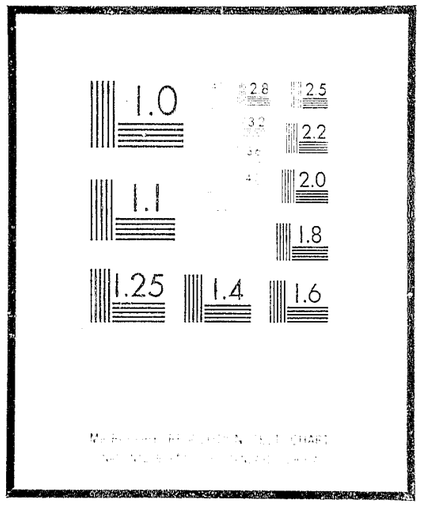


QUANTITATIVE TOOLS FOR CRIMINAL JUSTICE PLANNING

LEONARD OBERLANDER, Editor

UNITED STATES DEPARTMENT OF JUSTICE
LAW ENFORCEMENT ASSISTANCE ADMINISTRATION
Washington, D.C.

1975

RICHARD W. VELDE, Administrator

LAW ENFORCEMENT ASSISTANCE ADMINISTRATION

FOREWORD

A basic premise underlying the programs administered by the Law Enforcement Assistance Administration is that good planning is indispensable to the development and implementation of effective programs for improving criminal justice and reducing crime. The argument for this is that the chances for a rational allocation of scarce resources are greatly increased if a systematic planning process has preceded the allocation decision.

The premise is easy to state, but good criminal justice planning is extraordinarily difficult to put into effect at the State and local levels. The reasons why this is so are well known to those who have been involved in criminal justice planning. They involve factors over which not only planners have no control, but over which State and local governments also have little control. These include the basic structure of government in America. The Federal system, the separation of powers, and the reserved powers of States often appear to planners as barriers to effective planning. While I would agree that these kinds of factors pose special challenges to the ingenuity of planners, it seems to me that the largest barrier to effective planning is the absence of adequate and reliable data and the absence of the capability to analyze those data.

The most powerful tool at the planner's disposal is the data he has collected, organized, displayed, and analyzed in such a way that the relevance of the data to the problem at hand is clearly apparent. It is unfortunately the case that many planners are unable to realize the full potential of the data which exists because they have not developed the skills they need to organize and analyze the data and show its relevance to the critical problems of the moment. This is not to say that there is a failure to recognize the importance of planning. State Planning Agencies, Regional Planning Units, and local planners across the nation are engaged in that process. They have begun to reach some agreement on the elements of that process. They know that they must begin with an analysis of the crime and criminal justice problems they face and that they must thereafter make clear: (1) what resources and capabilities currently exist to meet those problems, (2) what capabilities are currently lacking, (3) what goals they will attempt to accomplish, (4) what standards they will attempt to establish and achieve, (5) what priorities they will set, (6) what long range programs they will pursue, and (7) what specific programs they expect to pursue in the coming year. Planners have done a steadily better job, with help from LEAA, with the latter stages of that process. In some states and in some communities, they have had success with the analysis of crime and criminal justice problems.

It is the case, however, that the earliest steps in the planning process are where the greatest difficulty has been and still is. Data have been hard to come by and expensive to obtain. The consequence is that the data bases have been fragmentary, incomplete, and sometimes highly unreliable. Planners and others have concentrated on the basic task of finding the data, and then of establishing the systems, also with LEAA help, for collecting and organizing those data. Analysis of the data has often had to wait, and has often not occurred except in relatively superficial ways.

We have now reached a stage in the development of criminal justice planning, however, when it is reasonable for us to take stock of where we are, and of what we have learned, and to make available to others what we have learned about analysis. This book begins that process. It places heavy emphasis on ways to organize, utilize, and analyze data and information about crime and criminal justice. It sets forth much of what we know now about the techniques of analysis and is careful to specify their limitations.

Some of the authors in the book draw conclusions about the utility of a particular approach or the lack of utility of another, which all of us do not share. The book is not an attempt to set forth all the answers to all the questions planners have about the ways to analyze crime and criminal justice data. It is rather an attempt to make available to all who have a concern with criminal justice planning what many who are experienced and many who are expert and many who are both have learned about the methods available for analysis. LEAA does not endorse every word in this book, but does believe that it will be a useful compendium of quantitative tools from which planners and others may select those which they find helpful to them.

We have called on the expertise of researchers, analysts, planners, statisticians, and of greatest importance, people who have had direct personal experience with State and local planning processes. They have done an outstanding job of telling us in clear and largely non-technical language what they know and how to use what they know in other settings.

This book will not finish the task LEAA has set itself of helping planners to do a better job. LEAA expects to offer training in analysis, in planning, and in evaluation. It expects to make its guidelines more supportive of good planning for the future. It expects to offer technical assistance in planning, assistance which will be tailored to meet individual needs wherever possible. This book is a beginning, and I commend it to its readers, with that understanding of its purpose.

Richard W. Velde

RICHARD W. VELDE,
ADMINISTRATOR

PREFACE

The literature on crime statistics is expanding rapidly, and interest in the use of statistics and quantitative models for analysis of crime and the criminal justice system is on the rise. A rapid growth in technology is occurring in the criminal justice profession, spurred on by the development and adoption of quantitative methods for analysis in research, operations, planning, and management. Outstanding scientists, law enforcement specialists, practitioners in courts and corrections systems, as well as emerging academicians, are all contributing to it. A major reason for the rapid advances in use of quantitative methods and information systems, along with the expansion of criminal justice statistics, has been the leadership and support for these developments by the Law Enforcement Assistance Administration with the support and appropriations provided by the Congress of the United States.

We are still a long way from any attempt at a definitive work on quantitative tools for application to criminal justice problems and their solutions. However, it is intended that this book will serve as a step in that direction. Since the nineteen-thirties in the United States, an increasing number of disciplines have been concerned with the problem of crime in the nation. Concepts from such fields as urban planning, organization theory, policy sciences, management science, quantitative geography, and public administration have been applied to the emerging criminal justice technology.

This book is addressed to criminal justice planners, practitioners, and students in applied or professional settings with an interest in the use of criminal justice data and quantitative methods for describing, analyzing, and communicating a broad array of issues concerning crime and the criminal justice system. It is intended that it be of particular interest and use to statisticians, operations researchers, systems analysts, and those with a special role in planning and evaluation at state and local levels of government.

The best information from expert sources has been compiled and assembled in the time available. The book should serve some of the functions for which textbooks are intended in more traditional fields. A compendium of materials on quantitative tools in the criminal justice field will not be as esoteric or advanced as a text in the field of economics. Future editions will build upon the first edition by benefiting from readers' comments which will identify the errors and omissions allowed in this volume. The value of this book derives not so much from what can be communicated from these pages to its readers, as from what its readers can contribute to furthering a systematic body of knowledge on practical quantitative tools for the criminal justice profession.

The Law Enforcement Assistance Administration is grateful to the contributors mentioned elsewhere in this book for their generous efforts and cooperation in meeting a demanding publication schedule, as well as to those who have given their time to review and comment on the draft manuscript.

The book was written at the suggestion of many criminal justice planners and practitioners, and was accomplished under the overall direction of

James M. H. Gregg, Assistant Administrator, Office of Planning and Management, and Blair G. Ewing, Director, Division of Planning and Evaluation Standards, Office of Planning and Management. As the editor of the book, I maintain responsibility for any faults resulting from my efforts to edit the chapters into a unified volume with a common format and transitional flow for facilitation of reading, comprehension, and practical useability.

Leonard Oberlander
Office of Planning and Management
Law Enforcement Assistance Administration
Washington, D.C.

August 15, 1975

CONTRIBUTORS

ALFRED BLUMSTEIN	Carnegie-Mellon University, School of Urban and Public Affairs, Urban Systems Institute, Pittsburgh, Pennsylvania 15213. Chapter 13. A Model in Planning for the Total Criminal Justice System.
SIDNEY BROUNSTEIN	Institute for Law and Social Research, 1125 15th Street, N.W., Suite 625, Washington, D.C. 20005. Chapter 10. Analysis of the Criminal Justice System with the Prosecutors Management Information System (PROMIS).
RONALD E. CRELLIN	Program for Data Research, Social Science Research Institute, University of Southern California, Los Angeles, California 90007. Chapter 14. Data Use Tools and Techniques Available to Criminal Justice Professionals.
JOYCE DEROY	Institute for Law and Social Research, 1125 15th Street, N.W., Suite 625, Washington, D.C. 20005. Chapter 12. Data Reliability and Data Purification.
DAVID L. DOUGHERTY	Office of Criminal Justice Plans and Analysis, District of Columbia Government, Washington, D.C. 20000. Chapter 5. UCR in Crime Specific Research and Planning; footnote to Chapter 13. A Model To Aid in Planning for the Total Criminal Justice System.
BLAIR G. EWING	Law Enforcement Assistance Administration, Washington, D.C. 20531. Chapter 1. Quantative Tools for Criminal Justice Planning.
GEORGE L. FARNSWORTH	Program for Data Research, Social Science Research Institute, University of Southern California, Los Angeles, California 90007. Chapter 14. Data Use Tools and Techniques Available to Criminal Justice Professionals.
FEDERAL BUREAU OF INVESTIGATION	U.S. Department of Justice, Washington, D.C. 20535. Chapter 3. Uniform Crime Reporting: Uses for Criminal Justice Planning.
BRIAN FORST	Institute for Law and Social Research, 1125 15th Street, N.W., Washington, D.C. 20005. Chapter 11. Statistical Techniques and their Limitations in the Analysis of Criminal Justice Data.
THOMAS GIACINTI	Denver Anti-Crime Council, 1313 Tremont Place, Suite 5, Denver, Colorado 80204. Chapter 6. Predicting Crime Incidence and Determining Change.

WILLIAM HAMILTON	Institute for Law and Social Research, 1125 15th Street, N.W., Washington, D.C. 20005. Chapter 10. Analysis of the Criminal Justice System with the Prosecutors Management Information System (PROMIS).
SUSAN KATZENELSON	Institute for Law and Social Research, 1125 15th Street, N.W., Washington, D.C. 20005. Chapter 8. Analysis of Crime with Offender Based Transaction Statistics (OBTS); Chapter 9. Analysis of the Criminal Justice System with Offender Based Transaction Statistics (OBTS).
FRANK LOCKFELD	County of Santa Clara, Office of the County Executive, Center for Urban Analysis, 19 County Administration Building, 70 West Hedding Street, San Jose, California 95110. Chapter 16. Tools for Forecasting.
BURT NANUS	Center for Futures Research, Graduate School of Business Administration, University of Southern California, Los Angeles, California 90007. Chapter 7. A Planning-Oriented Measure of Crime and Delinquency.
LEONARD OBERLANDER	Law Enforcement Assistance Administration, Washington, D.C. 20531. Chapter 1. Quantative Tools for Criminal Justice Planning.
R. ALLEN PAYNE	Law Enforcement Assistance Administration, Washington, D.C. 20531. Chapter 18. Crime-Oriented Planning and Program Development.
LUTHER PERRY	Center for Futures Research, Graduate School of Business Administration, University of Southern California, Los Angeles, California 90007. Chapter 7. A Planning-Oriented Measure of Crime and Delinquency.
JOHN R. PFEIFFER	Director, Statistical Analysis Center, Massachusetts Committee on Criminal Justice, Boston, Massachusetts 02100. Chapter 4. Crime Statistics Analysis in Massachusetts.
LEO SCHUERMAN	Program for Data Research, Social Science Research Institute, University of Southern California, Los Angeles, California 90007. Chapter 14. Data Use Tools and Techniques Available to Criminal Justice Professionals; Chapter 15. Small Area Population Estimation.
WESLEY G. SKOGAN	Assistant Professor of Political Science and Urban Affairs, Northwestern University, Evanston, Illinois 60201. Chapter 2. The Use of Victimization Surveys in Criminal Justice Planning.
CHARLES D. WELLER	Executive Director, Denver Anti-Crime Council, 1313 Tremont Place, Suite 5, Denver, Colorado 80204. Chapter 17. Crime Analysis: Organization of a Team To Impact the Local Criminal Justice System.

CONTENTS

Part I.

INTRODUCTION

CHAPTER 1

QUANTITATIVE TOOLS FOR CRIMINAL JUSTICE PLANNING

Leonard Oberlander and Blair G. Ewing

INTRODUCTION

Criminal justice planners need tools that permit decisions to be made on a basis which is more sophisticated than those founded solely on experience, judgement, and intuition. Without quantitative information, defining the nature and extent of crime and making decisions about the application of resources to the problem must almost entirely be based on intuition. There is no underestimating the value of quantitative data to those who operate agencies and make decisions.

QUANTITATIVE TOOLS: A DEFINITION

Quantitative tools are the devices used to construct and to analyze specific groupings of information in terms of quantity or numbers. The tools that criminal justice practitioners have at their disposal include quantitative data about crime, the criminal justice system, behavior, and the environment, as well as a wide selection of methods for organizing and analyzing these data. These tools help provide a quantitative base of knowledge upon which criminal justice practitioners can bring to bear their own expertise for planning, evaluation, program development, and policy analysis.

A partial list of the kinds of quantitative tools that might be utilized includes:

Data on Crime

1. Victimization survey data
2. Uniform Crime Reports
3. Modus Operandi files
4. Offender-Based Transaction Statistics
5. Other local crime data

Data on the Criminal Justice System

6. Offender-Based Transaction Statistics
7. Management and administrative data

Data on Behavior

8. School truancy data
9. Mental health data
10. Social services data
11. Alcoholism and drug abuse data
12. Economic trends, including unemployment and those persons outside the labor force

Data on Environment

13. Census data on population and population characteristics
14. Environmental data concerning commercial and residential areas, types of building structures, and land topology
15. Dual Independent Maps Encoding (DIME) files, and other geographic based files.

Methodological Tools

16. Techniques for statistical analysis
17. Justice System Interactive Model (JUSSIM)
18. Prosecutor's Management Information System (PROMIS)
19. Geocoding

Although there are other data and methodological tools available upon which the criminal justice practitioner may draw, this list provides a general perspective of the wide variety of tools that do exist and are accessable to those charged with analytic and planning tasks in state, regional, and local units of government and other institutions concerned with criminal justice.

THE DATA BASE FOR CRIMINAL JUSTICE PLANNING

Rationale for the Data Base

Planning for the management of the criminal justice system requires a base of knowledge from which informed decisions can be made. The availability of data on how many reported offenses there are and how many arrests have been made, and on the handling and disposition of arrests permits analysis to be made of the way in which the criminal justice system as a whole performs; how it deals with its workload; where the major areas of discretion are; where the bottlenecks exist; and which agencies appear to be effective or ineffective in handling cases, persons, and special problems.

Analysis of data pertaining to the relationship between crime and its social and environmental context leads to an increased understanding of the kinds of settings in which crime is more or less likely to occur. That understanding is a source of hypotheses about what kinds of programs or projects might best be tested under certain conditions and in certain areas. This kind of analysis is best done by relating specific types of crimes to demographic and physical environmental variables, and usually employs information such as land-use maps, master plans, and census data. Furthermore, such analysis can increase understanding of the ecology of crime; the mobility of offenders; the circumstances of victimization; and the role of age, race, sex, and the socioeconomic status of offenders. Although this type of analysis does not lead to definitive statements about the causes of crime, it does provide the basis for an objective understanding of the circumstances and characteristics of criminal events and offenders.

Operational planning focuses primarily on the allocation of resources and manpower to deal with specific workloads, including workload concentrations during peak hours or at high activity locations. This kind of planning rests on the availability of data about when and where different types

of crime occur. Such data is routinely collected by most police departments in conjunction with their regular reporting of criminal events. The data that is collected and analyzed, in some cases in very sophisticated, automated or computerized systems, can also be used to develop justification for increases, decreases, or shifts in manpower resources. Analysis of data can also be of considerable value for forecasting, including forecasting for operational purposes and overall forecasting for crime trends, population shifts, and other events such as economic trends and unemployment rates that may relate to crime rate forecasts.

Kinds of Collections in the Data Base

Much of the data for criminal justice planning is available as a product of ongoing standardized data collections. This data consists of uniformly collected statistics that concern specified populations for one specific point in time or for certain time periods. The nature of standardized data allows comparisons to be made in a systematic way between and within populations, based on comparable data. The ongoing nature of standardized data collections allows comparisons and trends to be computed for sequential time periods.

Other data and information that may be acquired from numerous sources are non-standardized data that have been collected for singular purposes. These data collections usually concern an interest in a specific segment of a population, a specifically defined subject area of research, and usually are unique to one defined time period. It is difficult to make comparisons of statistics in this kind of data collection with statistics for other populations or points in time because of the lack of a standard of comparability. This data is useful, however, for planning purposes because it provides insights into problems of crime and the criminal justice system when systematically aggregated with other data to highlight areas of consistancy and differences resulting from the examination of statistics for various populations and time periods.

How the Data Can Be Acquired

Data sources for the criminal justice planners' data base are varied. Some data collections are readily available from Federal, state, and local agencies or from private and public organizations such as information clearing houses, university data centers, professional association libraries, and research centers.

The following are sources from which the planner can most efficiently acquire some of the data referred to in this book.

1. Victimization Survey Data: This data, collected for the Law Enforcement Assistance Administration (LEAA) by the U.S. Bureau of the Census, is available in the form of published tabular data in a series of reports. These reports are available from the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402. Information concerning the nature and contents of these and related reports can be acquired by communicating with the National Criminal Justice Reference Service, Law Enforcement Assistance Administration, Washington, D.C. 20531. In addition,

public-use computer tapes (concerning victimization data) with accompanying documentation and instructional materials will become available in early 1976. DUALABS, 1601 North Kent Street, Arlington, Virginia 22209 can be contacted for information regarding the acquisition and use of these tapes. See Chapter 2 for a discussion of victimization data.

2. Uniform Crime Reports: Chapter 3 of this book, "Uniform Crime Reporting Uses for Criminal Justice" by the Federal Bureau of Investigation, discusses the means for acquiring and utilizing data generated by the Uniform Crime Reporting Program.

3. Modus Operandi (MO) Data: (detailed in Chapter 7 of this book) suggest that MO files can be an important resource to the planner. MO files are prepared according to a standard policy guide by trained officials of police departments. The use of police MO files, which contain no names, addresses, or personal identifiers of any kind, are in common use in many police agencies.

4. Offender-Based Transaction Statistics (OBTS): This data, where it is available is usually located in the State Statistical Analysis Center, a module of the LEAA Comprehensive Data Systems program. See Chapters 8 and 9 for discussions about OBTS.

5. Management and Administrative Data: Such data includes manpower, cost, and program or activity data among other kinds of information. The most important elements of this type of data are those which permit analysis of the capabilities of agencies to meet needs and achieve objectives and goals. The manpower, cost, and program description data, with manpower and costs allocated by program if possible, are most useful in this regard. These data can be obtained from most of the kinds of local agencies or departments listed below.

- Departments of finance
- Youth authorities
- Departments of corrections
- Personnel agencies
- Judicial agencies
- Police officers' training units
- Public employers' retirement systems
- Emergency services departments
- Comptrollers of budget offices
- Departments of rehabilitation

There are many agencies like these at local levels of government from which a considerable amount of discrete data may be used. The data may already be published, or available in easily accessible records.

It may be necessary to obtain expenditure and personnel data directly from local units by questionnaires; but such data is usually compiled routinely in the course of the normal budget process, and contacts with local jurisdictions can often indicate that it can be supplied without difficulty.

THE UTILITY OF THE DATA ON CRIME

When the data base has been accumulated, the planner can then execute his plan for analysis and utilization of the results to identify the nature and extent of the crime problem, and to examine and compare such components of the criminal justice system as workload productivity and workload objectives. The following discussion summarizes some of the uses of the data base for the criminal justice planner.

Victimization Data

Some of the uses of victimization survey data are discussed by Skogan in "The Use of Victimization Surveys in Criminal Justice Planning." Briefly, victimization surveys can be used to estimate crime rates for a jurisdiction. They can be utilized to examine local patterns of reporting and non-reporting, which can help planners to understand differences between increases in the crime rate and increases in the reporting rate. Victim survey data can also contribute to the study of the details of events and to isolate deterrable crimes. This data can also be analyzed to identify high-risk subgroups in the population. By obtaining "feedback" about citizens' views of the operation of the criminal justice system, the data can be used to measure citizen satisfaction with the performance of the criminal justice system. Victimization surveys can also be used to establish benchmarks for evaluating the effectiveness of criminal justice programs. The data can also be used as an aid to establish a pool of crime victims who can be "tracked" through the criminal justice system.

