whether to embrace all elements (e.g., to include case auditing and defendant interviews) and whether to measure performance in all of the listed issue areas (e.g., to include the measurement of case processing efficiency in the prosecutor's and public defender's office, as well as in the court). What would be vital to proper tailoring is a clear enunciation by agency officials of the management and policy issues on which performance measurement should focus.

This study is a *first step* toward the design of an IPMP. We have enumerated required data elements and performance measures and, with varying degrees of success, have devised and applied statistical models to explain key performance measures. We have also designed and applied mail and personal interview questionnaires to the four classes of lay participants. More work needs to be done, however, and its nature is discussed below.

In this study we have *not* made a serious attempt to analyze the resource implications of inaugurating an IPMP. Later in this chapter, however, we make very rough cost estimates for two bounding cases, based on actual costs incurred in this study and on rough guesses of costs of activities not covered in this study. Actual costs would vary from jurisdiction to jurisdiction and would depend on the scale of performance measurement desired and the extent to which *existing* or *planned* information systems were used to measure performance as opposed to inaugurating a new and more comprehensive information system designed solely for an IPMP. In any event, the statistical analysis necessary (developing statistical models and applying techniques such as multivariate regression) for more fully explaining the factors associated with performance changes would require a competent statistician or econometrician (at least part-time) who has acquired detailed knowledge of criminal court systems.

[1] Examples of such innovations include altering the plea bargaining policy, shifting the charging threshold, introducing a program for the special handling of habitual offenders, changing the policy of case assignment to the public defender's office, modifying the court calendar system, and instituting arrangements for lessening unwarranted sentence variation.

In this study we have demonstrated that it was *feasible to apply* statistical performance measures in two jurisdictions. This demonstration was more successful in applying performance measures to certain policy issues than to others because of inherent differences in the precision of the performance measures (e.g., in those that measure charging threshold compared with those that measure delay) or because of differences in the availability of data (e.g., the availability of data on sentence agreements in Multnomah compared with its unavailability in Dade for measuring plea bargaining effects). Whether the institutionalization of an IPMP is *practical* and has *utility* (i.e., whether its benefits outweigh its incremental costs) is as yet undetermined. It would require a pilot demonstration in at least one jurisdiction over some period of time, after which officials in the using agencies would have to make their own judgments about practicality and utility.

DATA AVAILABILITY AND METHODOLOGY

Case Audits

Our pilot-case auditing exercises (of cases in which there was a plea of guilty) in the demonstration jurisdictions strongly suggest that case auditing provides *complementary* information about qualitative factors that aid in the interpretations of the statistical performance measures. (By their very nature, case audits are much more expensive per case included than the data collection required to develop statistical performance measures. Thus, with limited resources, audit samples are inevitably too small to stand alone as a *substitute* for statistical performance measures.)

We believe that the average practitioner regards case auditing to be a natural and nontechnical way of revealing performance. His confidence in the correctness of what is shown by statistical performance measures is undoubtedly increased when the results of (even quite limited) case auditing confirm the statistical story. Another possible benefit of case auditing is that it may help reveal the explanations for the "behavior" of statistical indices. And, finally, it may considerably strengthen the credibility of *inter*jurisdictional comparisons by means of statistical measures. It would appear that, in general, case audits would best be conducted in the *initial* phases of a project to measure court performance. (Our suggestions for broadening case auditing to test its value more fully are discussed below.)

Data Availability and Sampling Problems

A salient lesson in our attempt to demonstrate the application of performance measures in two selected jurisdictions was that many necessary or desirable data elements normally recorded in various files were missing from the customary records, and some were simply not recorded at all. And this is likely to be the situation in other jurisdictions as well.

Among the data elements that had been (at best) *incompletely recorded* and preserved were defendant-related characteristics such as ethnicity, prior criminal record, occupation and employment, family status, income, and transiency; the number of appearances per victim or other witness in the course of a proceeding; data describing how judges apportion their time among judicial tasks; and attribution of continuances to the responsible movant(s).