Uniform Crime Reporting (UCR) Data

Uniform Crime Reporting data are a valuable tool to the planner. The Federal Bureau of Investigation, in "Uniform Crime Reporting Uses for Criminal Justice Planning" (Chapter 3), characterizes the UCR Program, the sets of UCR data that can be utilized by criminal justice practitioners, and the uses of UCR data. UCR data can be used to obtain crime rates and trends for the many sets of UCR data. UCR data can also be used in conjunction with other sets of data, such as school dropout rates; unemployment rates; economic levels; and gross population characteristics relative to age, sex, and race to examinate rates and trends that correlate arithmetically with the crime data.

Some of the uses of UCR data in conjunction with other data relating to sociological phenomena include analysis of offenses known. By concurrently viewing similarly constructed data from other data sources with this data set, significant patterns may be discerned. Clearance information is to some degree a measure of law enforcement efficiency. Careful analysis of prosecution and court information, overlayed with clearance data, can provide useful information about workload for future planning. Data on youthful offenders can be utilized by correlating it with population composition, detention facilities, population configuration, parole, probation, and welfare caseload composition, and other data sets to which youth or juvenility are intrinsic. Moreover, by utilizing the youthful offender data in time-series, certain observations can be made concerning the time lapse between the encounter of the youthful offender by law enforcement and the processing of the individual at later stages of the criminal justice system.

Additional observations or projections may be made concerning workloads, detention facilities, and alternative programs.

By utilizing arrest information, the planner can analyze fluctuations of law enforcement activity. Such data can be of aid to assessing the effectiveness of programs initiated not only within the law enforcement module but also within the other modules of the system. Disposition information can be helpful to the planner who desires to test offense information, aligning it with clearance data and, in the case of Crime Index offenses, to observe in summary number form the actions of the joint effort of courts and prosecution. In addition, law enforcement agency profiles can be constructed by utilizing UCR data and law enforcement employee data. These profiles can serve administrative needs.

Pfeiffer, in "Crime Statistics Analysis in Massachusetts" (Chapter 4), discusses how the criminal justice planning agency in Massachusetts has employed UCR data to analyze crime patterns and trends for use in planning and program development. He describes an approach used by the Massachusetts State Planning Agency to develop the structure and content of a crime report.

Doughtery, in "UCR in Crime Specific Research and Planning" (Chapter 5), discusses how UCR has been helpful in examining homicide, an area in which he has relied heavily on this data for analysis.

Modus Operandi (MO) Files

Modus Operandi data may be helpful as a starting point for generating a seriousness index of crime that can be used to evaluate existing policy, monitor activities, and help guide planning decisions in such areas as manpower authorizations and budget allocations. Nanus and Perry, in "A Planning-Oriented Measure of Crime and Delinquency" (Chapter 7), discuss the uses of MO files.

Offender-Based Transaction Statistics

The gathering of the required data for measurement essentially involves identifying and collecting data elements as set forth in the Offender-Based Transaction Statistics System (OBTS) module of the LEAA Comprehensive Data Systems Program (Chapter 8 by Katzenelson). By using the concept of the Comprehensive Data System, one could examine detailed recidivism indicators separately by crime category, by demographic characteristics, and by type of institution.

There are many alternative sources of data for criminal justice planning and crime analysis. Each is best suited to given types of applications; each has advantages. All provide the quantitative data that is necessary to supplement subjective experience and individual judgement. Which type of data to use, and when to use combinations of data sets supplemented by additional information as well, depend on the purpose of the analysis to be accomplished.

THE UTILITY OF DATA ON THE CRIMINAL JUSTICE SYSTEM

There are analytic tools that can be used in a variety of ways to develop a better understanding of justice system operations and to improve its

management. Some of these data tools and techniques are discussed in detail in Part III of this book. In addition to approaches that are discussed, such as OBTS, PROMIS, and JUSSIM, a data base of management and administrative statistics (including personnel and expenditure information) is necessary for an adequate analysis of criminal justice system activities.

OBTS

OBTS can be applied to assess the workload, productivity, and objectives of the criminal justice system. It can be used to examine a justice system or to compare systems. Questions that OBTS can help answer include: What is the relative frequency of guilty pleas, jury trials, and bench trials? What is the direction of change in most serious charges from arrest to disposition? What is the quantitative nature of the flow of offenders through each component of the system?

The Prosecutors Management Information System

The Prosecutors Management Information System (PROMIS) permits a prosecutor's office to accumulate information on each case and to receive reports and analyses based on such data so that prosecutors can concentrate on priority areas and control their own workloads. PROMIS can also be used to provide data for planning and research purposes. Chapter 10, by Brounstein and Hamilton, "Analysis of the Criminal Justice System with Prosecutors Management Information System," describes how PROMIS can be used, the data methods to be employed, and the reports than can be developed with PROMIS.

The Justice System Interactive Model

The Justice System Interactive Model (JUSSIM) is a model to aid in planning for the total criminal justice system. Chapter 13, by Blumstein, describes the planning process in each of its steps as it relates to uses of the JUSSIM model as a tool for analysis and future projections. In addition, the many uses of JUSSIM are detailed.

Management and Administrative Data

The data base for the criminal justice planner should include management and administrative data.* Such data sets may contain information on agency expenditures, workload, personnel, and other resources in a format that will expedite analysis relating these factors to available demographic, fiscal, and other criminal justice system data. One of the purposes for expanding a data base to include administrative data is to meet the needs of a different type of user than the enforcement agencies who are thought to be the primary users of the majority of criminal justice data presently.

---

*We are grateful to the California Bureau of Criminal Statistics for its reports which provided a valuable input to this discussion of management and administrative data.

One group of users who seem particularly important are the local administrative officers, department heads, and supervisors who prepare the criminal justice agency budgets. The insight this small group of officials may secure from having certain data available that helps relate changes in cost, workload, and personnel in a meaningful way could have considerable effect on the local criminal justice system.

There is more involved in making criminal justice policy than just budgeting expenditures. Yet, the decisions made and the information considered in the course of developing a department budget may do more to influence local action on criminal justice agencies than any other local governmental process.

Small groups of officials who are making budget decisions are often working with very little organized information about their criminal justice agencies of the type discussed in this chapter. The cost of building and maintaining large, centralized data information systems is considerable. Therefore, statistical output of an analytic unit should be planned to serve not only the traditional needs of the contributing agencies, but fiscal, administrative, and elective officials as well.

The utility of management and administrative data may be thought of in the context of a number of considerations. These include data that would be useful within the context of the local budgeting process; identification of the sources of data needed outside of the analytic unit, and their availability; an assessment of the kind of assistance that could be needed and received from local units in preparing the reports; suggested arrangement of report formats; suggested organization of research activities of the analytic unit staff; and scheduling of the reports.

As important as comparative workload and staffing information are to the budget process, it is very difficult for local administrators and their staffs to carry out the amount of research necessary to conduct comparative studies for a significant number of departments and keep the information current. In approaching specific problems within their individual jurisdictions, local administrators have developed many uniform and consistent standards for measuring change and rates of growth in a department's basic functions. Many of the standards are applicable to all jurisdictions or can be extremely useful in designing uniform standards. The work that has been done in this area for criminal justice agencies on a scattered, individual basis can greatly simplify the task involved in an analytic unit developing a data base of administrative and management data.

One approach is to organize selected data elements into a number of tables. The data elements will differ for different jurisdictions and agencies, and the sample procedure described below is simply intended for a general overview.

The data elements selected may be organized into four tables:

- Selected Demographic, Economic, and Criminal Justice Characteristics of Geographic Areas
- Annual Expenditure and Workload Data

- Authorized Personnel by Functional Classification
- Expenditure, Personnel and Workload Ratios

A brief discussion for each data element should explain the purpose of the data element, its source, and some of the problems encountered in collecting the data.

The main purpose for the data in each table is that it can help users qualify the data in each of the other three tables. Some of the differences that emerge from that comparison with the other tables can be better understood, when something is known about the geography, population, and wealth of an area. Differences in the amount of taxable wealth, for example, are often the critical factor in explaining why one area may appear to provide a higher level of law enforcement than another; when by geography or amount of crime, they look similar.

For every data element used in the tables, time series information can be shown, along with the absolute change and percent change. In time, and as the data base builds, the report formats could be easily expanded to show changes over longer time periods.

Although many of the services and functions performed by criminal justice agencies have similarities, it is possible to develop and maintain a reliable base of information about these agencies that could add an important dimension of reporting and analytical capability to a planner's data base. Some hints that might be beneficial in developing and using a data base of administrative and management-oriented data follow.

- It may be useful to put primary emphasis on developing analytical interpretations of the data presented rather than expending effort to present increasing amounts of detailed information. Budgeting "measures" in many jurisdictions usually tend to be very general rather than specific.

- Data is usually most effectively presented for individual jurisdictions on a department-by-department or agency-by-agency basis, although if the data can be presented also by program category, it would be even more effective.

- When reports are updated, it is usually best to avoid changes in the report formats, keeping the data elements constant, and maintaining consistency in the analysis.

- Reports should be timed to the starting date of the jurisdictional budget cycle.

- In publications about crime and criminal justice, data should be treated without implication that workloads or personnel ratios are being suggested as recommended standards. Standards are a matter for policymakers to determine, perhaps based in part on local data.

THE UTILITY OF DATA ON THE POPULATION, GEOGRAPHY, AND BEHAVIOR

The linking of crime data to data routinely gathered by many other public agencies can enable an analysis of crime statistics with socioeconomic, demographic, behavioral, and land-use information. Analysis of this kind can help decision makers assign weights of importance to different factors. An analysis of crime might indicate several key factors present, such as age, sex, and race of victims and offenders; population density; unemployment; locations and types of premises; school truancy; and drug abuse. The diagnosis of crime in one area will point to some combination of these and other factors. Analyzing crime data with other relevant data sets can provide information on which factors are important to a particular area. This information is important for the formulation of objectives and goals, and to program development. The planner then can utilize manpower, cost, and program or activity data to assess the capabilities of agencies to deal with defined problems, and to make recommendations for allocations of resources to address specific needs.

Part IV of this book discusses data on population and tools for analysis. Population data provides an invaluable frame of reference for examining crime data and crime patterns, and for identifying probable causes and the factors linked with these causes.

In Chapter 14, Crellin, Farnsworth, and Schuerman emphasize the importance of geographic data in the understanding and analysis of crime. They point out the value of census data for crime analysis in geographic and geopolitical areas of varying size. The Dual Independent Maps Encoding File (DIME) represents a technology that uses computers to create coded geographic maps. Computer programs that sort, structure, link, and display map feature data from a variety of sources are described. A list of these computer programs and their sources is provided, along with a list of reports on the uses of small area geographic files. The authors include a helpful glossary of geographic terms and a list of Standard Metropolitan Statistical Areas (SMSA's) with Geographic Based DIME Files.

In Chapter 15, Schuerman discusses small area population estimation. The chapter was written for a person with a background in advanced statistical techniques, but for those without it, the chapter remains a good orientation to tools and techniques. The Bouge-Duncan Composite Method is used for creating separate estimates for subgroups and for summing these to arrive at totals. The Ratio Correlation Method uses multiple regression analysis for estimating purposes. A variety of other techniques is included, with cautions about their use.

Computer mapping, geocoding, and forecasting with geocoded data are described in Chapter 16 by Lockfeld. The uses of time series analysis and spectral analysis are described for looking at recurring events with characteristics that vary in a pattern, e.g., weekly and seasonal variations in crime frequencies.

Part II.

CRIME: DATA AND TECHNIQUES FOR ANALYSIS

CHAPTER 2

THE USE OF VICTIMIZATION SURVEYS IN CRIMINAL JUSTICE PLANNING

Wesley G. Skogan

INTRODUCTION

Victimization surveys are a new research tool for criminal justice planners. In a victim survey, interviews are conducted with samples of the population of a jurisdiction. These interviews probe the respondents' experiences with crime, their reaction to these incidents, and their evaluations of the performance of criminal justice agencies. If the samples are carefully drawn, these reports can be extrapolated into population estimates of the frequency of crime and the perceived responses of community institutions to the needs of the populace. These estimates may then be used to assist planners in allocating resources to various criminal justice activities.

The first large-scale surveys of the population designed to measure the incidence of crime were conducted by the Crime Commission in the mid-1960's. A national survey conducted by Philip Ennis and a study of high-crime areas by Albert Reiss established that victimizations were much more common than police statistics indicated and that the burden of crime was unevenly distributed in the population. Since the Commission surveys, a number of studies of victimization have been conducted (some of which are listed in the Bibliography following this chapter). In 1972, the Bureau of the Census began a national crime investigation for the Law Enforcement Assistance Administration. This continuing survey will produce a new set of social indicators monitoring quarterly fluctuations in the volume and nature of crime in America.

LOCAL USES OF VICTIMIZATION SURVEY DATA

At the local level, there are at least seven ways in which population surveys can be employed to assist criminal justice planners:

1. Victim surveys can be used to estimate crime rates for a jurisdiction. This is extremely important, for police statistics usually underrepresent the total volume of crime in an area. Citizens do not report many crimes to the police, and not all reported crimes are accurately reflected in official reports. Victimization surveys can be used to generate crime data which are independent of criminal justice agencies and which circumvent many of the non-reporting problems which plague official totals. These rates may be calculated in several ways. In addition to the usual "crimes per 100,000," surveys enable us to examine household burglaries per 100,000 households, bicycle thefts per 100,000 bicycles, or payroll thefts per 100,000 payrolls. Survey data thus can be used to generate more useful estimates of the "risk" of various kinds of crimes, since these data can be used to estimate the number of targets for specific crimes as well as the number of victims.

2. Surveys can be used to examine local patterns of non-reporting. Many crimes are not reported to the police; nationally, about 40 percent of all personal crimes and 60 percent of all property offenses are not brought to the attention of the authorities. One role of victim surveys is to provide information about these offenses so that the resources of the criminal justice system can be distributed in response to the "true" distribution of crime. Another is to increase our understanding of why so many offenses go unreported, and to devise programs to increase citizen cooperation with the police. Only crimes which are reported can be deterred by official action, so efforts to increase citizen reporting plays an important role in crime control. Finally, surveys can be used to estimate the impact of changes in reporting rates upon official statistics. The first impact of new anti-crime programs may be on reporting habits, and if citizens begin to report more crimes it will paradoxically appear that the victimization rate is skyrocketing. A careful study of patterns of reporting and nonreporting will enable planners to tell the difference between increases in the crime rate and increases in the reporting rate.

3. Surveys may be used to study the details of events and isolate deterrable crimes. Victim surveys can be used to elicit many details about criminal events from their victims. These details can be significant in several ways. First, they provide estimates of the impact of crime on community institutions. The victims of crime are heavy consumers of the services of hospital emergency rooms. Moreover, the poor are often forced to turn to local welfare agencies for assistance in overcoming the burdens of crimes such as property theft or robbery. Information on these costs of crime are often hidden, but surveys enable planners to estimate more accurately the true cost of crime to the community.

In addition, interviews with crime victims may reveal information about the structure of criminal events which will assist the police in allocating scarce resources. For example, most interpersonal violence between friends, neighbors, and relatives is probably undeterrable by the police. It usually takes place indoors, in private space protected from the unwanted intrusion of the government, and it is spawned by bitterness and conflicts which often are not subject to rational calculation. The police may perform valuable crisis-intervention functions, separating the contending parties and negotiating a peace, but they will not deter the incident by intensive neighborhood patrolling. The police can do something about street crime (robbery, purse-snatching, and most assaults by strangers), and they may deter many auto thefts, commercial robberies, and household burglaries by intensive random patrolling. Victim surveys can be used to measure the "deterrability" of crimes based on these characteristics, and resources may then be allocated in response to the probability of their effectivensss.

4. Victim studies can be used to identify high-risk subgroups in the population. Crime is not randomly distributed in the population; some groups of persons or commercial establishments are much more likely than others to be victimized and to be victimized on repeated occasions. Other characteristics which affect the relationship between various groups and the criminal justice system are: trust in the police, confidence that their complaints will be taken seriously, and insurance claims which need to be documented. Surveys can be employed to isolate the attributes of high-risk subgroups, investigate the nature of their victimization experiences, and identify characteristic difficulties in their relationship with criminal justice agencies. The allocation of resources in the direction of high-risk groups may be highly cost effective. A very large number of potential offenders may be severely inconvenienced by programs aimed at hardening a few such high-risk targets as convenience grocery stores. Survey data can reveal the hours of operation of such establishments, counter-measures which they have taken against victimization, and the characteristics of high-risk operations that lead to their apparent vulnerability. While most businessmen can be counted on to cooperate with the police, the same cannot be said for many high-risk individuals. For example, young males under 25 years of age suffer over one-half of all assaults, but they report only about one-third of them to the police, and attitude studies conducted by the Bureau of the Census reveal that they are usually dissatisfied with police performance. Knowing who the targets of crime are is only part of the story; victim surveys would enable planners to estimate the problems involved in supplying them with various kinds of services as well.

5. Surveys may be used to measure citizen satisfaction with the performance of the criminal justice system, and elicit suggestions of needed reforms. Victim surveys not only count criminal incidents; they are also a vehicle for contacting large numbers of citizens in a jurisdiction and obtaining "feedback" about the operation of the criminal justice system. This feedback may come in several forms. Citizens may be asked to report their experiences with criminal justice agencies: How fast did the police come when called? How long did they wait to have their traffic case heard in court? Were they notified of the outcome of a criminal case in which they were involved as a witness or a victim? This kind of information can be valuable in evaluating the effectiveness of criminal justice service-delivery systems. Several communities, in fact, have inaugurated regular programs for interviewing crime victims in order to monitor system effectiveness. Each month, samples of those who call for police service are questioned about the quality of their experiences (see the articles in the Bibliography by Furstenberg and Bordua). More general evaluations of services can also be obtained through interviews. Survey respondents can be asked to rate the courteousness of the police, the orderliness of courtrooms, the flow of traffic, and their safety on the street and in their homes (see the article by Skogan in Crime and Criminal Justice Policy). People can also be asked to rank order the priority which local government should give to various criminal justice problems, or to volunteer specific suggestions about what is wrong (and

right) about the community and what could be done about it. Unlike many policy-making areas, most criminal justice decision-making is relatively removed from day-to-day politics and most decisions have very low visibility. Surveys would enable ordinary citizens to make inputs into the planning process, and might engender considerable symbolic satisfaction among respondents as well.

6. Victimization surveys will establish benchmarks for evaluating the effectiveness of criminal justice programs. One of the major uses of victimization surveys is program evaluation. First, they may be conducted independently of the implementation of a program. The measurement process need not be tied to the success or failure of a project or an organization, and the findings may be more trustworthy as a result. Also, many reforms are designed to affect what is being measured--crime--and our ability to measure it--crime reporting and recording. Victim surveys are more useful than official statistics in evaluating the effectiveness of those programs expected to increase the rate at which citizens report crimes or to increase the effectiveness of the police in recording and analyzing the data. Surveys also enable us to measure the non-crime-related consequences of new programs. Many of the benefits of effective crime control are attitudinal and perceptual: the presence of good police on the street in large numbers may serve to enhance people's feeling of security, increase their use of community facilities, and decrease their desire to move out of town. The symbolic consequences of official activity are an important component of governance, and the effects of criminal justice programs on people's use of the streets and public spaces in the community, their perceptions of danger, and their confidence in government, are crucial factors in the evaluation of those programs. The research designs which are devised to evaluate policies must satisfy several technical and scientific standards, and this component of an evaluation program should be considered by experts (see, for example, the discussion by Campbell and Stanley). The willingness to do evaluations and to structure citizen input into the evaluation design is a more fundamental planning issue.