Among the data elements that were *not recorded* at all were the apportionment of time among the principal activities of prosecutors, public defenders, and jurors; background characteristics of suspects whose cases were screened out before arraignment on felony charges; full information on the outcome of plea bargaining, including the nature of any sentence agreement reached; judicial statements of the rationale for sentences in individual cases; detailed reasons for case dismissals in lower court; duration of appearances of victims and other witnesses; and information on the attitudes of lay participants and defendants toward their experiences and toward the performance of the court agencies.

Our demonstration work also imbued us with a deeper appreciation of the need to tailor case sampling to the type of the data element sought. Events of interest in the performance measurement of criminal proceedings differ dramatically in their expected frequency of occurrence. When data on rare events are required (e.g., data on the outcomes of jury trials for a specified offense wherein a minority defendant is represented by retained defense counsel), one must employ well-planned "oversampling." Fortunately, many key events in court proceedings occur frequently enough that moderate (on the order of 100) case sample sizes suffice as a basis for reliable inferences.

DESIRABLE EXTENSIONS

We feel that a fuller foundation for the design of an operational IPMP would be provided by the following extensions in scope and in refinement in methodology to our demonstration work:

- Classes of data that were not recorded or were incompletely recorded in Multnomah and Dade counties should be collected and analyzed elsewhere. *Evenhandedness in screening* should be analyzed with a proper body of data containing appropriate defendant-related characteristics. The *allocation of prosecutors' and public defenders' time* to their various activities is another performance area warranting examination and would need a proper body of data.
- The assessment of case auditing should be broadened in the *screening area* (to include rejected cases) and also extended to the *trial area*, so that our inferences as to the value of case auditing as a complement to statistical performance measurement can be tested more fully.
- Improved statistical models should be constructed to help explain performance outcomes in criminal proceedings. Those we developed for explaining sentence outcomes and delay in proceedings worked fairly well but need further refinement. Because we were unsuccessful in explaining the determinants of conviction probability, we believe much more theoretical and empirical work is necessary. We speculate that data on the seriousness of the crime incident, on mitigating and exacerbating circumstances of the defendant and the crime incident, and on factors describing the strength of the case at the time of screening are relevant for constructing better conviction probability models.

POTENTIAL CAPABILITIES OF OTHER INFORMATION SYSTEMS FOR PERFORMANCE MEASUREMENT

A number of state- and local-agency-level information systems already exist or will be installed soon in several state and local jurisdictions. Given the considerable resources that will be devoted to these systems, it is clearly important to indicate their potential capabilities for the kind of performance measurement envisaged for an IPMP. In particular, it is important to compare their potential performance measurement capabilities (e.g., how many issue areas can be analyzed and in what depth?) with an IPMP under two conditions: (1) if only those data elements already collected by these information systems are available; (2) if a modest, inexpensive extension in data collection is added.

Examples of current or planned information systems are:

- CCH (Computerized Criminal Histories): a component of the LEAA's Comprehensive Data System (CDS) Plan.
- OBTS (Offender-Based Transaction Statistics System): a component of LEAA's CDS Plan.
- SJIS (State Judicial Information System).
- PROMIS (Prosecutor's Management Information System).

All of these systems are designed to track defendants or cases through that part of the criminal justice system with which they are concerned and all rely on the local criminal justice agencies for input data. The first three are state-level or multistate systems, whereas the fourth is intended for use by local prosecutors' offices. CCH focuses on information concerning the identity location, characteristics, and description of the known criminal offender. OBTS is a statistical system describing the aggregate experiences of arrested individuals from their encounter with the police through court processing and entry into, and exit from, the correctional system. CCH and OBTS are components of LEAA's Comprehensive Data System Plan. The relevant part of SJIS (i.e., the entry and passage of people and cases through courts of general jurisdiction) is designed to evaluate the organization, practice, and procedures of the courts in a state; assist with dispatch of judicial business; and facilitate technical assistance and long-range planning activities. The conditions for a state's participation in SJIS include a commitment for the state judicial system to provide the information needed for a "comprehensive criminal justice data system." As defined by LEAA guidelines, such a system must include a CCH file, an OBTS file, and a statistical analysis center.