7. Victim surveys will isolate a pool of crime victims who can be "tracked" through the criminal justice system. Data of this sort would provide the base for an extremely sophisticated analysis of the operation of the criminal justice system and an evaluation of the effectiveness of its components. The key point is that a criminal justice system is a system--it is a series of agencies and actors who must work in coordinated fashion if they are to be effective. Perhaps the primary failing of the criminal justice system in America is that it is not such a system. Interviews with victims (and witnesses) provide one mechanism for evaluating the operation of the system as well as its constituent parts. The design would involve the periodic reinterviewing of participants who were uncovered during a community-wide survey. At each step, citizens' perceptions of what was happening to them and their evaluations of that experience could be recorded. The data would also enable planners to independently gauge the rate at which individual problems move through the system, to identify where blockages develop, and to describe what kinds of cases appear to get "lost." New data on everything from

response time to the number of continuances which forced victims and witnesses into court reappearances could be gathered. Because of its survey base, the data would be based upon the experiences of the consumers of justice activity, not upon an eclectic collection of units like incidents (the police), indictments (the courts), and intakes (the jails). Virtually nowhere is officially collected data suitable for evaluating (or even simply describing) the criminal justice system qua system; the victim survey makes an "end run" around these institutional constraints upon the data at the same time that it links them to consumer satisfaction.

WHAT TO LOOK FOR: VICTIMS AND THEIR CRIMES

Victimization surveys can be used to measure the incidence of many criminal events. Some infractions of the criminal law are not suited to survey measurement, however, and many things about which people want "something done" are not violations of the law. The problem of "what crimes to look for" is also related to sampling: some kinds of crime are quite rare, and it would be difficult for a survey to uncover enough of them to analyze; other kinds of crime are quite common, but their targets are not easy to define for sampling. In practice, victim surveying involves many compromises between our desire to estimate the incidence of "every" crime and our ability to do so.

Most victim surveys focus upon a relatively small subset of criminal behaviors. At the top of a scale showing seriousness of crimes, homicide is usually excluded; aside from the obvious fact that it leaves no victim to be interviewed, it is also quite rare and already well-reflected in official statistics. At the bottom of the scale, surveys usually exclude such incidents as obscene telephone calls and vandalism; this reasoning is questionable, however, since certain subgroups in the population (women, the elderly) in fact often feel quite threatened by such incidents. It is important that the survey focus upon crimes which have individual victims--interviews are not suited to measuring the incidence of drunken driving, disorderly conduct, vagrancy, or any of the "crimes without victims." Crimes such as these victimize (if anyone) the collectivity, not specific individuals. Other kinds of crime, such as carrying stolen property or arson, are usually uncovered only through police investigations, and "crimes known to the police" define the universe of known events. Finally, it is possible to use personal interviews to measure the incidence of victimizations only when the victims are aware of their status. Doubtless most consumer fraud, price-fixing, and many unsuccessful attempts at property theft escape the notice of their victims.

Asking individual respondents to define the problem is often of little help. If you ask "What should be stopped?" most interviews will elicit a list of social irritants--unwashed hippies, radical priests, noisy youngsters, demonstrators--which generally do not fall within the purvue of criminal law. Methodological investigations of the victim-interview technique also suggest that other types of events are defined outside the "criminal" sphere even though they closely resemble illegal activity: disputes over the ownership of automobiles, property conflicts between ex-spouses, and many assaults within the family or between friends or relatives, fall into this category. In general, the following are (in my experience) incidents which are relatively easy to measure through a survey:

- any incident involving a gun or knife
- any personal attack resulting in serious injury
- any assault or rape by a stranger
- successful as opposed to attempted crimes
- property thefts involving entry into the home
- property thefts involving large losses (over about $50, although this varies by the wealth of the victim)
- successful auto theft
- events which are unambiguously "crimes" (which clouds the interpretation of many lost wallets: was it pick-pocketing, or was it lost?)

SAMPLING REQUIREMENTS

In order to uncover the victims of these crimes, it is necessary to sample the potential targets of crime. The universe of potential targets and procedures for selecting samples of them for interviewing is not always clear, however. Households can be sampled using simple area-probability procedures, for they are relatively permanent physical units. Individuals are usually contacted through household selection. In the surveys being conducted for LEAA, interviews are completed with each member of a sampled household, while other studies randomly select respondents within the group. Commercial enterprises of many sorts can also be sampled geographically, although isolated shops or service establishments (such as hair dressers working out of private homes) may be undersampled. Large

office buildings present tractable sampling problems, since individual rooms can be sampled within them. Rare and hard-to-sample commercial operators like taxi-cab drivers have serious crime problems, and contribute disproportionately to crimes in such serious categories as armed robbery. To track them down requires imaginative sampling designs. Organizations also present a challenge. The victims of crime may include Boy Scout troops, lodges, and other entities which are social rather than physical units and which are tied neither to place nor person. Like taxi drivers, some organizations will require sampling designs of some sophistication and imagination.

Sampling to find events can be costly. Many of the incidents in which we are most interested (rape, armed robbery) are relatively rare. When the Crime Commission first considered funding a victimization survey in 1966, this was a major concern. Official statistics suggested that the robbery rate, for example, was only about 120 per 100,000 persons in the population. Given this "low density" of events, very large samples would be required in order to turn up enough incidents to talk about. Fortunately (but unfortunately), those encouraging the surveys were correct in their assumption that the "dark figure" of unreported crime was quite large, and that more workable samples would suffice. The victimization survey conducted in Cincinnati in 1974 indicated victimization rates of the following magnitudes:

Crime	Rate
household burglary	14 per 100 households
household larceny	31 per 100 households
personal robbery	1.5 per 100 residents twelve and older
personal assault	5 per 100 residents twelve and over
rapes	0.2 per 100 residents

By this standard, in order to accumulate a random sample of about 140 burglaries (still a small number for many analyses), in many communities it would be necessary to quiz informants in 1,000 households or more. Simple thefts are more common, but the development of a decent sample of assault victims would require screening the experiences of several thousand individuals.

Surveys to investigate the attitudes or perceptions of the population would require much more modest samples. Unlike the relatively uncommon victimization experiences, opinions are part of the personal baggage we carry around with us, and almost everyone has them. A sample of approximately 1,200 respondents will reasonably represent the attitudes of a large community, and the group will be large enough to support many detailed analyses. It may be necessary to oversample certain key subgroups in a population in order to examine their responses, however. In many communities, students or blacks may not be numerous enough to appear in large numbers in a random sample of the jurisdiction. Neighborhoods in which they dwell may be oversampled, and their responses can be down-weighted in analyses which are being projected to the city as a whole. Sampling strategies are thus related to the substance

of the problem at hand (e.g., how frequent are the attributes of interest) and the analyses in the data which are planned (e.g., what subgroups of the population will be examined in detail). Once these judgements have been made, sample selection is a problem best left to experts.

WHAT TO ASK

The problem of what to ask is tied to the purposes of a particular victim survey. In this section I will suggest some examples of the kinds of variables which should interest most surveyors, and some of the uses to which they may be put. In general, planners will want to know (1) what have the respondents experienced, (2) why have they had those experiences, and (3) what can be done to alleviate the crime problems of high-risk populations.

Commercial Targets

Commercial establishments should be a particular focus of attention in crime surveys, for they have extremely high victimization rates. One national sample of businesses indicated that in 1973, commercial establishments were burglarized at a rate of 20 per 100 places of business. The commercial robbery rate was 4 per 100, six times the personal robbery rate. These victimizations were not randomly distributed: some types of stores were hit much more often than others, and a few were victimized many times while others were never attacked. The crucial targets appear to be gas stations, convenience grocery stores, and liquor stores. The reason for their particularly high victimization rate is problematic, but victimization appears to be related to being open after dark and having few customers or employees about to make trouble for potential offenders. In general, perhaps 5 percent of commercial establishments experience about 75 percent of all commercial robbery, and this in turn makes up a substantial proportion (perhaps 25 percent) of all the armed robbery in an average community. This is a "pressure point" where the application of police resources might pay substantial dividends. It is also a fertile field for investigating the effect of various anti-theft efforts upon victimization rates. Within a sample of commercial establishments will be found all manner of crime-prevention devices: burglar alarms, floor safes, armed guards, and electronic shoplifting detectors. With the proper controls for other factors (hours of operation, employees, neighborhood), it may be possible to evaluate the effectiveness of various combinations of self-protective measures.

Household Targets

For some kinds of incidents it is useful to think of the household rather than an individual as the victim of a crime. Burglary, for example, is distributed in response to the vulnerability and desirability of the target household. Dwellings are desirable targets if they promise to

contain goods of high value which can be fenced. They also are vulnerable if they are protected by weak locks or doors, or if no one is home during the day, or if they are located in large and anonymous buildings without adequate security arrangements (see the study of residential crime by Reppetto listed in the bibliography). While burglary does not generate the kind of community insecurity which spills over from personal crime, it strikes a large number of people. In 1973, the household burglary rate in the United States was over 9 per 100 targets, higher than all the robbery, personal theft, and interpersonal violence put together.

Because of its high incidence, burglary may be a particularly useful subject for study. A modest local survey (3,000 households or so) would generate a relatively large number of events for analysis. Because it is not often solved (the clearance rate for burglary is under 20 percent) and not often reported to the police (less than half of the household burglary in 1973 was reported), efforts to reduce burglary must proceed on the basis of some new information. Surveys could collect information on the timing and style of entry, which could be used to alert police patrols. The impact of such programs as Operation Identification (property marking) could be assessed by comparing burglary rates for matched households which displayed and did not display signs signaling their participation. One innovative survey in Portland attempted to gather data to gauge the impact of street lighting on household crime by having interviewers count the number of streetlights visible from the doorstep.

Individual Targets

While individuals are frequently the targets of crime, data on their experiences is often of little utility for operational planning. Individuals, unlike households or stores, are not pegged to a particular geographical location. While it is informative to know the relationship between such factions as personal income and the probability of being victimized, police departments are organized along patrol car beats. Victim data is more useful if information on the location of the offense can be gathered in addition to information on the residence of the victim. This will not be easy, for individuals vary greatly in their knowledge of a community and how to describe it to an interviewer. Some will vaguely place an incident on "the East Side," others in some socially defined neighborhood, while some will be able to give an approximate street address. The goal should be to place each incident within some useful administrative unit like a police precinct; there it can be related to other area data and used for allocation purposes. Individual-level data can also be used to evaluate differential response time by neighborhood. This is important, for current measures reflect only administrative response time--the gap between the reporting of an incident and the arrival of a patrol car. In deterrence studies it is more important to know the length of time which elapses between the commission of the crime and the arrival of an officer on the scene, for almost all important arrests are made by beat patrolmen who apprehend offenders before they can flee the site. Personal interviews also provide a vehicle for assessing reports of where people will not go at night (suggesting neighborhoods that could use some visible patrols), and their fear of various kinds of victimization (suggesting an alternative to the use of incident totals for allocating resources).

SOME PRACTICAL SUGGESTIONS AND COST ESTIMATES

The decision to undertake a community victimization survey is a serious one, for they are expensive. The direct cost of personal interviews (in 1974 prices) is about $15-$18 per completed questionnaire. For a workable sample (say, 4,000 interviews) this amounts to about $70,000. The field interviewing in a large scale survey should be contracted to a market research firm or to a university survey research unit. For an additional fee such an organization will also draw a sample. A block-probability household sample is expensive to construct, and the fee using such a sample may amount to several thousand dollars. Analysis of the data can be contracted to a marketing firm or to a university, and will cost about half as much as the interviewing for a quick-pass examination. A serious survey analysis should also include an examination of official crime records. These records can be compared to survey rates for the same classes of incidents, and can be used to supplement survey data on those crimes for which the latter is an inappropriate measurement technique. Pulling police records, geo-coding them to match the survey data, and recording the data contained in the file, are clerical tasks of some magnitude. In the end, it may cost as much as $180,000 to conduct a community crime study of this size.

At present there are efforts underway which should help reduce the cost of conducting victimization surveys. Several years of effort have been expended by the Bureau of the Census and LEAA in developing a reliable survey questionnaire for measuring the incidence of more serious offenses; this could be adopted and modified for local use at little cost. The Police Foundation in Washington, D.C. is now sponsoring an investigation of the reliability of telephone surveys of victimization. Telephone surveys can be conducted for about 35 percent of the direct cost of interviewing, and if they prove valid they may greatly facilitate the use of victim surveys at the community level. Analysis expenses could be reduced greatly by the formation of consortiums of users who desire the same information. Running many data sets in the same fashion is not much more expensive than analyzing the results of a single survey. This is precisely the kind of coordinated activism that could be greatly facilitated at the state level by actions of criminal justice planning agencies.

BIBLIOGRAPHY

Bordua, David J., and Larry L. Tifft. "Citizen Interviews, Organizational Feedback, and Police-Community Relations Decisions," Law and Society Review, VI (November 1971), pp. 155-182.

Campbell, Donald T., and Julian C. Stanley. Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally, 1963.

Furstenberg, Frank F., and Charles F. Wellford. "Calling the Police: The Evaluation of Police Service," Law and Society Review, VII (Spring 1973), pp. 393-406.

Law Enforcement Assistance Administration, National Criminal Justice Information and Statistics Service. "Crimes and Victims: A Report on the Dayton-San Jose Pilot Survey of Victimization." Washington, D.C.: Law Enforcement Assistance Administration, National Criminal Justice Information and Statistics Service, June 1974.

Maltz, Michael. "Measures of Effectiveness for Crime Reduction Programs," Operations Research, XXIII (May-June 1975), pp. 452-474.

Reppetto, Thomas A. Residential Crime. Cambridge, Mass.: Ballinger Publishing Co., 1974.

Reynolds, Paul D., and Dale A. Blyth. "Sources of Variation Affecting the Relationship Between Police and Survey-Based Estimates of Crime Rates," chapter 16, Victimology: A New Focus. Edited by Emilio Viano and Israel Drapkin. Vol. III. Lexington, Mass.: Lexington Books, 1975.

Skogan, Wesley G. "Measurement Problems in Official and Survey Crime Rates," Journal of Criminal Justice, III (Spring 1975).

Skogan, Wesley G. "Public Policy and Public Evaluations of Justice System Performance," chapter 8, Crime and Criminal Justice Policy. Edited by John Gardiner and Michael Mulkey. Lexington, Mass.: D. C. Heath, 1975.

Skogan, Wesley G. "The Victims of Crime: Some National Survey Findings," Criminal Behavior and Social Systems. Edited by Anthony L. Guenther. 2d ed. Chicago: Rand McNally, 1975.

Smith, Paul E., and Richard O. Hawkins. "Victimization, Types of Citizen-Police Contacts, and Attitudes Toward the Police," Law and Society Review, VIII (Fall 1973), pp. 135-152.

CHAPTER 3

UNIFORM CRIME REPORTING: USES FOR CRIMINAL JUSTICE PLANNING

Federal Bureau of Investigation

INTRODUCTION

The criminal justice planner must maintain a perspective of the entire criminal justice system. While the criminal justice system has many facets not easily defined, it is generally accepted that the system is composed of several modules. Law enforcement, prosecution, courts, probation and parole, and corrections each play a vital role in the criminal justice process. Historically, these modules have acted with a high degree of uniqueness. It has been found that the interrelationship among these various modules has been minimal. To properly fulfill his responsibilities, the criminal justice planner must look to each entity while still maintaining a perspective of the whole system before efficient, effective planning can proceed. In nearly every instance, individuals responsible for committing criminal acts enter the system at the law enforcement level. Because of this fact, the planner must be fully knowledgeable of the information available to him from this module inasmuch as it will impact, eventually, all other modules of the system. The most readily available and most comprehensive set of data accessible to the planner relating to law enforcement is that provided by the Uniform Crime Reporting (UCR) Program. For the planner to effectively utilize this information, he must be cognizant of its scope and limitations. Further, he must be conversant with the terminology and procedures of this program if it is to be used in drafting futuristic plans for the entire system.

BACKGROUND

The UCR Program was conceived, developed and implemented by law enforcement for the express purpose of serving law enforcement as a tool for operational and administrative purposes. (Throughout this presentation, the term law enforcement is used to refer to the traditional police function of the criminal justice system.) Under the auspices of the International Association of Chiefs of Police (IACP), the UCR Program was developed in 1930. Prior to that date, no comprehensive system of crime information on a national scale existed. This was, primarily, due to the fact that the criminal statutes varied so greatly from state to state in their use of terminology to define criminal behavior. To overcome this problem, a set of definitions for specific criminal acts was devised. These definitions were developed following an in-depth research effort relative to each of the state criminal statutes extant. It was determined that law enforcement would tabulate the number of criminal acts as defined by the UCR Program as these acts were brought to the attention of law enforcement. Recognizing the problem of coping with mere volume, it was decided that only those criminal acts deemed serious would be counted. It was necessary to define "serious." In this regard, a criminal act is considered serious if it meets a set of criteria; namely, that the act would occur regardless of geographical location; that it would be an offense most likely to be

reported to law enforcement; that it would affront the moral sensitivities of our society's rational being; and that it would occur with sufficient frequency so that counting it would be worthwhile statistically. In applying the concept of seriousness to the multitude of criminal acts identified, a select group evolved. This group of criminal acts is referred to as the "Crime Index."

The Crime Index offenses and their definitions are set forth below:

1. CRIMINAL HOMICIDE.--(a) Murder and nonnegligent manslaughter: All willful felonious homicides as distinguished from deaths caused by negligence. Excludes attempts to kill, assaults to kill, suicides, accidental deaths, or justifiable homicides. Justifiable homicides are limited to: (1) The killing of a person by a law enforcement officer in the line of duty; and (2) The killing of a person in the act of committing a felony by a private citizen. (b) Manslaughter by negligence: Any death which the police investigation established was primarily attributable to gross negligence of some individual other than the victim.

2. FORCIBLE RAPE.--The carnal knowledge of a female, forcibly and against her will in the categories of rape by force, assault to rape, and attempted rape. Excludes statutory offenses (no force used--victim under age of consent).

3. ROBBERY.--Stealing or taking anything of value from the care, custody, or control of a person by force or violence or by putting in fear, such as strong-arm robbery, stickups, armed robbery, assaults to rob, and attempts to rob.

4. AGGRAVATED ASSAULT.--Assault with intent to kill or for the purpose of inflicting severe bodily injury by shooting, cutting, stabbing, maiming, poisoning, scalding, or by the use of acids, explosives, or other means. Excludes simple assaults.

5. BURGLARY--BREAKING OR ENTERING.--Burglary, housebreaking, safecracking, or any breaking or unlawful entry of a structure with the intent to commit a felony or a theft. Includes attempted forcible entry. The UCR definition does not include auto burglaries, burglary of moveables, or a wide variety of such incidents as included in some state statutes.

6. LARCENY-THEFT (EXCEPT MOTOR VEHICLE THEFT).--The unlawful taking, carrying, leading, or riding away of property from the possession or constructive possession of another. Thefts of bicycles, automobile accessories, shoplifting, pocket-picking, or any stealing of property or article which is not taken by force and violence or by fraud. Excludes embezzlement, "con" games, forgery, worthless checks, etc.

7. MOTOR VEHICLE THEFT.--Unlawful taking or stealing or attempted theft of a motor vehicle. A motor vehicle is a self-propelled vehicle that travels on the surface but not on rails. Specifically excluded from this category are motor boats, construction equipment, airplanes, and farming equipment.

Since the inception of the UCR Program the FBI has acted as administrator, by Congressional mandate, of the program. Since 1930, timely

publications have been issued presenting the results of the monthly crime data collection effort.

STATE UCR PROGRAMS

During that period of time when UCR was still a concept, it was recognized that the individual states would also need crime information of particular interest to the state but of no great importance to the national view of crime. It was not until the latter part of the 1960's that funds became available for states to consider the development of their own individual reporting systems.

The purpose for state UCR programs is multifaceted. First, with personnel administering a state program, more direct and meaningful contact with individual contributors is realized. Second, the ability to expand contributorship exists due to state personnel readily available to lend assistance. Nearly every state thus far enjoying the services of a state UCR program has enacted mandates requiring law enforcement agencies to participate. Third, mandatory participation insures a law enforcement agency will either enhance an already existent records system or will implant one capable of producing the needed data. Fourth, with state personnel reviewing information emanating from law enforcement contributors and this information being checked at the national level, the validity as well as completeness of data is further insured. Fifth, individual state programs can address problems that are unique to the state. For example, numerous Northern states are vitally concerned over the theft of snowmobiles while this data is of little or no interest to those states in more Southern climates.

As of 1975, 32 states have operational UCR programs. It is anticipated that at least nine more states will be operational during calendar year 1976.

The procedures utilized in data handling for state programs and found to be highly effective are for the state program to continue the collection, in summary form, of basic UCR data. Upon acquiring at least one year of experience of such collection, a state program is urged to sophisticate its collection efforts. There are states currently embarked on incident-based collection programs. In these programs, individual law enforcement agencies provide the state program with a copy of an actual incident report which is standardized throughout the state. Under this concept, the individual contributor is relieved of UCR data compilation responsibility. With the incident reporting form in hand, the statistical effort can address such crime-related problems as stranger-to-stranger offenses, etc.