PROMIS is designed to aid local prosecutors in identifying the more serious criminal cases to which prosecutorial priority should be given, to aid in controlling or eliminating impediments to effective prosecution, to aid in regulating the exercise of prosecutorial discretion so as to maintain evenhandedness and consistency, and to aid in conducting relevant analyses of screening and prosecution of criminal cases. PROMIS tracks the arrested person from arrest through processing and disposition in the lower and felony courts. As indicated at the outset, an IPMP would *not* track open cases, but is designed to do retrospective performance analysis.

Table 8 displays the data elements or categories collected by these four information systems that are relevant to the kind of performance measurement and analysis demonstrated in this study. For purposes of comparison, we also display those data

elements or data categories collected in our demonstration work, as well as those that *would* be collected if the desirable extensions to an IPMP, discussed above, were implemented. From Table 8 we observe that *none* of the existing or planned systems obtains data on lay participant attitudes and experiences, on how practitioner or lay participant time is used, or on all potentially relevant characteristics of defendants (e.g., CCH/OBTS and SJIS omit data on education, employment, transiency, family status, and income level; and PROMIS omits data on education, income, and family status). However, PROMIS is designed to collect data on *all* filed and final charges and counts and on other aspects of plea bargaining such as sentence bargaining. CCH/OBTS and SJIS do not collect data on rejection actions and their reasons, nor on the fact and nature of sentence bargaining.

Given these differences and similarities in data categories to be collected, what are the implications for performance measurement?[2] To answer this question, we consider four options for each of the following: the combined CCH/OBTS, SJIS (with CCH/OBTS), and PROMIS.

Option I: The basic system with existing data elements, assuming that a simple software package is available that can compute the values of performance measures and produce summaries and cross-tabulations.

Option II: Option I *plus a few added data elements* that can be collected inexpensively.

Option III: Option I *plus an additional software package* (e.g., statistical models and standard multivariate statistical analysis routines) for estimating the independent effect of important factors on delay and sentence severity. Additional research is needed to develop a similar package for explaining conviction probability.

Option IV: Option II plus Option III.

To our knowledge, more sophisticated statistical software packages (as in Options III and IV) are not planned for any of these systems; at best, simpler software packages of the Option I variety will be available.

Potential Capabilities of CCH/OBTS

Table 9 compares the potential capabilities of the four CCH/OBTS options with two versions of an IPMP. For a specified issue area, an entry of "No" indicates that the option cannot measure performance; an entry of "Partial" indicates that certain relevant data items are still needed or that cross-tabulations are the best available tool or both; and an entry of "Yes" indicates that the available data elements and the software packages are adequate for performance measurement.

This table is a means of showing that adding a few more data elements (listed

[2] The following comparisons of performance measurement capabilities of the various systems are viewed only in terms of the ensemble of issue areas we considered for an IPMP. There may be *other* issue areas (e.g., bail and O.R. policy) of interest to a jurisdiction, too; thus, comparisons among the systems might look different depending on whether or not an IPMP was designed to address these issues.

Table 8

A Comparison of Data Elements and Categories Collected That Are Relevant to Performance Measurement[a]

Data Element or Category	Information System					
	CCH	OBTS	SJIS (with CCH/OBTS)	PROMIS	IPMP (as illustrated in this study)	Improved IPMP (with extensions)
Defendant-related characteristics:				(For filings *and* rejections)	(For filings only)	(For filings *and* rejections)
Age	Yes	Yes	Yes	Yes	Yes	Yes
Ethnicity	Yes	Yes	Yes	Yes	Yes	Yes
Prior record	Yes	Yes	Yes	Yes[b]	Yes	Yes
Education	No	No	No	No	Yes	Yes
Income	No	No	No	No	Yes	Yes
Employment	No	No	No	Yes	Yes	Yes
Family status	No	No	No	No	Yes	Yes
Transiency (years living in jurisdiction)	No	No	No	Yes	Yes	Yes
Pretrial custody status	No	Yes	Yes	Yes	Yes	Yes
Type of defense counsel	Yes	Yes	Yes	Yes	Yes	Yes
Date and nature of each major event from arrest through disposition (including sentencing, if applicable)	Yes	Yes	Yes	Yes	Yes	Yes
Most serious arrest offense	Yes	Yes	Yes	Yes	Yes	Yes
Full specification of arrest offenses	Yes	No	Yes	Yes	Yes	Yes
Factors describing the seriousness of the crime incident	No	No	No	Yes	No	Yes
Mitigating/exacerbating circumstances of defendant or crime	No	No	No	Yes	No	Yes
Strength of case at screening	No	No	No	Yes	No	Yes
Screening actions:						
Rejection	No	No	No	Yes	Yes	Yes
Reasons for rejection	No	No	No	Yes	Yes	Yes
Most serious charge filed	Yes	Yes	Yes	Yes	Yes	Yes
Full specifications of filed charges and counts	Yes	No	Yes	Yes	Yes	Yes
Final most serious charge	Yes	Yes	Yes	Yes	Yes	Yes
Full specifications of final charges and counts	Yes	No	Yes	Yes	Yes	Yes
Other types of bargain agreements:						
Fact of sentence assurance or agreement	No	No	No	Yes	Yes	Yes
Nature of sentence agreement	No	No	No	Yes	No	Yes
Agreement to drop other pending cases	No	No	No	Yes	Yes	Yes
Full specifications of sentence elements imposed	Yes	Yes, but fine excluded	Yes	Yes	Yes	Yes

Continuances:						
Number per case	No	No	No	Yes	Yes	Yes
Duration of each continuance	No	No	No	Yes	Yes	Yes
Attribution to defense, prosecution, court, or joint	No	No	No	Yes	Yes	Yes
Data elements required to estimate weighted caseload performance measures:						
Judge[c]	No	No	Yes (gross information only)	Partial[d]	Yes	Yes
Prosecutors	No	No	No	Partial[d]	No	Yes
Public defenders	No	No	No	Partial[d]	No	Yes
Use of lay participant time:						
Number of victim/witness appearances per disposition[e]	No	No	No	No	Yes[f]	Yes[g]
Duration of victim/witness appearance[e]	No	No	No	No	Yes[f]	Yes[g]
Proportion of juror time spent:						
In idleness	No	No	No	No	Yes[f]	Yes[g]
In voir dire, criminal	No	No	No	No	Yes[f]	Yes[g]
In trial, criminal	No	No	No	No	Yes[f]	Yes[g]
In voir dire, civil	No	No	No	No	Yes[f]	Yes[g]
In trial, civil	No	No	No	No	Yes[f]	Yes[g]
Questions asked in defendant personal interview questionnaire and mail survey questionnaires for victims, other witnesses, and jurors[h]	No	No	No	No	Yes	Yes

SOURCES: We draw descriptive information and information on data elements collected by CCH, OBTS, and PROMIS from the National Advisory Commission on Criminal Justice Standards and Goals, *Criminal Justice System*, U.S. Government Printing Office, Washington, D. C., 1973, pp. 100-101; *PROMIS Briefing Series*, Institute for Law and Social Research, Washington, D. C., October 1974, especially No. 1, "Management Overview of PROMIS," and No. 17, "Interface with Other CJIS;" and "Data Base and Data Element Dictionary" (App. D to Vol. I of the six volumes of PROMIS software documentation). Information on data elements to be collected by SJIS is drawn from *State Judicial Information System Project, Requirements Analysis Subcommittee Final Report*. The Institute of Judicial Administration, Inc., New York, April 1975.

[a]The entry indicates whether the data element is recorded.

[b]Arrests and convictions only, not sentences imposed.

[c]See Chap. 2 of this report.

[d]PROMIS contains the data elements associated with the computation of weighted caseload *except* for time data, that is, the time associated with each activity or proceeding connected with a case (letter from William Hamilton, President, Institute of Law and Social Research, April 30, 1976).

[e]Separately for victims and other lay witnesses.

[f]Based on memories of victims, other witnesses, and jurors who responded to mail surveys.

[g]These data elements would be collected by personnel in prosecution, defense, and court agencies at the time when these lay participants appear.

[h]See the questionnaires in Apps. G and I, R-1918-DOJ.