The state programs are expected to feed back to individual contributors the necessary information required by the agencies for administrative and operational purposes. State programs are urged to maintain close and direct contact with the contributors to insure the needs of law enforcement are being met.

TYPES OF INFORMATION COLLECTED

Offenses

Law enforcement agencies are requested to provide, on a monthly basis, Crime Index offense information with selected breakdowns. The collection device is structured in order that the agency may tabulate all Crime Index offenses brought to its attention. It is further designed to permit a tabulation of those offenses which upon investigation have been found to be false or baseless reports. The total volume of all bona fide actual Crime Index offenses is then tabulated at the national level.

Clearances

For UCR purposes, a clearance is defined as the arrest of one or more individuals responsible for the commission of an Index offense. There are exceptions to this rule wherein a clearance is scored by exceptional means. This is true in the event that the individual, for some reason beyond law enforcement's control, cannot be arrested or otherwise handled. Clearances of reported actual offenses committed by those individuals under the age of 18 who have been arrested or otherwise handled for the commission of an Index offense are also counted. This is done in an effort to identify the youthful offender problem.

Value Loss

In order to supplement the monthly offense and clearance data, the value of property stolen and recovered by specific type of property is collected. Value loss is also collected as it relates with each of the Index offenses along with additional detail for each of the offense classifications.

Arrests

Arrest information is a valid measure of one aspect of law enforcement activity. To fully capture this data set, the UCR Program requests, on a monthly basis, the number of persons arrested, exclusive of traffic violations. This arrest information is broken down by age, sex, and race of the person arrested.

Law Enforcement Officers Killed and Assaulted

To fully explore the perils associated with the law enforcement profession, data is gathered on a monthly basis relative to those situations wherein officers are assaulted. This data collection effort attempts to identify the type of officer activity, type of weapon used by the offender, type of officer assignment, extent of officer injury, and time of day for each of the assaults. Law enforcement agencies are readily guided by this data set in the development of training programs and acquisition of equipment designed to minimize situations that are dangerous. This information further acts as a warning device to officers given specific assignments.

An adjunct to the assault studies is the detailed analysis of situations wherein an officer dies as the result of a felonious attack.

Disposition Data

On an annual basis, law enforcement agencies are requested to provide summary totals on the number of persons formally charged by type of criminal act. The results of these charges are requested by guilty verdict, guilty of a lesser offense, acquitted or dismissed, or whether the chargee was turned over to juvenile court.

Law Enforcement Employee Data

On an annual basis, each contributing law enforcement agency is requested to provide the number of full-time, sworn and civilian, employees by sex, on duty as of October 31 of the specific year. Through 1975, certain data elements of shift assignment have been collected.

Bomb Information

While not a direct part of the UCR Program, semiannual publications are issued relative to actual and attempted incendiary and explosive bombing situations occurring throughout the United States and Puerto Rico. This data presents such occurrences by geographic region, division, and state, as well as by ascertainable motivating factors.

TYPES OF INFORMATION PUBLISHED

Each quarter, on a calendar year basis, the UCR Program issues a report reflecting trend information on the Crime Index offenses. The presentation is by percent change of the current time period with the same period of the prior year. The trends are exhibited on the basis of total volume as well as by violent and property crimes and also by each of the Index offenses. A breakdown by population grouping, by cities and counties in terms of trend, is set forth. A five-year, long-term trend table for each of the offenses is presented. Also in the quarterly report is the useable Crime Index information for all cities over 100,000 population.

On an annual basis, the UCR Program issues a comprehensive annual report entitled "Crime in the United States." This document sets forth the total United States' crime picture based upon the UCR Program. Tabulations by region, division, and state; Standard Metropolitan Statistical Area; and individual cities having over 10,000 in population are presented in detail relative to the Crime Index. Correlating clearance, arrest, and disposition data are presented. Tables dealing with short- and long-term trends relative to offenses, clearances, and arrests portray the crime experience of the United States. While much of the information, other than that dealing with offenses, is not published agency by agency, this data is presented by population grouping.

The presentation of arrest information is not published agency by agency as in the case of offenses. The mere volume precludes such

publication. However, available data that is useable is provided upon request. Arrest information is presented in summary form by geopolitical breakdowns as well as by age, sex, and race. Both current volume and trend data by population grouping are tabulated.

Law enforcement employee information relative to sworn and civilian employees is presented in summary fashion. Tables reflecting interquartile ranges, by population groupings and by officer to population rate, show the law enforcement strength. Tables are constructed by agency for those places with a population of 25,000 or more inhabitants.

A special section of the annual publication, "Crime in the United States," deals with the criminal repeater. In 1963, a program entitled "Careers in Crime" was initiated. Individuals arrested for the first time for a Federal offense were tagged for tracking. The statistical effort did not deal with identities but only with numbers of offenders. As time passed and these individuals were rearrested, the statistical data was accumulated and certain observations were apparent on the basis of the rearrest experience. The 1969 issue of "Crime in the United States" renders a comprehensive view of the Careers in Crime effort.

With the advent of the Computerized Criminal History (CCH) File, administered by the FBI, an updated data set was established. The 1972 publication of "Crime in the United States" presents more recent tabulations on repeaters in terms of rearrest experience. Profiles of offenders rearrested as well as the mobility factor of rearrested persons are shown for the time period 1970-1972. Follow-up studies with this data base appear in the annual publication for 1974.

LIMITATIONS OF DATA

The UCR Program is not a perfect program. While it can be of invaluable assistance to the criminal justice planner, there are definite limitations with which the planner must be familiar.

Not all law enforcement agencies enjoy the services of a records system capable of producing UCR data as a by-product of the system's overall design. As a side note, the IACP, in developing the UCR concept, recognized the need for a law enforcement records system within each agency. It was felt the individual law enforcement agency would develop such a records system in order to be an active participant in this program. Their feelings have proven correct. In 1930, there were only a little over 300 participants and in 1975, there are in excess of 12,000 law enforcement agencies providing UCR data, at least in part. The fact still remains, however, there are some agencies which do not participate and, consequently, there is not 100 percent coverage of the Nation.

Another limitation deals with the incompleteness of data. Certain agencies are able to provide limited segments of the total UCR request but lack the facility to respond totally. The inability to provide complete data is generally caused by the type of records system utilized by the agency. The system may well serve the departments' needs efficiently but is not designed to capture total UCR data.

It must be borne in mind that from the standpoint of criminal incidents, only seven are tabulated on frequency of occurrence. Therefore, the UCR Program addresses only a portion of the total crime picture reported to law enforcement.

The recommended procedures for developing UCR data at the contributor level are set forth in a handbook provided by the FBI. As in any statistical program of this nature, the risk of misinterpretation is always present. This risk, combined with a degree of subjectivity on the part of those responsible for capturing UCR data, may create a bias in the data. This shortcoming is partially overcome by correspondence between the FBI and contributors and in numerous training sessions held throughout the country.

A more subtle and pervasive limitation deals with program and policy changes within an individual department. For example, an agency's burglary volume may increase substantially, not because of an actual increase in the number of burglaries occurring but because the department has embarked on a broad base program to reduce burglaries. The public, developing an awareness to the interest of law enforcement, proceeds to report burglary offenses that in the past it may have ignored. The planner, when studying trend data for individual agencies, must be alert to such aspects which could affect crime data.

Along similar lines as the aforementioned, influence can be exerted on crime data by the general attitude of people and unique legislation in given geopolitical areas. Numerous places are interpreting drunkenness as a noncriminal activity. Some areas have decriminalized certain drug abuse violations. Consequently, this will have an impact on arrest information. Again, the planner must be alert to these situations.

The criminal justice planner must be cautioned to not utilize UCR data as a stand-alone body of knowledge. There are numerous factors which influence the occurrence of crime. As a caveat, each annual publication of "Crime in the United States" presents a partial listing of those sociological phenomena which are influential in crime causality.

CALCULATION OF RATES AND TRENDS

To develop meaningful plans, the criminal justice planner must, at some point, identify areas requiring immediate or long-term intensive attention within the plan. He must, at some point, array numerical data in such a way as to provide a picture which will identify areas of need. While is is recognized that examining the volume of crime occurrence either totally or by individual offense over a period of time for a particular agency is important, the planner must also view this experience for a selected group of agencies. An agency equalizing factor must be developed. In short, he needs a common denominator. The geographical area he is studying will undoubtedly consist of law enforcement agencies of various strengths and serving widely varying sizes of population. To accomplish a leveling technique, the concept of rate is suggested.

The crime rate is defined as the number of Index offenses occurring per 100,000 people over the period of one year. In the way of example, a community with 100,000 people that experiences a reporting of 2,500 Crime Index

offenses for a given calendar year is said to have a crime rate of 2,500. A second community having a population of 20,000 people that experiences 500 reported Crime Index offenses in a given calendar year is also said to have a crime rate of 2,500.

A common factor for any community can be acquired by dividing the community's population by 100,000. Once this factor is determined, it becomes a relative constant until such time as the population of the community changes appreciably. The crime rate can be ascertained by dividing the total reported Index crime by the factor. If the planner wishes to identify the crime rate for an individual offense, the total number of that reported offense is divided by the factor.

Frequently, a planner may desire to compute a crime rate on partial year data. When this is done, only a fractional portion of the jurisdiction's population should be used, correlated with the fractional part of the year to be observed. If a city has 100,000 people and the crime rate for the first three months is desired, then only one quarter of the population would be used in computing the crime rate. In this instance, 25,000 would be divided by 100,000 (one quarter of the city's population) and the factor becomes .25. The offense counted is then divided by this factor and the result would be the crime rate.

Crime trend information is a presentation of the fluctuation of data from one time period to another. Trends can be computed on the basis of volume or rates. In effect, a trend is presented as a percent change. The important factor to consider here is that similar data bases and similar time periods are used for comparison in determining the trend. To compute a trend, always subtract the current period from the prior period. The result is then divided by the prior period. The result of this division represents the percent change or, in other words, the trend.

Trend compilations can be applied to any of the many sets of UCR data, just as long as the aspects of jurisdiction and time-frame are similar.

The development of trend data requires certain degrees of discipline within the collection methodology. The observer must be assured of the constancy of data conversion. There should be no evidence of policy or procedural changes in data capture within a given agency. The observer needs two or more set time periods for data observation. If he can rely on these two aspects, then a trend observation can be made. Trend tabulations are no more than an aligning of what has occurred within a given recent time period with what occurred in a prior identical time period, all identifying characteristics being equal.

Assuming that seven law enforcement agencies in a given geopolitical area experienced 600 reported Index offenses for a three-month period and all seven had population coverage of 80,000 people, only the following can be observed.

Agencies	Population	Offenses Reported	Period
7	80,000	600	3 months

If, in a later period, preferably a year or more, the same set of agencies has 750 offenses reported for the same three-month period, a more meaningful presentation will develop.

Agencies	Population	Offenses Reported	Period	Percent Change
7	80,000	Period 1(600) Period 2(750)	3 months	25%

The rate change, affected by an adjustment for partial year coverage of population, would be a crime rate in the first period of 3,000 and in the second period 3,750. The percent difference, or rate trend, becomes 25 percent.

Historically, trend data is computed on the basis of volume for Crime Index offenses on a quarterly basis. Rate differentials are reserved for annual observance.

One of the measures of law enforcement effectiveness is the clearance rate. The clearance rate is calculated by dividing the number of offenses cleared, as per this program's definition of clearance, by the number of offenses known with the result being multiplied by 100. If an agency should experience a total of 72 robberies, 38 of which were cleared, a clearance rate of 52.8 percent would exist.

While the crime rate is computed on the basis of the number of Index offenses per 100,000 people, the police employee rate is developed on the basis of 1,000 people. Both concepts utilize the same arithmetical functions. To compute the police employee rate, divide the jurisdiction's population by 1,000, then divide the number of police employees by the result. When the employee rate is referred to, it is always construed to be the number of police employees per 1,000 people.

As mentioned earlier, UCR information represents one data base. The planner must consider other data bases in order to acquire a well-rounded picture of the sociology of the given jurisdiction. He must look to such sets of data as school dropout rates; unemployment rates; economic levels; gross population characteristics relative to age, sex, and race and from these data sets, develop rates and trends that correlate arithmetically with the crime data manipulations.

APPLIED USES OF UCR DATA

The criminal justice planner must respond to the needs of the criminal justice community in his area. Before he is able to respond to these needs, he must first identify them, determine priorities, and proceed in such a fashion that, while meeting the needs of one aspect or module of the criminal justice system, he does not unnecessarily intensify existing needs or

create additional needs in some other module of the system. In an effort to preclude such situations, a starting point for study must be found. It would appear the logical point would be within that module first dealing with matters of interest to the entire criminal justice system; namely, law enforcement. Another consideration would also be to consider that module first which is able to provide the most comprehensive picture of its activities. In many instances, this would again be law enforcement.

In each of the following suggested uses of UCR data, it is assumed that efforts have been made to acquire other information relating to sociological phenomena outside the purview of law enforcement.

1. Offenses Known--Create a line chart covering a five-, and if available, a ten-year illustration of the volume variation of reported offenses. Now, do the same for the crime rate. By overlaying similarly constructed data from other data sources both from within and without the criminal justice system, significant patterns may evolve. Certain projection characteristics may offer the planner a guide that would have an impact on all modules separately and/or collectively.

2. Clearance Information--To some degree, this is a measure of law enforcement efficiency. To use this data, structure graphs showing time-series experience by individual agencies and groups of agencies. Make correlations with the national experience and with the appropriate population group experience. The annual publication, "Crime in the United States," gives this information.

Prosecution and court information, overlayed with clearance data, will give an in-camera view of the workload correlation and provide for future planning.

Care must be taken in time-series analysis to allow for impact of such things as precedent setting court decisions which have had a direct effect on caseload responsibility. An example of this would be the effect of a court decision impacting the length of time to try a case.

3. Youthful Offenders--By use of additional background, the criminal justice planner can look to three areas of UCR data collection for information on youthful offenders. First, one segment of clearance information addresses those offenses cleared by persons arrested under the age of 18. Second, age, sex, and race characteristics of persons-arrested data present arrest information in such a format that specificity of age can be determined. Third, in conjunction with the collection of arrest data, individual agencies are asked to provide data on juveniles (as defined by state statute) that are (a) Handled within the department and released; (b) Referred to juvenile court or probation department; (c) Referred to welfare agency; (d) Referred to other police agency; and (e) Referred to criminal or adult court.

The planner may desire to structure tables and graphs reflecting the law enforcement experience with the youthful offender, correlating the aggregate data with population composition; school dropout rates; family size and composition; detention facilities; population configuration; parole, probation, and welfare caseload composition; and other data sets wherein youth or juvenility are intrinsic.

By utilizing this data, again in time-series, certain observations can be made. To some degree, the time lapse can be determined between encountering the youthful offender by law enforcement and the processing of the individual at later stages of the criminal justice system. Workload projections will be inferred by pictorially presenting this information with bar graphs and other such analytical devices. As in other compilations, care must be taken to make the observations on the basis of volume, rate, and trend.

By wedding the various data bases, the planner may identify needs not heretofore encountered by any one or all of the criminal justice modules. It is conceivable the planner may decide to suggest a strengthening of a law enforcement agency's manpower commitment to youthful offender crime. He may, in turn, envision a need for intensified foster home programs or halfway houses. He may discover that while detention facilities currently being used are sufficient, in the next five years they will be taxed to the detriment of the overall system.

4. Arrest Information--As mentioned earlier, this data is partially a measure of police activity. By constructing various visual analytical devices, utilizing time-series, the planner can identify the fluctuations of law enforcement activity. To some degree, such compilations test the effectiveness of programs initiated not only within the law enforcement module but also the other modules of the system.

By testing law enforcement's arrest experience against the age composition of the population and correlating this information with other data sets, some of which have been previously mentioned, and again applying time-series, limited projections will evolve.

It should be recalled that clearance information does not relate directly to arrest information; a clearance can be scored only by the arrest of one or more persons and conversely, several clearances can result from the arrest of only one person. With this thought in mind, associating total arrest numbers for specific offenses with the clearance rate, certain observations can be made. Offenses which inordinately cause the arrest of numerous people can be isolated and innovative programs designed to address this problem. This is one of many instances where the criminal justice planner may be in a position to recommend programs to agencies and groups which are outside the purview of the criminal justice system.

Arrest data on the sex of persons arrested, utilized in ways previously mentioned, may suggest to the planner the need for future detention facilities, personnel composition of parole and probation services, and the development of innovative programs within law enforcement agencies to handle female offenders not heretofore needed.

Certain implications are involved when analyzing the population composition of persons arrested. By correlating such data with economic, employment, education levels, housing, etc., data and applying age, sex, and race characteristics to these data sets and again using time-series, it may be determined that a small, suburban law enforcement agency can anticipate a significant effect on activity at some point in the future. Population shifts and mobility must be taken into account. This is only one application dealing with only one module of the criminal justice system. Obviously, all modules will eventually be impacted.

5. Disposition Information--Of all the sub-data sets in UCR discussed, disposition information is generally the most difficult for law enforcement to acquire. This is primarily because this data set is ordinarily outside the normal channels of information within a law enforcement agency. Once a matter is presented for prosecutive action, the law enforcement agency, to a degree, loses control of pertinent statistical information relating to the disposition of offenders. It is only in those instances where a smooth, cooperative information exchange system among the law enforcement, prosecution, and court modules exists that the law enforcement agency is able to report efficiently on the disposition of cases handled. Consequently, the criminal justice planner may be unable to rely on law enforcement to provide this data. In those instances where it is available, obvious benefits for the planner are forthcoming. The planner will be in a position to test offense information, aligning it with clearance data and, in the case of Crime Index offenses, observing in summary number form, the actions of the joint effort of courts and prosecution. He may be able to identify specific areas which have a debilitating effect on the caseload of prosecutors. This may suggest to the planner a need for expansion of the prosecutor's office, added training for prosecutors; it may, in effect, identify a need at the law enforcement level, and it may suggest to the planner certain items of legislation to be recommended to the appropriate authority.

6. Constructing Law Enforcement Agency Profiles--A table can be constructed utilizing three data sets from each individual law enforcement agency. By correlating law enforcement employee data with the number of offenses reported, the volume being reduced to a rate, a percent change or trend can be constructed. The second portion of the table would reflect the number of arrests per offense and a ratio tabulated. The number of offense arrests can then be computed on the basis of per police employee. The third portion of the table deals with clearance information. The total number of clearances, correlated with the total number of offenses, correlated with the number of arrests, and further correlated with law enforcement employees completes the profile. By constructing such profiles over time, administrative observations can be made. (See Figure 1.)

AVAILABILITY OF DATA

The planner's attention is drawn to the limitations of this program. Certain data sets are not available from certain agencies. However, inquiry to the state agency responsible for state data collection or to the FBI will resolve any questions of data availability. If the planner is dealing with a limited number of agencies in close geographical proximity, direct contact with those agencies will, in most instances, provide a quick response to data requests. In the event more expansive coverage is required, or if the agency refers the planner to the FBI, he should correspond with the FBI's UCR Section, Washington, D.C. 20535, to determine what data is available. In many instances, the FBI will provide reported information, and for those agencies not actively participating, an estimate is available. The FBI will withhold information if some question of completeness exists. Completeness and validity can be affected by a change in records systems, annexation or deannexation of geography and/or population, or some other phenomena that would cause a current crime count to deviate from a prior comparable period.

Another source of UCR data is from state UCR Programs. Thirty-two states have such efforts and are listed in the annual publication, "Crime in the United States."