Table 9

Potential Performance Measurement Capabilities of CCH/OBTS[a]

Issue Areas	Option, Data Elements Collected, Software Packages				Reference IPMP Options	
	I Existing Cross-Tabs Only	II Improved[b] Cross-Tabs Only	III Existing Cross-Tabs and Regression Models	IV Improved[b] Cross-Tabs and Regression Models	As Demonstrated in This Study	Improved with Extensions Specified Above
Charging threshold	No	No	No	No	Yes	Yes
Charging accuracy	Yes	Yes	Yes	Yes	Yes	Yes
Evenhandedness in screening	No	No	No	No	No	Yes
Effects and concessions in plea bargaining	Partial	Partial	Partial	*Yes*[c]	Yes	Yes
Determinants of conviction probability:						
Legitimate factors	No	No	No	No	No	Yes
Illegitimate factors (evenhandedness)	Partial	Partial	Partial	Partial	Partial	Yes
Estimation of sentence variation	Yes	Yes	Yes	Yes	Yes	Yes
Determinants of sentence variation:						
Legitimate factors	Partial	Partial	Partial	*Yes*	Yes	Yes
Illegitimate factors (evenhandedness)	Partial	Partial	Partial	*Yes*	Yes	Yes
Estimation of delay	Yes	Yes	Yes	Yes	Yes	Yes
Determinants of delay	Partial	Partial	*Yes*	*Yes*	Yes	Yes
Analysis of continuances	No	No	No	No	Yes	Yes
The use of jurors' time	No	No	No	No	Yes	Yes
Duration and number of appearances of victims and witnesses	No	No	No	No	Yes	Yes
The use of practitioners' time:						
Judges	No	No	No	No	Partial	Yes
Prosecutors and public defenders	No	No	No	No	No	Yes
Attitudes of lay participants and their determinants	No	No	No	No	Yes	Yes

[a]The entries indicate the extent to which the specified option can measure performance in the indicated issue areas.

[b]Additional data elements that can be added inexpensively to those already specified in the existing system are:

- Defendant community ties (education, income, employment, family status, transiency).
- Other types of plea bargains: the fact and nature of sentence agreements; agreement to drop other pending cases.
- Time trend of inmate population in local jails and state prisons.

[c]An italicized entry indicates that the specified option improves performance measurement capability over Option I for the indicated issue area.

at the bottom of Table 9) without adding a more sophisticated software package (Option II) would not appreciably improve the performance measurement capability. Adding *both* (Option IV) would (by design) clearly provide the most benefits. *Option IV would enable better analysis of the plea bargaining balance, the independent effects of important factors on delay, and the independent effects of legitimate and illegitimate factors on sentence severity imposed.* However, compared with an improved IPMP, *none* of the four options are capable of performance measurement in the following issue areas: charging threshold, evenhandedness in screening, the effect of legitimate factors on conviction probability, continuances, the use of practitioner and lay participant time, and the attitudes (and their determinants) of lay participants.

Potential Capabilities of SJIS

Because SJIS must include CCH/OBTS files, an inspection of Table 9 reveals that the capabilities of the various SJIS options (assuming the same added data elements for Options II and IV) would be nearly identical with the corresponding CCH/OBTS options, with only one major difference: The basic *SJIS is capable of very gross estimates of the use of judicial time.*

Potential Capabilities of PROMIS

The basic system (Option I) will have *better* potentialities for performance measurement than the basic CCH/OBTS (Option I), as shown in Table 10. In addition to all of the CCH/OBTS Option I capabilities, PROMIS Option I has a good capability in the charging threshold area and a partial capability in analyzing evenhandedness in screening and the effects of legitimate factors on conviction probability. Adding only the additional data elements (Option II) listed at the bottom of Table 10 leads to no significant improvement. However, adding *both* additional data elements and software (Option IV) provides the most benefits. *Option IV would improve the capability to analyze evenhandedness in screening, the plea bargaining balance, the independent effects of important factors on delay, and the independent effects of legitimate and illegitimate factors on conviction probability and sentence severity imposed.* However, compared with an improved IPMP, none of the four options are capable of performance measurement in the following issue areas: the use of practitioner and lay participant time and the attitudes (and their determinants) of lay participants. However, because PROMIS already collects much of the data needed for estimating practitioners' weighted caseload, if practitioner time data associated with various activities and proceedings in each case were sampled outside PROMIS, any of the PROMIS options would be able to measure the use of practitioner time.