AGENCY PROFILE

City of _____ Sworn Officers _____

197_ Population _____ Total Employees _____

OFFENSE TAB

Crime Index	Number	Per 100,000	% Change 1971-1972
Total	3,067	2748.2	24.1-
Murder	6	5.3	
Rape	17	15.2	
Robbery	223	199.8	20.0-
Assault	97	86.9	12.1-
Burglary	1,377	1233:8	30.6-
Larceny	706	632.6	10.9-
Motor Vehicle Theft	638	571.6	24.8-

ARREST TAB

Crime Index	Number	Per Offense	Per Officer
Total	635	.2070	1.702
Murder	5	.8333	.013
Rape	15	.8824	.040
Robbery	106	.4753	.284
Assault	84	.8659	.225
Burglary	170	.1234	.455
Larceny	195	.2762	.522
Motor Vehicle Theft	57	.0893	.152
Part II Arrests	5,181		13.890

CLEARANCE TAB

Crime Index	Number	Per Offense	Per Arrest	Per Officer
Total	506	.1649	.796	1.3565
Murder	6	1.0000	1.200	.0160
Rape	14	.8235	.933	.0375
Robbery	65	.2914	.613	.1742
Assault	85	.8762	1.011	.2278
Burglary	148	.1074	.870	.3967
Larceny	137	.1940	.702	.3672
Motor Vehicle Theft	48	.0752	.842	.1286

Figure 1. LAW ENFORCEMENT AGENCY PROFILE

CHAPTER 4

CRIME STATISTICS ANALYSIS IN MASSACHUSETTS

John R. Pfeiffer

INTRODUCTION

Analysis of crime patterns and trends is gaining more attention as an important criminal justice planning tool. It is recognized as a prerequisite to sound criminal justice planning at state and local levels of government. An increasing number of state and local criminal justice planning bodies are intensifying efforts to utilize data about crime more fully in their planning. Since analysis of crime statistics has been the province of academicians until recently, guidelines for such analysis are not available in laymen's terms. The following discussion explains how criminal justice planning agencies can analyze crime patterns and trends in a way that is useful to planners and program developers. A model is offered that is based on that used by the Massachusetts State Planning Agency (SPA) but should be easily applicable to other states.

The discussion is divided into three parts. The first part reviews major considerations in planning the crime report. The second part discusses the organization and content of the crime report in more detail. The final part reviews some mechanics of data analysis and points out some of the limitations involved in using Uniform Crime Report (UCR) statistics for policy analysis and planning.

PLANNING THE REPORT

The first step should be the development of a good plan for preparing the report. Careful planning of data to be presented, methods of data analysis, and report format greatly facilitate successful project completion. Extra time and effort spent planning can save the time and labor involved in redoing the data analysis or revising parts of the report. There are several important questions to be answered at this stage. We shall comment on each of them briefly. They are all somewhat interdependent, the answer to one being contingent on the answer to another. Consequently, their relative priority is not as important as having answers to all of them.

What Do We Want To Include in the Report?

Many organizations, including the FBI, present little more than tables of statistics in their reports. This approach is simpler and less time consuming than adding analysis and narrative description of the data. But even if you only provide statistical tables, you must decide what data to include and how the tables will be organized. Whether or not the report is to contain much narrative, some organizing scheme must be developed and limitations set on the topics covered. More about this later.

How long the report should be is also a significant consideration when deciding what to put in it. Even a seemingly narrow topic like the geographic distribution of reported offenses can easily fill 50 to 100 pages of text, charts, and tables. Given the potential readership's interests and attention span, shorter, efficiently presented reports might be better received and have more impact than a longer, more comprehensive but more forbidding volume.

What Can Be Included in the Report?

The data available obviously limit the contents of the report. The two most widely known data sources are the FBI's annual Uniform Crime Reports and the 1970 U. S. Census volume for your state.[1] A good report can be prepared from just these two sources where other data are not accessible. To go beyond reported crimes and their demographic correlates--for example, to include data on the organization and activities of different parts of the state criminal justice system--may be difficult presently in some areas. It may require asking other state and local officials for information, consulting other agencies' annual reports, or conducting a formal survey. These efforts add to the time needed to complete the report, but can be very worthwhile. If the data are of poor quality or of limited utility to readers, it may be desirable to initiate methods for improving the data.

The Sourcebook of Criminal Justice Statistics: 1973 and Expenditure and Employment Data for the Criminal Justice System: 1972-1973 contain data on criminal justice operations in all states and larger counties and cities.[2] These volumes can be helpful in attempting more extensive data collection.

How Will the Data Be Compiled and Analyzed?

Two options are offered here. One can use paper and pencil and a hand calculator--or punchcards and a computer to do the analysis. The latter approach sometimes avoids arithmetic errors and allows for more complex and interesting analyses of the data. The number of cases (cities and towns) will be so few and the kinds of statistics generated will usually be so simple that computer costs would be quite small.

How Much Staff Time Is Available for the Project?

A conceptually simple report can take a surprising amount of time to complete. Original time estimates should probably be doubled to be realistic. For example, one person could reasonably take two to three months working full time to prepare a straightforward report on the geographic distribution of reported crime, including planning, preparation of charts and tables, writing the narrative, and editing.

REPORT ORGANIZATION AND CONTENT

This part reviews the organization of the Massachusetts crime analysis report. Essentially, it provides an annotated table of contents for one type of crime report. The Massachusetts crime report focused on the distribution of reported UCR crime in the state and examined the relationship of crime distribution to some basic demographic indicators. Both the UCR statistics and the demographic statistics were available in nationally published documents. This approach dealt only with the crime problem as reported by local police. It excluded other possibilities. In particular, it did not cover the criminal justice system's reactions to reported crime. For example, it did not discuss rates of arrest, conviction, or sentencing. Moreover, the report did not discuss the impact of crime on victims or on the criminal justice system. For example, no mention was made of dollar loss to crime victims or of the costs of operating police, courts, and corrections agencies. Nor did the report compare victimization statistics (another indicator of the actual incidence of crime) with the numbers of crimes reported to police.

There are several reasons that these alternatives were foregone. First, the other types of data were either unavailable or difficult to obtain. Second, a briefer report with one theme was likely to have more impact. Later reports could explore the other issues. Third, time to prepare the report was limited. Inclusion of only one or two of the other topics might have given undue emphasis to the crime problem as somehow the responsibility of a particular part of the criminal justice system. However, careful analysis and interpretation of additional data will often yield better information for understanding crime and the criminal justice system comprehensively.

An area-specific rather than a crime-specific approach to organizing the report was chosen. That is, each section of the report was written about crime trends either in the state as a whole or in particular parts of the state rather than about a particular crime and its distribution. The area-specific approach facilitated analysis of the total crime problem within a given area and comparisons among different parts of the state, but it did not preclude examination of trends in reporting of specific offenses.

Consequently, the report was organized as follows. There were two major parts. The first part contained the analysis of state crime trends along with supporting tables and graphs. The second part contained a series of maps of crime distribution in the state and a listing of all communities for which crime data were available, along with their demographic characteristics and crime statistics.

Part One of the Report

The introduction to Part One included a brief statement of the purpose of the report, suggested some of its possible uses, and described its organization. There was also a commentary on the weaknesses of the kinds of data used in the analysis. It is extremely important that readers be clearly informed about limitations inherent in the Uniform Crime Reports as measures

of crime. Some of the limitations will be described in the final part of this paper.

Next, the report presented an overview of trends in statewide crime rates in recent years, using data from several of the latest UCR volumes and compared the trends in the state with those in the nation. Comparisons with neighboring states or with states having similar socioeconomic characteristics also would have been appropriate here.

After the review of crime trends in the state as a whole, sections were presented describing crime patterns in each of the officially designated planning areas and sub-areas of the state. Each section started with a description of the demographic characteristics of the area. The total volume and rate of crime in specific areas relative to those in other areas and in the state as a whole also were discussed here. Then there were comparative descriptions of the patterns of each of the violent crimes and of each of the property crimes. When there were any large cities in the area, the section ended with a discussion of cities' impact on crime statistics for the area.

A section on correlational analysis of the relationships between crime rates and socioeconomic indicators for each town for which crime data were available was planned for the report. The section had to be dropped. The planned analyses were not undertaken when it became apparent that the statistical form of demographic variables was incompatible with the form in which the crime statistics existed. In this case a lesson was learned after the fact on how to accommodate data that contain both median statistics and percentages. The situation offered a lesson in the complexities of statistical analysis and interpretation, but the explanation goes beyond the intent of this section. The reader is referred to Robert Gordon's article.[3]

The final section of the first part of the report summarized findings about crime patterns in the state and outlined policy implications of the data. We pointed out that UCR statistics can only indicate where crime is being reported but not why. The data could indicate where in the state particular programming efforts probably should be focused but not what form those efforts should take. Consequently, the policy recommendations suggested particular parts of the state where programs to fight certain crimes should be concentrated. Additionally, specific recommendations were made for improvements in collection and analysis of crime data in the state.

Part Two of the Report

The second major part of the report contained a series of maps of the state divided according to geographic planning areas. Individual planning areas were then shaded in to indicate variations in the distribution of specific crimes and demographic factors. In addition, an alphabetical listing of all cities and towns for which crime data were available was included in the second part of the report. Each listing showed in which planning area the community was located, the number of crimes reported, the crime rate for all offenses combined and for each individual type of offense, and the rank order of that community within its planning area and within the

state on each crime. Ranks were given for both the crime volume and the crime rate. For example, one small resort town was ranked first in the state for its crime rate but its rank for volume of crime was much lower. Similar statistics and ranks were listed for the demographic variables. Figure 1 illustrates the listing format.

CITY	CRIMES	Offense
	NUMBER	Number reported
AREA #	RANKS	(State Area Region)*
REGION #		
	RATE	Offenses 100,000 population
	RANKS	(State Area Region)*
	DEMOGRAPHIC DATA	Demographic characteristic
	VALUE	City value on that characteristic
	RANKS	(State Area Region)*

Figure 1. FORMAT FOR LISTING EACH COMMUNITY'S CRIME AND DEMOGRAPHIC CHARACTERISTICS WITH RANKS

*City's rank for that offense or demographic characteristic in the state, area and region.

DATA PREPARATION AND LIMITATIONS

Coding the Data

Before computer analysis can begin, the data for each city and town must be abstrated from the UCR listings from the U. S. Census volume and other sources, written on coding sheets, and keypunched onto cards. This effort requires development of a scheme for coding the data after the variables to be analyzed are chosen. Decisions must be made about where each variable is to be recorded on a punchcard; how many card columns will be allowed for each variable; and if more than one card is needed for each community, which card each variable is to be punched on.

To keep track of all this and to make it easier to do the next year's data coding, it is very helpful to prepare a codebook. A codebook is like an index. It tells which variables were coded and where they are located in a set of data cards. Indication of the sources from which particular variables were obtained would also be worth including. A "Codebook for State Crime Analysis, 1975" could easily be written using a format like that in Figure 2 for each variable.

Card # Columns	Variable Name	Source
Card 1		
32-34	Homicide Volume	1973 UCR, Table 75
35-36	Homicide Rate	1973 UCR, plus 1970 Census, Table 10

Figure 2. FORMAT FOR VARIABLE CODEBOOK WITH EXAMPLES

After the tedious job of coding is completed, the analysis can begin. The statistics needed for the report probably will consist of simple cross-tabulations or correlation coefficients. Programs for computing these may be available at a state computer center. If not, they are available at most colleges. Specialized programming assistance might be needed only if special printouts are planned for publication in the report.

Limitations of UCR Data

It was mentioned earlier that it is extremely important that readers and users of the report be clearly informed about weaknesses in the data and about the kinds of interpretations that may be legitimately derived from the data. Some readers may not be experienced in the complex use of statistics. They should not be misled into thinking the data explain more than they actually do.

The most serious interpretation problem likely to arise--especially if crime report statistics are correlated with socioeconomic or other indicators--is that readers may think such data show what or who causes crime. Everyone wants to know that, but the UCR data cannot tell us. UCR data show only which communities are reporting how much crime. Knowing the location of crime is valuable information in its own right but, by itself, it does not provide valid evidence on the causes of such crime.

To illustrate: A positive correlation of reported crime with cities having a higher percentage of population older than 60 years of age simply shows that crime is more frequently reported in cities with a large proportion of older people. That correlation cannot be used to conclude that old persons are more likely to be victims of crime or to commit crime or to report crime, but it does suggest the usefulness of further data and analysis to test the two more reasonable hypotheses--higher victimization and reporting rates. There is no way to tell from the statistic itself why such cities report more crime. Hypotheses can be developed about why the correlation occurs but other data have to be investigated to find an answer.

Another difficulty is in distinguishing crime rates from crime volume. For example, a small town can have a higher rate of crimes reported per 100,000 population than a larger city, while the absolute volume of crimes reported can be much greater in the big city. Another example: proportional increases in the volume (absolute number) of crimes in a city over time can be quite large, while proportional increases in the crime rate in the same city are not nearly as dramatic. The discrepancy occurs if a city grows significantly in size during the time period. In such cases, crime volume is likely to increase much more than crime rate. Both rate and volume are important and should be considered together in analyzing a city's crime problem.

Even when readers do interpret the findings properly, they should keep in mind several qualities of the UCR data themselves that may affect their interpretation.

- The FBI publishes only crime data for individual cities and towns with populations of 10,000 or more, and even in these cities, data are only presented for the more "serious" (as defined by the FBI) crimes of murder, rape, robbery, aggravated assault, larceny, burglary, and auto theft. Consequently, it is difficult to use UCR data to analyze crime trends in small towns and to analyze patterns of other (non-UCR) offenses that may constitute a large proportion of the crime reported in the state. The FBI does collect data from smaller towns but does not publish them. The FBI does, however, include the small town statistics for serious offenses in its calculation of the overall state crime rates.
- Not all communities do presently participate in the UCR program.
- The FBI data include only offenses reported to the police. Some studies have indicated that the actual amount of crime is greater than that reported to the police in any locality.

NOTES

[1]Clarence M. Kelley, Director, FBI, UNIFORM CRIME REPORTS for the United States. (Washington, D. C.: U. S. Government Printing Office, Printed annually); U. S. Bureau of the Census, "Characteristics of the Population, Part (Number) (State)," Census of Population: 1970, Vol. 1 (Washington, D. C.: U. S. Government Printing Office, 1975).

[2]Michael J. Hindelang, et al., Sourcebook of Criminal Justice Statistics 1973. (Washington, D. C.: U. S. Government Printing Office, August 1973); U. S. Law Enforcement Assistance Administration and U. S. Bureau of the Census, Expenditure and Employment Data for the Criminal Justice System: 1972-73. (Washington, D. C.: U. S. Government Printing Office, 1975).

[3]Gordon, Robert A., "Issues in the Ecological Study of Delinquency," American Sociological Review, XXXII, (December 1967), pp. 927-944.

CHAPTER 5

UCR IN CRIME SPECIFIC RESEARCH AND PLANNING

David L. Dougherty

INTRODUCTION

The usefulness of Uniform Crime Reports is extensive for law enforcement agencies. Information provided by the Report can be applied to areas of statistics and hard data, as well as to aspects of planning and research. This chapter provides examples of using UCR data to help develop crime patterns and trends that are both necessary and useful for planning and research needs.

THE USEFULNESS OF UNIFORM CRIME REPORTS

The Uniform Crime Reports have utility for crime specific research and planning as (1) indicators of specific crime trends calling for further investigation, and (2) data to begin more detailed study.

HOMICIDE AS AN EXAMPLE

An example of a study which might rely heavily on UCR data is an analysis of trends in homicide. Homicides are considered among the most reliably reported crimes in the UCR Program and additional information regarding homicides reported is provided in the attachment to the reports prepared by local law enforcement agencies.

The attachment to a Uniform Crime Report may have less reliability in that it is prepared soon after the crime occurs, usually prior to the completion of investigation of the case; and, although in some cases it is not completed for all reported homicides, enough information is generally provided to direct the researcher toward trends regarding weapons used and other factors. When used in volume, trends both over time and across jurisdictions can be determined.

The UCR reporting form itself is highly reliable with respect to homicide reporting. A simple count of increases or decreases over time and across jurisdictions can easily provide trend data and point to the need for programs to respond to changing trends or for further investigation of factors related to those trends.

The UCR Report and the attachment are especially useful in conjunction with arrest reports or rap sheets to provide an in-depth analysis of trends. When these three sources are available, the simple count of increases or decreases may indicate a period or (if compared across jurisdictions) a location in which homicide is increasing. This would direct the researcher to further analysis of the attachments related to that finding which may provide information on such changing factors as the types of weapons used,

the location and time of the crimes, and, in some cases, the relationship of the victim to the alleged offender. That information can be further expanded by analysis of arrest reports and rap sheets for at least a portion of the homicides of interest. The arrest reports and rap sheets provide additional information about the circumstances of the crimes under study and may show patterns in the nature of the crimes or in the offenders' backgrounds amenable to control through new policies or practices of agencies of the criminal justice system. Recommendations based on a thorough study of trends, factors, and patterns discovered from UCR reports and other sources might include increased police patrols in locations and at times where homicide is increasing, (especially if stranger-to-stranger homicides are involved), more intensive investigation of homicides at the scene and by the coroner, alternative sentencing for offenders whose crime patterns are similar to those who have committed homicides, or new controls on weapons.

UCR AND PART II CRIMES

A more difficult but still valuable use of the UCR program for crime specific research is for Part II crimes such as drug use. Since Part II offense reports are generally based on arrests made, not crimes reported, research with respect to those offenses tends to show more about law enforcement activity than actual offense conditions in the community. Nonetheless, trends can be shown in relation to increasing arrests.

Arrest reports and rap sheets can be especially useful in regard to Part II offense patterns. Much as Part I reports point to trends and patterns calling for further investigation, Part II reports provide similar indicators and when such trends appear to be significant, analysis of arrest reports and rap sheets can show factors which may account for changes and suggest activities to respond to the changes. Such things as circumstances of the crimes (from arrest reports), offender background, and patterns of arrests for other crimes (from rap sheets) show patterns requiring new responses from criminal justice agencies.

It should be noted that, due to the nature of the reporting, Part II data is much less reliably compared over time and across jurisdictions, therefore such comparisons should be made only with full understanding of possible changes in reporting and enforcement practices over the period or between the jurisdictions studied.

CHAPTER 6

PREDICTING CRIME INCIDENCE AND DETERMINING CHANGE

Thomas A. Giacinti

INTRODUCTION

If planned improvements are to be incorporated in the criminal justice system, there is a need to estimate the current extent of crime as well as to forecast future crime rates and apply these estimations and forecasts to a determination of whether or not the actions have had an effect on the rate of crime. The utility of quantitative methods for estimating and forecasting crime incidence and for interpreting observed changes in crime patterns is becoming increasingly apparent.

There are both simple and complex statistical methods that can be employed to predict crime incidence on the basis of previous annual crime statistics. These methods are associated with predicting crime as a function of a straight line extrapolation. This chapter describes methods for performing both visual preliminary estimations and computations of crime rate statistics and forecasts.

LINEAR PREDICTION METHODS

A Visual Estimating Method

The albegraic formula for linear regression ($Y = a + bX$) permits one to determine "Y" (the incidence of crime) from the relative slope of a line "b" or by a technique called the "least squares" method. The simple preliminary method of doing this involves preparing a graphic presentation of the annual frequencies of crime and a "visual estimation" or projection of future crime frequency. This is done by plotting annual frequencies of crime on a calibrated graph, linking the points together to show the historical trend of crime, and then plotting an estimated average increase (or decrease) trend line which extends through the points indicating past annual crime frequencies to a future year.

Customarily, such a graphic display of data is presented utilizing X and Y coordinate axes, with the Y axis being a vertical line on the left, depicting frequency intervals of crime, while the X axis is the graph's horizontal base (abscissa) depicting elapsed time in days, weeks, months, or years.

In Example A (Figure 1), the Y axis denotes the frequency of rape in increments of 15, going up from 0 to 600. The X axis depicts the years for which crime data is available, in this case 1966 through 1974 with the year 1975 added as the year for which a prediction of crime frequency is desired. The steps to construct this graph are few and uncomplicated.