* * *

In summary, none of the existing or planned information systems would have the breadth and depth in performance measurement capabilities of an improved IPMP for two reasons: Many data elements are not collected and, to our knowledge, none of the existing and planned systems will have the full array of statistical models and software packages required for analyzing the independent effect of important parameters on delay, dispositional, and sentence outcomes. Each basic

Table 10

Potential Performance Measurement Capabilities of PROMIS (with or without CCH/OBTS, SJIS)[a]

Issue Areas	Option, Data Elements Collected, Software Packages				Reference IPMP Options	
	I Existing Cross-Tabs Only	II Improved[b] Cross-Tabs Only	III Existing Cross-Tabs and Regression Models	IV Improved[b] Cross-Tabs and Regression Models	As Demonstrated in This Study	Improved with Extensions Specified Above
Charging threshold	Yes	Yes	Yes	Yes	Yes	Yes
Charging accuracy	Yes	Yes	Yes	Yes	Yes	Yes
Evenhandedness in screening	Partial	Partial	*Yes*[c]	*Yes*	No	Yes
Effects and concessions in plea bargaining	Partial	Partial	Partial	*Yes*	Yes	Yes
Determinants of conviction probability:						
Legitimate factors	Partial	Partial	Yes	*Yes*	No	Yes
Illegitimate factors (evenhandedness)	Partial	Partial	*Yes*	*Yes*	Partial	Yes
Estimation of sentence variation	Yes	Yes	Yes	Yes	Yes	Yes
Determinants of sentence variation:						
Legitimate factors	Partial	Partial	Partial	*Yes*	Yes	Yes
Illegitimate factors (evenhandedness)	Partial	Partial	Partial	*Yes*	Yes	Yes
Estimation of delay	Yes	Yes	Yes	Yes	Yes	Yes
Determinants of delay	Partial	Partial	*Yes*	*Yes*	Yes	Yes
Analysis of continuances	Yes	Yes	Yes	Yes	Yes	Yes
The use of jurors' time	No	No	No	No	Yes	Yes
Duration and number of appearances of victims and witnesses	No	No	No	No	Yes	Yes
The use of practitioners' time:						
Judges	No	No	No	No	Partial	Yes
Prosecutors and public defenders	No	No	No	No	No	Yes
Attitudes of lay participants and their determinants	No	No	No	No	Yes	Yes

[a]The entries indicate the extent to which the specified option can measure performance in the indicated issue areas.

[b]Additional data elements that can be added inexpensively to those already specified in the existing system are:

- Sentences associated with prior convictions of defendant (if CCH unavailable).
- Data elements comprising the defendant community ties index that are not already collected in PROMIS, that is, education, income, and family status.
- Time trend of inmate population in local jails and state prisons.

[c]An italicized entry indicates that the specified option improves performance measurement capability over Option I for the indicated issue area.

system could function as a partial IPMP as is, and if upgraded (i.e., a few data elements are added for routine collection as well as the required software packages), their performance measurement capability could be substantially improved. In this connection, it appears that among the planned systems considered in this report (CCH/OBTS, SJIS, PROMIS), an upgraded PROMIS would have the most comprehensive performance measurement capability, primarily because it can address case screening issues as well as the issues that can be addressed by CCH/OBTS and SJIS.

Some Final Comments on Benefits, Costs, and Utility of Performance Measurement Systems

Given the preceding discussion on the potential capabilities of existing or planned systems, of planned systems that are upgraded in the ways we specified, and of an improved IPMP, how is a jurisdiction to decide which alternative to choose? The decision would depend on a variety of considerations: the availability of (and which) computerized information system is already installed or planned; the issue areas or policy issues of interest to court, prosecution, and public defender agencies; and the incremental costs and benefits of installing and operating CCH/OBTS, SJIS, and PROMIS, of upgrading these systems as we indicated, and of moving toward an improved IPMP.