First, draw the X and Y axis for the graph and determine which units of measure will be most appropriate for recording the frequency of crime. In

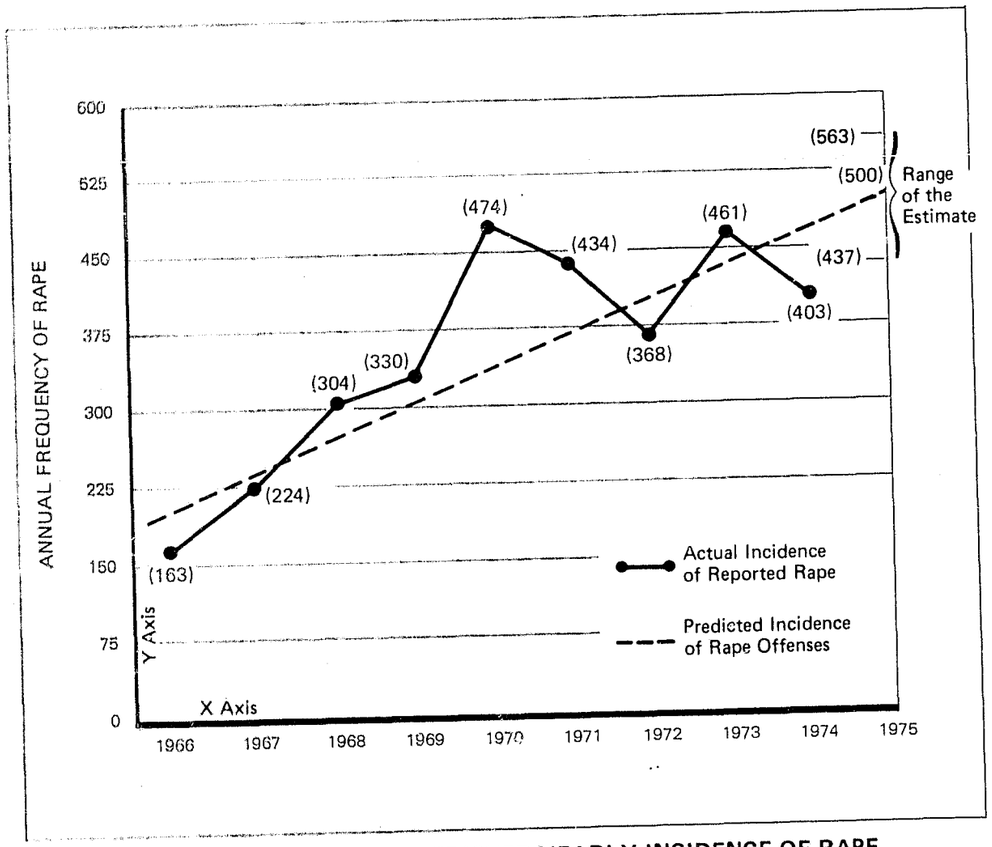


Figure 1. EXAMPLE "A", PREDICTED YEARLY INCIDENCE OF RAPE AS THE FUNCTION OF A STRAIGHT LINE

Example A, the units of measure were 15 crimes on the Y axis and 1 year on the X axis. Generally speaking, the smaller the units of measure are the more accurate the graph will be (in the example the emphasis was on predicting the yearly incidence of crime, so there was no need to present time in smaller units). The next step is to enter a coordinate point within the axes (X and Y coordinate values) which depicts the frequency of crime for each year.

In Example A, the year 1966 resulted in a frequency of 163 reported rapes. The use of calibrated graph paper makes this exercise accurate with minimal effort. After entering a point on the graph depicting the frequency of crime for each year, then link all the points with a solid line. Now the graph portrays a historical trend line of criminal events over time.

To "visually estimate" the predicted frequency of crime for the next year and into the future, another line must be added to the graph. It will be called the "regression line" (broken line in Example A). To construct this regression line place a straight edge or ruler over the trend line which has already been drawn so that one half of all the data points are above the ruler and one half of the data points are below it. Now, simply draw a straight line through the existing historical trend line and extend the line one year into the future so that the line describes what appears to the eye to be the angle or slope that the connected data points follow. The point at which the regression line crosses the next year's location on the graph will be an estimate of the frequency of crime for that year.

In Example A, a regression line (broken line) has been drawn through the historical trend line separating the yearly frequency points so that four points are below the line and five are above it. The regression line has been extended through 1975. The point at which it crosses 1975 is the 500 frequency level on the calibrated scale, and this represents the estimated frequency of rape for 1975 using the simplistic model or "visual estimate" method. This technique is meant to provide only a very rough indication of what one might expect to find. Although an efficient procedure, reliance should not be placed on it when estimates are needed for reporting purposes.

One should assume a "range of tolerance" around this 1975 predicted frequency of 500 approximately equal to the average differences of the annual frequencies for the years plotted, as discussed below. This method is also only a rough approximation for preliminary use in arriving at a "ballpark" estimate.

The average difference is calculated by taking each year's frequency of crime and subtracting from it the previous year's frequency; summing the differences; and dividing by the number of years of data for which the differences were calculated as follows:

	Year	Frequency		Difference
	(1966)	163 - no previous data	=	not applicable
1.	(1967)	224 - 163	=	61
2.	(1968)	304 - 224	=	80
3.	(1969)	330 - 304	=	26
4.	(1970)	474 - 330	=	144
5.	(1971)	434 - 474*	=	40
6.	(1972)	368 - 434*	=	66
7.	(1973)	461 - 368	=	93
8.	(1974)	403 - 461*	=	58
				568 ÷ 8 = 71

In determining the differences* between the frequencies, an absolute difference is calculated regardless of whether it was an increase or a decrease. In other words, ignore the plus or minus signs of the difference when calculating the average yearly difference. The exercise above reveals an average change in the frequency of rape from 1966 to 1974 of 71 events, so the "range of tolerance" for estimating the frequency in the next year (1975) is 500 + 71, or 500 - 71. The range for the visual estimate of rape in 1975 as portrayed in Example A would be 429 to 571. An actual frequency of rape which occurred in 1975 and was above 571 or below 429 would more than likely be a significant increase or decrease.

Algebraic Method

Because its accuracy is dependent upon a fairly consistent linear trend with small fluctuations from year to year, and because this visual estimating method is only a preliminary technique, a more accurate estimate of the frequency and the range of the predicted frequency should use the equation for linear regression and a formula for determining confidence intervals using the standard error of the estimate. The formula for predicting the frequency of crime is:

$$Y = a + bX$$

To actually arrive at the prediction, one must use the computational formula. In the following example we will compute the coefficients "a" and "b" for the descriptive formula above using the data from Example A. To develop the coefficients, it is easiest to compute "b" first because it can be used to compute the coefficient "a." The computational formula for "b" is:

$$b = \frac{N\Sigma XY - (\Sigma X)(\Sigma Y)}{N\Sigma X^2 - (\Sigma X)^2}$$

and the computational formula for "a" is:

$$a = \frac{\Sigma Y - b\Sigma X}{N}$$

The data we have available from Example A is:

X (is ordinal value of the year)	Y (is yearly frequency of rape)
1966 = 1	163
1967 = 2	224
1968 = 3	304
1969 = 4	330
1970 = 5	474
1971 = 6	434
1972 = 7	368
1973 = 8	461
1974 = 9	403

Ordinal values have been assigned to the years 1966 through 1974. These values are 1 through 9 and they are X values. For the Y values the actual frequency of crime in each of the X years is used. As indicated above, in 1966 the X value is 1 and the Y value is 163 rapes. The only other value we need to know is N, which is the total number of years which will be used in the computation. In Example A, the N is 9 years.

To compute "b" using the formula $b = \frac{N\Sigma XY - (\Sigma X)(\Sigma Y)}{N\Sigma X^2 - (\Sigma X)^2}$, $N\Sigma XY$ is the first part of the formula to be computed. To do this, first multiply each value of X times its corresponding value of Y as in the following:

X	Y	XY
1	x 163 =	163
2	x 224 =	448
3	x 304 =	912
4	x 330 =	1320
5	x 474 =	2370
6	x 434 =	2604
7	x 368 =	2576
8	x 461 =	3688
9	x 403 =	3627
		ΣXY = 17,708

All these values must be summed, which is what the symbol sigma (Σ) means. This equals 17,708. This number must then be multiplied by N, which is 9. ($N\Sigma XY = 9 \times 17{,}708$), which equals 159,372.

The next calculation will be to solve the right hand side of the formula. To do this, sum (Σ) all the X scores and sum (Σ) all the Y scores.

X	Y
1	163
2	224
3	304
4	330
5	474
6	434
7	368
8	461
9	403
$(\Sigma X) = 45$	$(\Sigma Y) = 3161$

and multiply these two numbers together, that is, $(\Sigma X)\ (\Sigma Y) = 45 \times 3,161 = 142,245$.

Now the bottom half of the formula for "b" will be computed. Multiply each of the X scores times itself (square them) and sum (Σ) the squared values, and multiply them by N, which is 9.

X				X^2
1	x	1	=	1
2	x	2	=	4
3	x	3	=	9
4	x	4	=	16
5	x	5	=	25
6	x	6	=	36
7	x	7	=	49
8	x	8	=	64
9	x	9	=	81
ΣX^2			=	285

$$N\Sigma X^2 = 9 \times 285 = 2,565$$

To solve the bottom right hand side of the equation, take the previously computed sum of the X's (ΣX), which was 45, and square it:

$$(\Sigma X)^2 = 45 \times 45 = 2,025$$

Now all the component parts of the formula to calculate the "b" coefficient have been computed.

$$N\Sigma XY = 159,372$$
$$(\Sigma X)\ (\Sigma Y) = 142,245$$

$$N\Sigma X^2 = 2,565$$

$$(\Sigma X)^2 = 2,025$$

These values merely need to be plugged into the formula to determine "b."

$$b = \frac{N\Sigma XY - (\Sigma X)\ (\Sigma Y)}{N\Sigma X^2 - (\Sigma X)^2}$$

$$b = \frac{159,372 - 142,245}{2,565 - 2,025}$$

$$b = \frac{17,127}{540}$$

$$b = 31.7$$

The "b" coefficient (31.7) can now be used to compute the "a" coefficient in the formula:

$$a = \frac{\Sigma Y - b\Sigma X}{N}$$

All the component parts of this equation have been computed when the "b" formula was calculated. It is known that:

$$\Sigma Y = 3,161$$
$$\Sigma X = 45$$
$$N = 9$$
$$b = 31.7$$

All that need be done is to plug these numbers into the equation for "a"

$$a = \frac{3,161 - (31.7 \times 45)}{9}$$

$$a = \frac{3,161 - 1,426.5}{9}$$

$$a = \frac{1,734.5}{9}$$

$$a = 192.7$$

We now have the value of the coefficients "a" and "b" to enter into our prediction formula $Y = a + bX$. To predict the 1975 incidence of rape from Example A, enter the values for "a" and "b" in the equation and enter the ordinal value for the predicted year. In Example A, the ordinal value for 1974 was 9 so the next year, 1975, will have the ordinal value of 10. Plugging the values into the linear prediction formula we have

$$Y = 192.7 + (31.7 \times 10)$$

$$Y = 192.7 + 317$$

$$Y = 509.7$$

so that the 1975 predicted frequency of rape using the data from Example A and the algebraic formula for linear regression will be 509.7 or 510 reported crimes.

Because this predicted frequency of rape (509.7) is an extrapolation beyond the data on which the prediction equation was based, it also is a rather rough estimate. However, one may determine the standard error of the estimate and use that information to build a rough accuracy range around the predicted crime frequency. A formula available in most social science statistics texts dealing with linear regression may be adapted for this purpose. It involves using the predicted value of Y (frequency of rape in 1975) and creating an interval of interest using the "t" score for the desired level of confidence (two-tailed) and the estimated standard error of the estimate. The interval around the predicted frequency of the crime of rape in 1975 would be from 601.2 to 418.2.* Any actual frequency which falls above or below this level might represent a meaningful change from the anticipated yearly occurrence of the crime. As can be seen, the range generated using the algebraic formula is close to the range generated by the simple visual method previously mentioned.

- Range for Simplistic Method = 571 to 429 Estimated Frequency = 500
- Range for Algebraic Formula = 601.2 to 418.2 Estimated Frequency = 509.7

Given access to a calculator, the statistical formula can be used to derive the estimated frequency of crime. However, one can be fairly accurate using the simplistic visual graphic procedure.

ADJUSTMENTS OF DATA FOR POPULATION SIZE

Because actual frequencies of crime are somewhat misleading, in using a linear prediction model (either the mathematical formula or simplistic method), it is advisable to adjust the UCR crime data. Depending on the crime category, one can adjust the data to population total, number of automobiles, number of women, number of residences, etc.

In the example used it would have been more appropriate to predict a rate of rape by using the yearly UCR data on the crime as a function of the population of women in the jurisdiction each year in the form of a yearly rate of rape per 100,000 women. This would be a more accurate representation of this crime problem and would also probably affect the size of the predicted "range" by reducing it. Similarly, a situation could arise with the crime of residential burglary where, for example, the frequency of the crime would rise from 15,000 per year to 19,000 per year. In any situation an increase of 4,000 residential burglaries from one year to the next would represent a serious increase in the crime. However, if the frequency of the crime is adjusted to a base of the number of housing units and it can be determined that the number of housing units increased, for example, from 75,000 to 100,000 from one year to the next, the rate of residential burglary per 100,000 housing units would be 20,000 for the first year and 19,000 for the second year.

*The actual calculation for the interval (or accuracy range) appears in the notes to this article. The confidence level selected was the .10, or 90%, confidence level.

LIMITATIONS OF THE LINEAR MODEL

One of the limitations of the linear prediction model is that the further the yearly frequencies (or rates) are from the prediction line, the poorer attempts at extrapolation will be. In other words, the greater the deviations about the regression line are, the greater will be the error of the estimate. In Example B (Figure 2), one can see graphic representation of this. In Figure 2, the accuracy range for the predicted burglary frequency is small because the distance between the yearly frequencies and the regression line is small (i.e., the yearly increases or decreases in the crime do not differ that much from the prediction line). In Figure 3, the range of interest for the predicted robbery frequency is great because the distance from each yearly frequency to the regression line is large. This also holds true using the algebraic formula for least squares method of predicting.

Another limitation of the linear prediction model is that interpolation (i.e., making estimates within the range of data already available) is necessarily more accurate than extrapolation (i.e., going beyond the range of data previously examined). Moreover, the further one attempts to predict into the future (extrapolate), the greater will be the potential for error. This has serious implications for the estimating procedures used. For example, in the algebraic regression approach outlined above, the standard error of estimate deals with interpolated data; accordingly, it will always underestimate the error of measurement associated with extrapolation. For this reason the regression approach outlined in this chapter should be viewed as a "rule-of-thumb" rather than a rigorous statistical procedure.

An additional limitation of the linear methods presented here is that they do not account for a directional change in the trend line. If, for example, as in the case of the rate of automobile thefts, the frequencies stop increasing each year and start decreasing, the utility of either predictor model is temporarily limited. Either method is designed for predicting the straight line function of the occurrence of some event. If, because of some drastic change like the advent of certain successful crime reduction measures (as in the case of the automobile with the innovations of anti-theft devices in the design of the new vehicles), the yearly incidence of crime begins to steadily decrease from year to year, then the linear prediction method becomes temporarily less useful.

If the direction of the line changes, as in Example C (Figure 4), a new starting point must be selected for predicting crime incidence. In the example, the new starting point is the most recent highest frequency year, 1971. The frequency of crime in this year was selected as the beginning point for the computation of the regression (predictor) line. It should be noted that it takes at least three data points (years in this case) showing a directional change to begin to plot a predictor line using either the simplistic method or mathematical formula. In any situation using linear procedure, the more data points there are, the greater the accuracy of the prediction will be, assuming that the basic relationship between predicted and predictor variables remains the same.

One question that arises in using the linear prediction methods described is the utility of the models in determining change over time if the

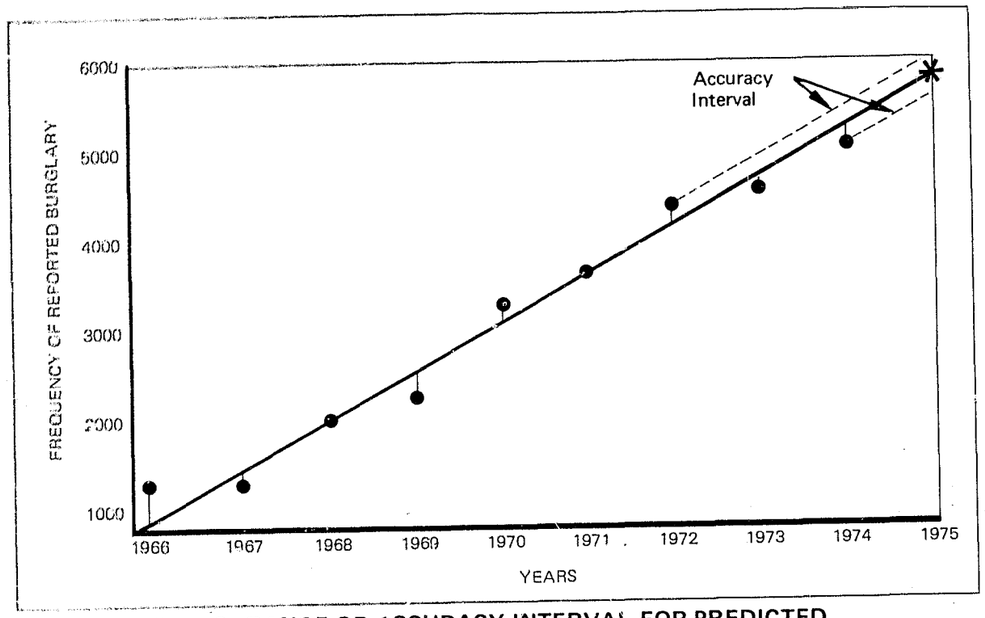


Figure 2. RANGE OF ACCURACY INTERVAL FOR PREDICTED BURGLARY FREQUENCY

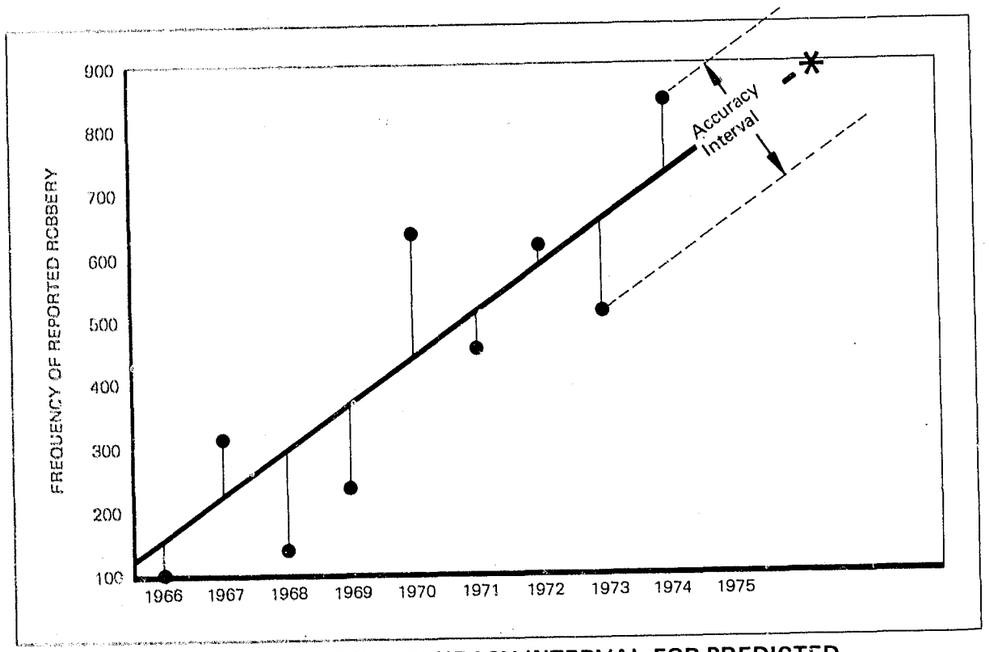


Figure 3. RANGE OF ACCURACY INTERVAL FOR PREDICTED ROBBERY FREQUENCY

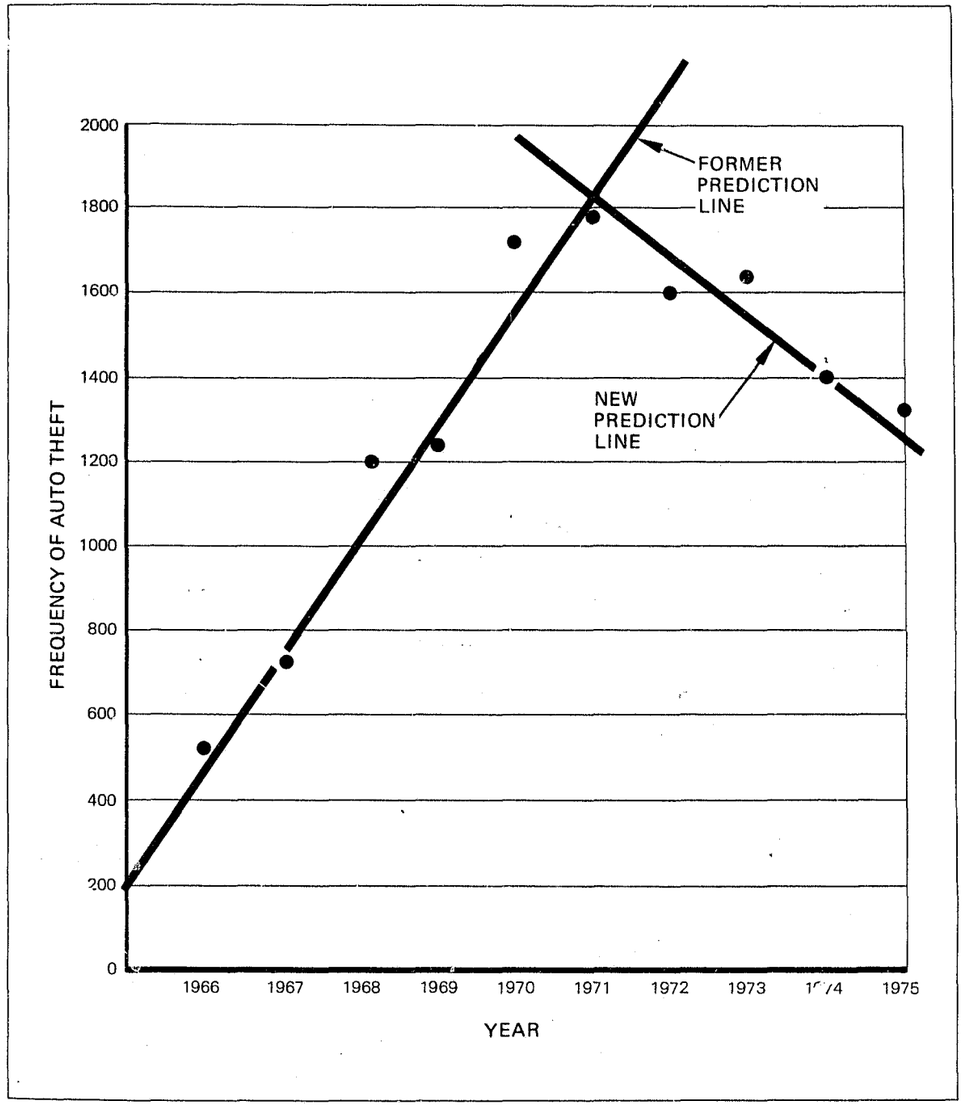


Figure 4. EXAMPLE "C", FREQUENCY OF AUTO THEFT

intervals of interest around the predicted values are large. Certainly, the utility of the prediction seems limited when the range of interest is as large as 183 (601.2 to 418.2), as was developed using the algebraic formula and the data from Example A. In this case, it should be noted that the lowest estimate of the interval of interest (418.2) is only 15.2 greater than the previous year's frequency of the crime (403). Any occurrence that would be equal to or less than the previous year's frequency of the crime would represent a drop in the crime which would be of interest. Wide fluctuations of the reported frequencies over the years make the incidence of the crime less predictable. Adjusting the frequencies to a population of potential victims or targets may help to lower the degree of yearly crime fluctuations.

QUALIFICATIONS OF THE MODELS

There are other mathematical formulae for accounting for directional changes or curvilinear relationships for prediction purposes using multiple regression techniques and second degree algebraic equations. However, because of the complexities of these models they do not warrant coverage here.

One must assume that all yearly reporting of crime frequencies is subject to influence by extraneous variables. Such variables could include changes to statutory definitions of crime, changes in police reporting procedures, and extreme changes in population demographics or the environment of the jurisdiction. These extraneous variables are not usually anticipated in predicting crime incidence. Fortunately, however, in many instances one can go back and adjust the data by accounting for the proportional change incurred by the intrusion of the extraneous variable. A case in point would be the change in the definition of the UCR category of larceny. Prior to 1972, the UCR reports of larceny separated reported crimes where the loss was $50 and over. In recent years, the two categories have been combined to include all larcenies in one category. To adjust previous years' data for prediction of ensuing years, it is necessary either to add the frequencies of larcenies under $50 for each of the previous years or increase the previous years' frequencies of larceny by the proportion of total larceny accounted for in the most recent year's incidence of the crime by larcenies of under $50.

In summary, to adjust crime data for the influence of the extraneous variable, first be aware of the presence of the extraneous variable and second, account for the proportion of the crime incidence influenced by the extraneous variable. No one will be able to accomplish both of these actions in all situations. In any event, any statement about yearly dramatic changes in crime incidence should be preceded by a vigorous investigation of the possible intrusions and effects of an extraneous variable.

NOTES

The formula for computing a confidence interval around a predicted score using the linear prediction model makes use of the predicted score plus or minus the standard error of the estimate of the predicted Y scores at each X point which has been multiplied by the t score for the desired confidence interval, or

$$Yp \pm \hat{\delta}Yp \times t$$

In Example A we computed Yp which was 509.7. To calculate the standard error of Yp we use the formula

$$\hat{\delta}Yp = \hat{\delta}Y/X \sqrt{\frac{1}{N} + \frac{(X - \overline{X})^2}{\sum_{i=1}^{N}(X_i-\overline{X})^2}}$$

To solve for $\hat{\delta}Y/X$ use the formula

$$\hat{\delta}Y/X = \sqrt{\frac{\sum_{i=1}^{N}(Y_i-Y)^2 - b\sum_{i=1}^{N}(X_i-\overline{X})(Y_i-\overline{Y})}{N-2}}$$

Using the data in Example A to compute the confidence interval (prediction range) around the predicted score at the .10 confidence level would involve the following computations:

1. solving the formula for $\hat{\delta}Y/X$

$(Yi - \overline{Y})$			$\sum_{i=1}^{N}(Yi) - \overline{Y})^2$
163 - 351	=	-188	35,344
224 - 351	=	-127	16,129
304 - 351	=	- 47	2,209
330 - 351	=	- 21	441
474 - 351	=	123	15,129
434 - 351	=	83	6,889
366 - 351	=	17	289
461 - 351	=	110	12,100
403 - 351	=	52	2,704
3,161			91,234

$$\overline{Y} = \frac{3,161}{9} = 351$$

$\sum_{i=1}^{N} (X_i - \bar{X})$		$\sum_{i=1}^{N} (X_i - \bar{X})^2$
1 - 5	= -4	16
2 - 5	= -3	9
3 - 5	= -2	4
4 - 5	= -1	1
5 - 5	= 0	1
6 - 5	= 1	
7 - 5	= 2	
8 - 5	= 3	
9 - 5	= 4	
45		60

$\sum_{i=1}^{N} (X_i - \bar{X})$		$(Y_i - \bar{Y})$	
-4x	-	188	= 752
-3x	-	127	= 381
-2x	-	47	= 94
-1x	-	21	= 21
0x		123	= 0
1x		83	= 83
2x		17	= 34
3x		110	= 330
4x		52	= 208
			1903

$$\bar{X} = \frac{45}{9}$$

$$\hat{\delta}Y/X = \sqrt{\frac{\sum_{i=1}^{N} (Y-\bar{Y})^2 - b \sum_{i=1}^{N} (X_i - \bar{X})(Y_i - \bar{Y})}{N - 2}}$$

$$\hat{\delta}Y/X = \sqrt{\frac{91,234 - 31.7 \times 1,903}{9-2}}$$

$$\hat{\delta}Y/X = \sqrt{\frac{91,234 - 60,325}{7}}$$

$$\hat{\delta}Y/X = \sqrt{4,415}$$

$$\hat{\delta}Y/X = 66.4$$

This value is used in the equation for the standard error for the predicted Y (1975 rape frequency)

$$\hat{\delta}Yp = \delta Y/X \sqrt{\frac{1}{N} + \frac{(N-\bar{X})^2}{\sum_{i=1}^{N} (X_i - \bar{X})^2}}$$

$$\hat{\delta}Yp = 66.4 \sqrt{\frac{1}{9} + \frac{(10 - 5)^2}{60}}$$

$$\hat{\delta}Yp = 66.4 \sqrt{.71 + .42}$$

$$\hat{\delta}Yp = 66.4 \sqrt{.33}$$

$$\hat{\delta}Yp = 66.4 \times .728$$

$$\hat{\delta}Yp = 48.3$$

The prediction range or confidence interval is then determined by multiplying the t score for (N-2) degrees of freedom by the standard error of the predicted score and adding and subtracting this from the predicted value of Y. In the computation for Example A, the 90% confidence level was used for t with 7 degrees of freedom (two-tailed test) so that the confidence interval for the predicted 1975 rapes is

$$509.7 \pm 1.895 \times 48.3$$

$$= 509.7 \pm 91.5$$

The prediction range or confidence interval equals

601.2 to 418.2.

REFERENCES

Blalock, Herbert M., Jr. Social Statistics, 3d. ed., McGraw-Hill Company, no date.

CHAPTER 7

A PLANNING-ORIENTED MEASURE OF CRIME AND DELINQUENCY*

Burt Nanus and Luther Perry

INTRODUCTION

A great deal of money is currently being spent on problems of crime and delinquency in our society. Some of the money is allocated to programs designed to improve the general system of criminal justice, but most is earmarked for specific policies or programs aimed at dealing with particular problem areas. In order to determine the ultimate impact of these programs on the overall level of crime and delinquency, one must first have a suitable and generally acceptable yardstick or measure, or as they have recently come to be known,'social indicators.' The purpose of this note is to propose a new measure that would be useful both for analysis and for guiding planning on current and future crime programs in a given region.

THE PROBLEM

Very likely at present the best, most widely used, and most readily accessible social indicator of crime and delinquency is the Uniform Crime Report (UCR) issued by the F.B.I. since the early 1930's. Its principal value is that it has existed for about four decades, is familiar to police officers and the public, and offers at least a broad overview of the general direction of the national crime level. However, there are many well-known deficiencies of the UCR, including its degree of aggregation, its inability to discriminate between various degrees of victimization and differences in reporting approaches over time by various jurisdictions.

Thus, the UCR, while widely quoted, is of limited use as a social indicator of crime activity. Other measures have been proposed including victim surveys, recidivism rates, self-reported crime studies, the 'fear index' of key groups such as shopkeepers, taxi drivers and police. All these measures have major shortcomings, not the least of which is cost and the absence of an historical data base for trend analysis.

Thus, it is apparent that there is a need for new crime measures including the following properties:

1. The measures should give a true reflection of the trend over time of the most serious crimes in a given geographic area. Serious crimes are important because they cause the most harm in terms

*Reprinted with the permission of the Journal of Criminal Justice, I, No. 3 (Fall 1973), pp. 259-263.

of human suffering and contribute the most to a sense of fear and insecurity in the populace. They are also the most likely to be reported.

2. The measures should be capable of being disaggregated as needed, down to very small geographical and time units so as to be useful for detailed analysis as well as projections. In addition, it should be possible to construct historical indices on the same basis for planning and decision-making purposes.
3. The measures should be derived from sources as close to the events as possible. The closest official source is the police since they receive the first report of a crime. There are drawbacks in using only crimes known to the police--variations in reporting rates among types of crimes possible, inaccuracies, etc.--but compared to other data, many police records are reasonably complete and historically uniform. Some records are also available in an automated or semi-automated fashion.
4. The measures should include consideration of the social costs or impacts of the crime, including costs to the victims and to society. This implies a need to weight crime for its seriousness.

One measure that appears to meet some of these criteria was developed by Sellin and Wolfgang and published in a landmark book called The Measurement of Delinquency.[1] These researchers discovered that there was a great deal of agreement among university students, police officers, judges and the general population in Philadelphia on the relative seriousness of different types of crimes.

Having established that it is possible to weight the seriousness of different crimes with consistency, they were then in a position to create a crime index. To calculate the index, one must first calculate a 'seriousness weighting' for each criminal 'event' that took place in the time period of interest. Rather than use the legal label for an offense, the Sellin-Wolfgang Index examines the offense itself for all the serious constituents of the act and assigns a weighting to each.

The Index is compiled by summing all the weighted events that occurred in a given geographic region during a specified period of time. In addition, various ratios can be calculated such as the average number of seriousness units per 10,000 population or per 10,000 juveniles, or the average number of seriousness units per event or per offender.

These measures appear to be very meaningful for evaluation of activities in small geographic areas or neighborhoods, particularly if they can be calculated monthly over a period of time. To explore further, a prototype project was undertaken in Los Angeles using a variation of the Sellin-Wolfgang Index.

THE PROTOTYPE PROJECT

The basic objectives of this experiment were to design a procedure to show trends in criminality in the Los Angeles area, and to design a prototype community crime seriousness compilation procedure that could be used for planning purposes. The compilation procedure had to meet four basic criteria,

1. The cost and manpower needed had to be minimal.
2. The data base had to be uniform over the historical time frame of interest.
3. The data base had to have a good relationship to actual crime.
4. It had to be possible to apply measures of seriousness to the crime data in such a way as to facilitate disaggregation and analysis.

It was discovered that a very useful source of information for social indicators exist in a location that is infrequently tapped for the purpose--namely, the modus operandi or MO files in a police agency that are normally kept to facilitate criminal investigations. There are several reasons why the MO files could be useful for this purpose.

First, there is a high level of detail and completeness. For example, time, to the hour of occurrence and the day of the week, and geographical location, to a few block neighborhood area, are completely specified. Data can be aggregated in any combination on the time and location variables, and can be reaggregated many times to suit various needs. Moreover, the MO files represent most reported crimes, the reporting-abstracting-coding process has been relatively uniform over a long time span, and the files are often automated.

Second, the form of data in the MO files contains sufficient detail on the crime itself to permit the application of weighting procedures similar to the Sellin-Wolfgang Index. Also, the MO files can be cross-referenced to land use and census files kept by various other agencies that have geographical bases very close to the police reporting districts. Time and geographical coordination between different agency data sets is fairly good.

There are three other advantages that should be mentioned. The MO files are based on original first-hand crime reports and are prepared according to a standard policy guide by trained officials--i.e., police officers. The use of police MO files poses no danger to individual privacy since there are no names, addresses, or personal identifiers of any kind in them. Finally, MO files are in common use in many police agencies providing a vast, largely untapped, historical base of information for studying and measuring crime trends.

To test these assumptions, a prototype project to trace trends in the seriousness of crime in the Watts Model Cities Area was undertaken using the Los Angeles Police Department (LAPD) data base from late 1966 through November of 1971. A city-wide total of about 950,000 crimes over the five-year period were available for analysis. An adaptation of the Sellin-Wolfgang Index of Seriousness was applied to the MO descriptors to check the feasibility of computer-based procedures and to obtain an indication of possible local trends.

These adopted weights introduced two weaknesses into the prototype project. First, the Sellin-Wolfgang Index has not been validated for Los Angeles or for the Model Neighborhood. Second, descriptors in the MO files, especially for bodily injury, did not match up directly with

the Sellin-Wolfgang measures. However, the researchers were less interested in the exact numbers than in determining whether it would be possible to create a social indicator that takes particular advantage of the characteristics and data available in a police MO file.

A rather straightforward computer program was written to examine each record in the MO files, assign weights as indicated, and then aggregated the weights for specific time periods or small geographic areas. Comparisons were then made between the Watts Model Cities Area and other areas in Los Angeles. The seriousness was calculated in three forms:

- raw seriousness
- seriousness per 100,000 population
- seriousness per MO card (approximating seriousness per crime)

Figure 1 shows the calculated monthly crime seriousness per 100,000 population in four Los Angeles areas. Regression lines are shown for comparison with the all-city average. The data suggests that the all-city seriousness index did not change radically over the three year period, but that some areas are increasing in seriousness while others are decreasing over time, and that month-to-month variations can be quite significant.

Figure 2 shows a comparison of the regression lines for a single division of monthly seriousness per 100,000 population as compared to the UCR Part I crimes per 100,000. In this particular case, the UCR suggests that crime is increasing in the division while the new measure suggests that seriousness is actually decreasing. While it was not found to be typical for the UCR to be moving one way while the seriousness index moved another, this example does show that it can happen.

In the process of the experiment several problems arose. It was not always possible to formulate a simple relationship between descriptors and weights. Several descriptors sometimes had to be combined logically to determine the appropriate seriousness weighting to apply. Cost was another potential problem. However, aside from setup costs (obtaining and validating the weightings and programming the system), the experiment showed that a procedure can be implemented to give desired crime measures for any neighborhood a few blocks in size, for any time period in the past five years, for the cost of a very short computer run (under $100). Moreover, it would not be a difficult matter to make the calculation of a few specific crime indices an automatic by-product of the normal running of the MO files.

SUMMARY

As this experiment has shown, it is feasible to use MO files to generate an index of crime that appears to be superior to traditional measures in several important respects. Once such an index is calculated, it becomes possible to use it for a variety of purposes, some of which are described below.

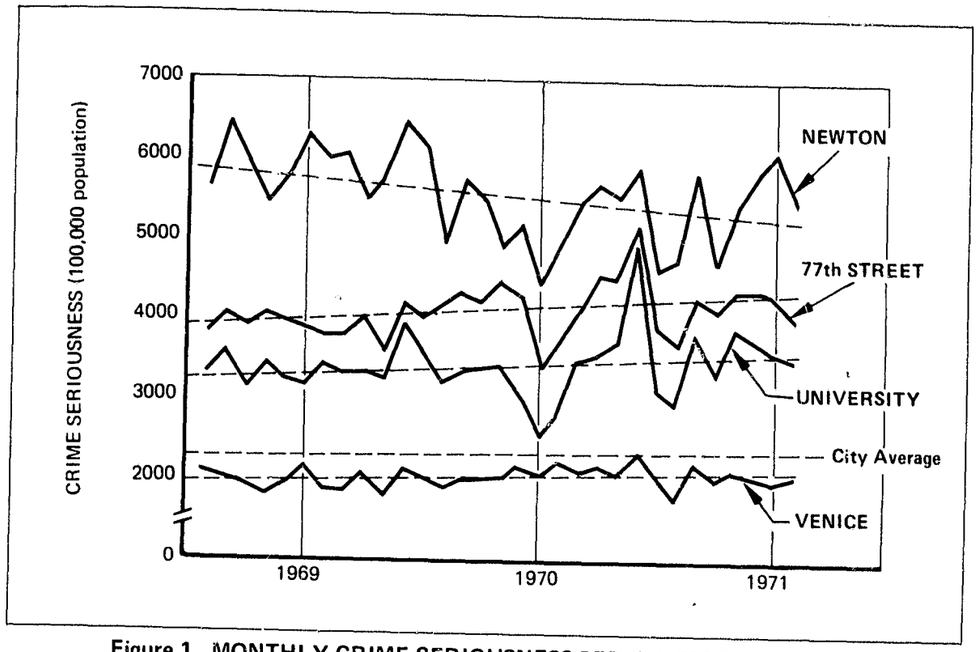


Figure 1. MONTHLY CRIME SERIOUSNESS PER 100,000 POPULATION IN FOUR AREAS OF LOS ANGELES

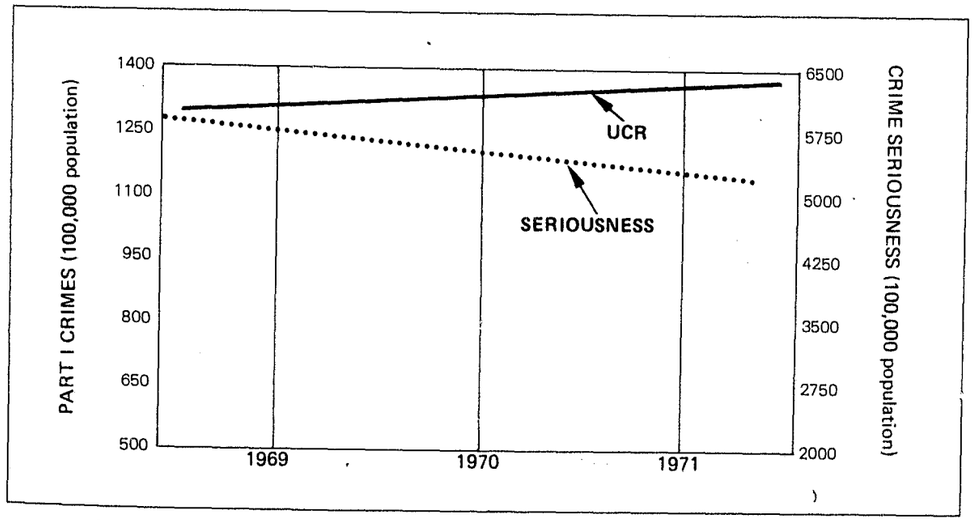


Figure 2. MONTHLY SERIOUSNESS/100,000 COMPARED TO MONTHLY PART I CRIMES/100,000 FOR THE NEWTON DIVISION

1. Policy Evaluation--As long as funds are available for social experimentation in high crime areas, it will be necessary to evaluate the impacts of these experiments. For example, if a social agency introduces a new counseling program for juvenile offenders or a new outreach program for potential offenders in a particular geographic area and desires to study the impact on crime in that particular sector, one could define the geographic boundaries precisely, perhaps down to the few city blocks involved and then calculate a crime index per 1,000 juveniles for the last five years. This index can then be projected for several years in the future to give a baseline for the area to suggest what crime would have been had the program not been implemented. Actual readings can then be taken after the program has been implemented for several years to see how the actual impact has differed from the projected impact. Similarly, a police department may wish to know whether crime in a particular area is getting more serious--that is, whether the crime index per criminal event is increasing or decreasing over time. This would be a simple matter to calculate with the proposed measure.
2. Monitoring Activities--It frequently becomes necessary to determine where in a city the rates of serious crime are increasing and where they are decreasing. The proposed index makes it possible to locate high crime and low crime areas dynamically with some precision and to set up control parameters that will alert officials when special action in a particular area appears warranted. Data can be calculated to observe trends over time, to correct for seasonality and even to analyze the frequency of events on particular times of the day or days of the week in certain areas. Obviously, this would be very helpful for scheduling the deployment of police officers.
3. Planning Decisions--It should be possible with an index of crime such as the one proposed in this report to tie budget allocations and manpower authorizations directly to demonstrated needs in terms of crime trends. The measures lend themselves directly to many operations research type of studies on resource allocation decisions, scheduling, dispatching rules, etc.