Issue Areas of Interest to Agencies. We cannot generalize about which issue areas would be of greatest interest. However, in Chap. 1 we noted that the practitioner interviews confirmed their importance, and we can indicate how the agencies reacted in the two demonstration jurisdictions in which we worked. Officials there received early drafts of our reports and then were briefed and interviewed by a Rand team. In general, officials were enthusiastic and felt that there was great value in the application of performance measures, primarily as an objective way of demonstrating what was going on, how well certain well-defined objectives were being met (e.g., arrest to trial standards), and in explaining why performance measures varied. The chief judge and court administrator in one of the jurisdictions were particularly interested in the applications to charging accuracy, plea bargaining, sentence variation and evenhandedness, the use of lay participant time, and the relationships between lay participant attitudes and the problems they faced (that could be manipulated by policy changes). The court administrator in the other jurisdiction was particularly interested in delay, but indicated that almost all of the findings in our pilot application of performance measures were "new," since the few statistics they did produce had to be manually estimated (due to the lack of computerization). The prosecutors in both jurisdictions were particularly interested in applications to screening, plea bargaining, sentence variation, evenhandedness, treatment of "habitual" or "career" offenders, and the attitudes and problems of lay participants. The public defenders in both jurisdictions were particularly interested in their performance relative to private attorneys, defendant attitudes, evenhandedness, delay, and the tentative finding that trials seemed to involve little or no sentence penalty compared with straight pleas. (In one jurisdiction, the public defender's office was also very interested in other issue areas not demonstrated in our work, such as the use of public defender time, the attitudes of judges toward public defenders as opposed to retained counsel, and the utility of support staff.)

Incremental Costs. We stated at the outset that careful estimates of the incremental cost implications of upgrading the capabilities of existing or planned infor-

mation systems or of implementing an improved IPMP from scratch were beyond the scope of this study. However, it may be useful to provide the reader with a breakdown of the resources we used in manual data collection (from agency case files), in computer processing and analysis of these data, in the administration of mail surveys (of victims, witnesses, jurors) and personal interviews (of defendants), and in the computer processing and analysis of these survey responses. Assuming that jurisdictions would be provided with the necessary standardized data collection forms, lay participant questionnaires, and the necessary software packages at no cost, the resources we allocated to the data acquisition, processing, and analysis can be viewed as a rough starting point for estimating the range of incremental costs that might be incurred by interested jurisdictions. We also include very approximate cost estimates for collecting and analyzing data necessary for examining weighted caseloads in the court, the prosecutor's office, and the public defender's office; these estimates are based on telephone conversations with personnel from consulting firms that have implemented weighted caseload measurement systems in such local agencies, since we did not ourselves collect these data in this study.

We consider two bounding cases: (1) jurisdictions with access to a computer, but *no* existing or planned information system, wishing to measure performance in all issue areas covered by an improved IPMP; and (2) jurisdictions with a CCH/OBTS, SJIS, or PROMIS system wishing only to upgrade to Option IV. In both cases, the rough estimates are *annual* costs for measuring performance *once* per year, although the data collection may be intermittent or continuous over the year.[3]

Rough cost estimates are shown in Table 11 for these two cases. For a jurisdiction *without* an existing information system, but with access to a computer, Case I might cost on the order of $50,000 per year to operate a full IPMP once it is set up. (First-year costs should be considerably higher, perhaps by 25 to 50 percent, because of nonrecurring setup costs.) This assumes that the software is made available free, that practitioner time devoted to data collection (e.g., having prosecutors fill out data sheets at each stage of the felony proceeding) is essentially "free,"[4] and that the number and size of data samples collected and analyzed are similar to those we collected in this study. If a jurisdiction wished to draw *additional* case file samples (e.g., for more offense types and/or for oversampling trials), total costs for analyzing the case file data should not increase very much, because most of the additional cost would be in data collection (a small part of the total) and not in data processing, analysis, and interpretation. Costs would also rise if larger samples of defendants were interviewed or if larger samples of victims, witnesses, and jurors were surveyed. The largest fraction of the total operating costs of an IPMP may well be attributed to the analysis of the use of judicial, prosecutorial, and public defender time. However, we have limited confidence in the estimates shown in the table, because we did not have first-hand experience in gathering and analyzing such data in this study.

For a jurisdiction *with* one of the information systems discussed, incremental

[3] The issue of how often performance should be measured cannot be resolved in general terms. It would depend on the resources available to each agency, the issue area or policy issue under consideration, the agency's perceptions about the acuteness of their problems, and perceptions about the public's interests in and attitudes toward the performance of the court system.

[4] Experience with collecting PROMIS and judicial weighted caseload data suggests that a judge or prosecutor might spend only a small fraction of an hour per day (approximately 15 minutes) filling out data sheets.

annual costs of about $10,000 might be incurred if it were upgraded to Option IV. Whichever alternative a jurisdiction chooses, a competent statistician or econometrician, who has acquired a detailed knowledge of criminal court systems, would be required (at least part-time) to perform the statistical analysis and interpretation.

We should emphasize that *all* of the estimated costs displayed in Table 11 are quite uncertain predictors of what jurisdictions would actually incur. Readers should view them only as very gross approximations. Given the probable low marginal costs associated with upgrading an existing or planned information system and its major benefits (an increased understanding of the independent effects of important factors on major performance measures of outcome), it is probably cost-effective for a jurisdiction to pursue this alternative. Whether a full IPMP is cost-effective is a judgment that can only be made by an implementing jurisdiction after such a system is installed and operated over several years and when its actual costs and benefits are assessed.

Table 11

Rough Cost Estimates for Performance Measurement

Case I: Implementing an IPMP from Scratch (assumes that software packages are provided free and that practitioners' time devoted to data collection is "free")	Approximate Range of Man-Months Required	Approximate Range of Direct[a] Dollar Cost
Case file data (three 100-case samples of filed cases, two 100-case samples of screening actions, two 100-case samples of rejection reasons, one 100-case sample of continuances and victim/witness appearances):		
Manual data collection (at $3-$5/hr)	1.5-2.0	1,000-2,000
Data cleaning, keypunching, and computer processing	—	10,000
Statistical analysis and interpretation (at $10/hr)	2.0-3.0	3,000-5,000
Subtotal		14,000-17,000
Weighted caseload use of practitioner time (assumes 3 man-months of a coordinator-analyst required in each agency and that practitione time for data input, coordination, etc., is "free"):		
Court (10-15 judge sample, 4-6 weeks each/year)	3	5,000-10,000
Prosecution (10-15 prosecutor sample; 4-6 weeks each/year)	3	5,000-10,000
Public defender (10-15 defender sample; 4-6 weeks each/year)	3	5,000-10,000
Subtotal		15,000-30,000
Mail surveys of 150-200 samples each of victims, witnesses, and jurors (assum ing questionnaires are free):		
Administration (initial mailing, follow-up mailing, and telephoning)	2.5	5,000
Data cleaning, keypunching, and computer processing	1-2	1,500
Statistical analysis and interpretation	1-2	2,000-3,000
Subtotal		8,500-9,500
Defendant interviews (45 interviews) (assuming questionnaires are free):		
Administration		1,000-1,500
Computer processing		500
Subtotal		1,500-2,000
Report writing (one month per agency)	3	5,000
Grand total		44,000-64,000 (approx.)
Case II: Upgrading Existing or Planned Systems (CCH/OBTS, SJIS, PROMIS) to Option IV	**Incremental[b] Man-Months Required**	**Incremental[b] Direct[a] Costs**
Case file data (same data samples as Case I) (assumes that software packages are free and practitioners' time devoted to data collection is "free"):		
Data collection	Negligible	Negligible
Data cleaning and key punching	Negligible	Negligible
Computer processing	—	2,000
Statistical analysis and interpretation	1-2	2,000-3,000
Report writing (one month per agency)	3	5,000
Grand total		9,000-10,000 (approx.)

[a]Excludes overhead and fringe benefit costs.

[b]Over and above Option I costs of existing or planned systems.

☆ U.S. GOVERNMENT PRINTING OFFICE: 1977— 241-090/57

END