The calculation of a seriousness index for crime and delinquency based on a city's MO files appears to be a flexible and powerful methodology for quantifying and analyzing patterns of crime in an urban society. Further work should be done to study the generalizability of these results, the replicability of the Sellin-Wolfgang studies and the various applications to which these measures can be put.

REFERENCES

1 Sellin, Thorsten and Marvin E. Wolfgang (1964). The Measurement of Delinquency. John Wiley and Sons, 605 Third Ave., New York, New York 10016.

CHAPTER 8

ANALYSIS OF CRIME WITH OFFENDER BASED TRANSACTION STATISTICS (OBTS)

Susan Katzenelson

INTRODUCTION

OBTS data allows criminal activities to be analyzed with respect to individual offenders and with regard to offense. Since OBTS information originates at the local level and is aggregated for increasingly large areas, a variety of offender- and offense-based issues may be addressed for the purposes of understanding criminal activities and crime patterns and trends.

Crime is a serious problem of industrialized, urban societies. OBTS is a method based on modern technology that tries to give a faster, more reliable and accurate picture of crime and its treatment by the criminal justice system. Assuming some prior knowledge of OBTS on the part of the reader, this Chapter will discuss the scope of OBTS implementation efforts, and how the data can be utilized for analyzing crime. It will attempt to suggest approaches that can be used with both manual and computer based OBTS data systems.

FORMS OF OBTS IMPLEMENTATION

The concept of OBTS in its original form was developed on a national level under the Project SEARCH* program by a special task force working in coordination with several participating states. Although there is considerable flexibility in the form and scope of OBTS in the states that are implementing OBTS, some important generalizations can be made about OBTS and its application to crime analyses.

Data for an OBTS system are originated by law enforcement agencies: the police, prosecutor's office, courts, and corrections. Each agency that collects data for OBTS is required to check its accuracy and completeness, to conduct periodic updating and auditing of its files, and to guarantee security and privacy.

Data collected for the criminal justice system can serve multiple purposes. As the National Advisory Commission in its report on

*For a more detailed description of OBTS, see "Implementing Statewide Criminal Justice Statistics Systems--A Model and Implementation Environment," Project SEARCH (Sacramento, Calif.: California Crime Technological Research Foundation, January 1972).

Criminal Justice Standards and Goals (1973) pointed out: "Identical data elements should be used to satisfy requirements for similar information to be developed for either an OBTS or the Computerized Criminal History (CCH) system over all areas of the criminal justice system" (Standard 7.1, page 98). Hence, "The collection of data required to satisfy both the OBTS and CCH systems should be gathered from operating criminal justice agencies in a single collection" (Standard 7.2, page 102).

The establishment of an OBTS record is triggered by an arrest, as the first transaction between the adult defendant/offender and a criminal justice agency (in that case, the police). The file is based on the person; it records his step-by-step transactions with the various agencies as his case is processed through and disposed of by the system.

The OBTS system identifies each offender by a finger print-based FBI number, used nationally to track and identify the individual throughout his criminal history irrespective of his present and prior states of arrest or residence. In addition, the states can use other identifying numbers for each OBTS unit.

As suggested by LEAA, the minimal elements of OBTS would include identification data, information on charges, police and prosecutor actions, criminal court actions, and finally, court and correctional dispositions (see Figure 1. in "Analysis of the Criminal Justice System with Offender Based Transaction Statistics").

There is some variation regarding the types of offenses leading to the establishment of an OBTS file. Project SEARCH in its Technical Report No. 4, page 15 (cited previously) suggested to use as the unit of count only adult felony defendants, defined as such by the most serious police charge at the time of arrest. But this method might raise several problems. First, it describes only one form of crime, felonies, and only part of the criminal justice system's functions, that of handling felonies. It does not cover the large number of misdeameanors (which probably constitute over 50 percent of all known crimes) and the large amount of resources spent on their processing. Second, the category of felonies covers different offenses in different states, and such a lack of uniform definition would hinder the exchangeability and comparative value of state-based OBTS records. Third, due to the possibility of reducing original felony charges to misdeameanors, OBTS figures might include a sample of offenses that eventually ended up as misdemeanors, but this sample would probably be biased, not representing the universe of known misdemeanors. Therefore, the suggestion is made to use National Criminal Information Center (NCIC) criterion offenses that include all felonies and most misdemeanors, with the exclusion of traffic offenses and drunkenness. Finally, although there is need to handle juvenile justice data separately, with many more restrictions on access, these cases constitute a very significant part of the crime problem and could eventually be incorporated with the OBTS system of adult crimes. A complete criminal history should include the defendant's juvenile justice record.

The form and level at which the automation of data is accomplished varies by state. Source documents can be sent directly by local agencies to the state. Special OBTS forms can be filled out or coded locally prior to their transfer. Local terminals can send information to a central computer, or records can be completely computerized at the local level. Local transactions can be reported to state judicial or correctional information systems, which in turn make data available for OBTS.

OBTS can be useful even for communities or agencies with a manual system of filing and recording. Their data can be sent to other state centers to be computerized; they can be manually studied and analyzed for special research purposes; or a data base can be built up as a first step toward future automation of a system.

To summarize, the OBTS system would hold information on a state basis on each individual arrested (as an adult) and charged with certain offenses (preferably defined by the uniform NCIC criteria); the step-by-step transactions recording contacts, as he proceeds through the criminal justice system; and over a period of time, the accumulated entries and criminal contacts of these individuals with the criminal justice system.

OBTS DATA ANALYSIS AND APPLICATIONS

In describing criminal activity, OBTS data can be conceptualized in at least two meaningful ways: (1) from the point of view of the individual defendant/offender, and (2) from the point of view of offenses.

From the moment of his first police encounter, a person may go through several steps of relabeling: as suspect, arrestee, defendant, convict, recidivist, etc. These labels only partially correspond with his initial behavior, i.e., whether he is a "criminal," actually having committed an illegal act. Rather, they reflect various decisions made about a person (e.g., the decision to arrest him makes him an arrestee and the decision to find him guilty makes him a convict). Further, even is we assume that only persons that commit a crime are arrested, prosecuted, and convicted, the charges brought against them at each point do not necessarily represent the person's actual criminal actions, especially in jurisdictions where extensive plea bargaining occurs. Accordingly, any analysis based on a system of official records, such as OBTS, will not describe all the actual criminal behavior in a given universe, but only the officially known portion of all criminality that resulted in an arrest and the formal legal charges (changing even while in the process) brought against the defendant.

Keeping in mind these qualifying elements, OBTS will enable the planner or researcher to look at "criminals," their demographic characteristics (age, race, sex, etc.), the types of crimes they are charged with, changes in the formal charges as they relate to the orginal charge, final disposition of the case, and correctional actions, if any, taken against the offender. On a continuous basis, OBTS would permit analysis of the development of criminal "careers," with the reentrance into the system of an individual previously identified. These "recidivists," depending on the purpose of the research, can be defined in terms of rearrest, reprosecution, reconviction, or reincarceration. Time analyses of recidivists would demonstrate existing career patterns; i.e., the "professional criminal" specializing in burglaries or prostitution; the "eclectic addict" sampling various types of property offenses to support his habit; or the "crime-graduate" who starts out with gang membership and minor assaults and gradually grows into armed robbery, rape, or homicide. These various patterns could then be related to personal characteristics and to previous dispositions of recidivists. A division of individuals into first and repeat offenders would give a picture of the proportion of crime, within a given time period or location, committed by recidivists as compared to "first timers."

When used in combination with other data bases, OBTS can aid the analysis of crime in general. In its relation to victimization surveys, OBTS can show what types of offenses tend to remain hidden and what types lead to more arrests, prosecutions, and convictions. By comparing victimization, Uniform Crime Reports (UCR), and OBTS statistics, the relative accuracy, reliability, and patterns of police charges versus prosecutorial charges or convictions can be analyzed.

OBTS data, when accumulated over time would also build the basis for several, scientifically meaningful, comparisons:

- crime trends in the same jurisdiction over time
- similarities and differences in crime among jurisdictions or geographic areas
- evaluation of policies and policy changes within and between jurisdictions or geographic areas
- analysis, on a statewide basis of individual offenders over time with a readily available and more accurate definition of recidivists, career criminals, etc.

The rates at which people are being arrested, charged, prosecuted, convicted, and incarcerated can be computed, using as the denominator the general population, personal characteristics of defendants, all the defendants in the previous step of processing, all defendants charged with that particular offense, etc. Each of these rates would give a different view of crime and its handling in a given community.

For example, by comparing the rate of

$$\frac{\text{all defendants incarcerated for rape}}{\text{all defendants arrested for rape}}$$

with the rate of

$$\frac{\text{all defendants incarcerated for armed robbery}}{\text{all defendants arrested for armed robbery,}}$$

one could study attrition during criminal justice processing, strength of evidence, conviction, degree of punitiveness, etc., in these two violent crime categories.

By computing rates of all defendants prosecuted, based on arrest; all defendants convicted, based on prosecution; all defendants incarcerated, based on conviction, etc.--one could make inferences about the way defendants drop out of the system. One could also use this method to compare jurisdictions as to their respective system rates and different points of dropping out. Additional systems rates giving a more general picture of crime could be computed by comparing victimization and UCR data (crimes committed and crimes reported) with OBTS figures. Such rates could illuminate problem spots in the system and suggest possible changes in policy and practice.

The wealth of data in an OBTS system lends itself to various descriptive and causal analyses. Some of these can be illustrated from PROMIS* data, that in many ways is similar to a local OBTS system.

Descriptive statistics could start with simple univariate methods, such as frequency distributions, measures of central tendency (mean, median, mode), and dispersion. As an example, one could see the number and relative frequency of various offense types in a given time period (Table 1) or the rate of offenders charged with these offenses.

---

*For a description of PROMIS, see "Analysis of the Criminal Justice System with the Prosecutors Management Information System (PROMIS)."

TABLE 1

ARRESTS IN 1973 BY OFFENSE TYPE OF
MOST SERIOUS CHARGE AGAINST THE DEFENDANT

	Number	Percent
I. Crimes Involving A Victim		
A. Personal Victimizations Involving Violence	5040	32.6
1) Homicide	259	
a) Murder	200	
b) Manslaughter	49	
c) Negligent homicide	10	
2) Assault	2891	
a) Aggravated	2002	
b) Simple	684	
c) Assault on a police officer	205	
3) Forcible Sex Offenses	450	
a) Female victim 16 and over	357	
b) Victim under 16	72	
c) Male victim	21	
4) Robbery	1440	
a) Armed	726	
b) Other	714	
B. Personal Victimizations Without Violence	1898	12.3
1) Larceny	1337	
2) Auto theft	372	
3) Fraud	189	
C. Crimes Against Residences or Households	1370	8.9
1) Burglary	1174	
2) Property destruction	164	
3) Arson	32	
D. Crimes Against Businesses or Institutions	2099	13.6
1) Robbery	217	
2) Burglary	372	
3) Larceny	1059	
4) Embezzlement and Fraud	305	
5) Auto theft	74	
6) Arson	8	
7) Property destruction	64	
II. Crimes Without an Identifiable Victim	4757	30.8
A. Weapons Offenses	1042	
1) Gun	827	
2) Other weapon	215	
B. Gambling	372	
C. Consensual Sex Offenses	836	
D. Drug Offenses	1872	
E. Bail Violations and Prison Breach	635	
III. Crimes Which Could Not be Classified	296	1.9
All Cases	15,460	100.0

Two and three dimensional contingency tables can be used to explore the relationship between selected variables; such as case disposition, by race and sex of the defendant (Table 2).

TABLE 2

PERCENT DISTRIBUTION OF OFFENSES BY TYPE OF DISPOSITION
AND OFFENDER'S RACE AND SEX

Offender's Sex & Race / Disposition	MALE			FEMALE		
	White	Black	Total %	White	Black	Total %
Open	8.	9.	8.9	8.9	9.4	9.4
No Paper	28.7	23.1	23.7	19.9	19.1	19.2
Dropped after papering	38.7	32.6	33.1	35.	37.5	37.1
Not Guilty	2.4	4.4	4.2	6.5	4.7	5.
Guilty (plea or other)	21.	30.1	29.3	28.3	28.9	28.8
Other	1.2	.7	.8	1.5	.4	.6
Total Percent (No. cases)	100 (1195)	100 (11405)	100 (12601)	100 (403)	100 (2011)	100 (2414)

Finally, more sophisticated multivariate methods, such as regression analysis, can be applied to describe the determinants of selected dependent variables. For example, one can study the strength and direction of the impact that several demographic and crime-related variables have on a case being accepted for prosecution.

USE FOR PLANNING

The first step in planning should be a description of the phenomenon of crime in order to obtain a knowledge of the scope, form, and type of the problems. The descriptive statistics derived from OBTS can fulfill this need by providing an aggregate picture of known crimes in a given area, or special cuts and segments of it, depending on the need and questions asked by the planner.

By deriving a demographic and criminal profile of offenders, learning about crime patterns, identifying certain offender groups (such as violent, professional, career, etc.), and exploring the effects of various dispositions and correctional actions on rehabilitation/recidivism, a planning agency will be able to define the problems to be attacked. It will be in a better position to determine priorities and to assess programmatic alternatives and project proposals. It can then make the crucial decisions about allocation of resources and manpower, need for cooperation or changes, etc. Further, by viewing the criminal justice agencies as a system that processes the individual offender from arrest to final release, the necessity of synchronized planning becomes evident. The success and effectiveness of any decision by a single agency might be severely limited if that decision is not accepted and consistently followed up by the policy and action of other, related agencies. For example, the police might make a conscious effort to apprehend and press charges against serious recidivists, but these arrests will have little impact if the district attorney's office does not have the resources for swift and thorough preparation and prosecution of these cases, if the courts cannot mete out speedy and evenhanded dispositions to offenders, and if corrections departments lack suitable programs recommended for rehabilitation.

The ability of OBTS to focus on the individual offender and follow his transactions step-by-step will facilitate "system" decisions that are aimed at improved interagency cooperation. OBTS would aid in a concerted planning effort by the whole criminal justice system to deal with crimes and eventually reduce their volume and seriousness.

USES FOR EVALUATION

Once an area under consideration has been accurately described, its crime problems identified, and plans have been made to deal with them, the next step would be to set up a program to evaluate the effectiveness of the plans implemented. Although the definition of effectiveness originates from various social and political value judgements, it often can be measured by objective, quantified, statistical methods.

For example, one value standard in our system is the administration of justice in an evenhanded way. An OBTS system can offer several empirical ways of defining and measuring evenhandedness. For example, this could

be done by comparing dispositions for various offender types, while controlling for other relevant variables, like offense type, seriousness, and prior convictions. It could also be accomplished through looking at variations in pretrial release, by race, age or sex, again controlling for legally relevant variables.

Sound empirical findings could help the criminal justice system to evaluate its existing practices and policies, as well as to measure the effectiveness of changes, and thereby serve as a useful tool in planning and decision-making.

Part III.

THE CRIMINAL JUSTICE SYSTEM: DATA, MODELS, AND TECHNIQUES FOR ANALYSIS

CHAPTER 9

ANALYSIS OF THE CRIMINAL JUSTICE SYSTEM WITH OFFENDER BASED TRANSACTION STATISTICS (OBTS)

Susan Katzenelson

INTRODUCTION

The Offender Based Transaction Statistics (OBTS) system can be used in numerous ways to understand, interpret, and evaluate criminal trends and their relationship to the criminal justice system. This Chapter suggests the types of analyses that could be used within an agency or for comparison purposes across agencies. These analyses may range from studies of single variables to more complex analyses in which given outcomes may be related to several variables that characterize the criminal justice system. The general structure of an OBTS file is presented in Figure 1.

As described previously (see Part I, Chapter 7) OBTS is a state-wide system based on individual files originated by an arrest. Each file includes transactions reflecting contacts of the person arrested with various agencies of the criminal justice system up to the point of exit from the system. Each file contains the relevant interaction between the defendant and the agency: the most serious charge brought against him, disposition of the charge, and correctional action. The OBTS structure suggested in Figure 1 is only a skeletal model with the minimum necessary information. Other items can be added according to the needs and interests of the jurisdiction adopting it.

GENERAL APPLICATIONS

Whereas the earlier section focused on the use of OBTS in analyzing patterns of crime, the purpose of this section is to show how OBTS can be used to analyze the working of the criminal justice system, including measures of its workload, productivity, and evenhandedness.

One of the major problems, often cited by practitioners and researchers alike, is that the criminal justice system, despite its name, does not function as a well-coordinated "system." Policymakers, operations personnel, planners, and researchers tend to view the system in a fragmentary way, dealing with each agency as a separate, independent entity. OBTS is especially well equipped to deal with this problem, because it lends itself to a "systems" view of the administration of justice. At the receiving end of the system, a process is initiated with the "input" of an arrest for a criminal incident that was cleared by charging an individual with committing the crime. At the other end of the process is the "output" of the criminal justice system, the final action taken against the arrested individual related to that particular criminal incident. Based on personal identifiers (FBI

OFFENDER BASED TRANSACTION STATISTICS (OBTS)

(OBTS)

IDENTIFICATION ELEMENTS

State Identification No.
FBI No.
State Record No.
Sex
Race

Date of Birth

POLICE/PROSECUTOR ELEMENTS

Arresting Agency No.

Sequence Letter
Date of Arrest

Charged Offense (Most Serious)

Police Disposition
Prosecutor Disposition
Police/Prosecutor Disposition Date

LOWER CRIMINAL COURT ELEMENTS

Court Identification No.

Initial Appearance Date

Disposition Date

Charged Offense (Most Serious)

(OBTS)

Lower Court Disposition
Release Action
Release Action Date
Final Charge (Most Serious)
Type of Charge
Plea (at Trial)
Type of Trial
Date of Sentence
Type of Sentence
Confinement Term (Days)
Probation Term (Months)
Type of Counsel

COUNTY PROSECUTOR GRAND JURY ELEMENTS

Prosecutor Identification No.
Date of Filing
Type of Filing
Filing Procedure
Date of Arraignment
Charged Offense (Most Serious)
Initial Plea
Release Action
Release Action Date

FELONY TRIAL ELEMENTS

Court Identification No.
Trial Date
Trial Type
Final Plea
Trial Ending/Disposition Date
Final Charge (Most Serious)
Type of Charge
Court Disposition
Sentence Date
Sentence Type
Confinement – Prison (Years)
Confinement – Jail (Days)
Probation (Months)
Type of Counsel

CORRECTIONS ELEMENTS

Agency Identifier
Receiving Agency
Date Received
Status
Date of Exit
Exit

Figure 1. GENERAL STRUCTURE OF AN OBTS FILE

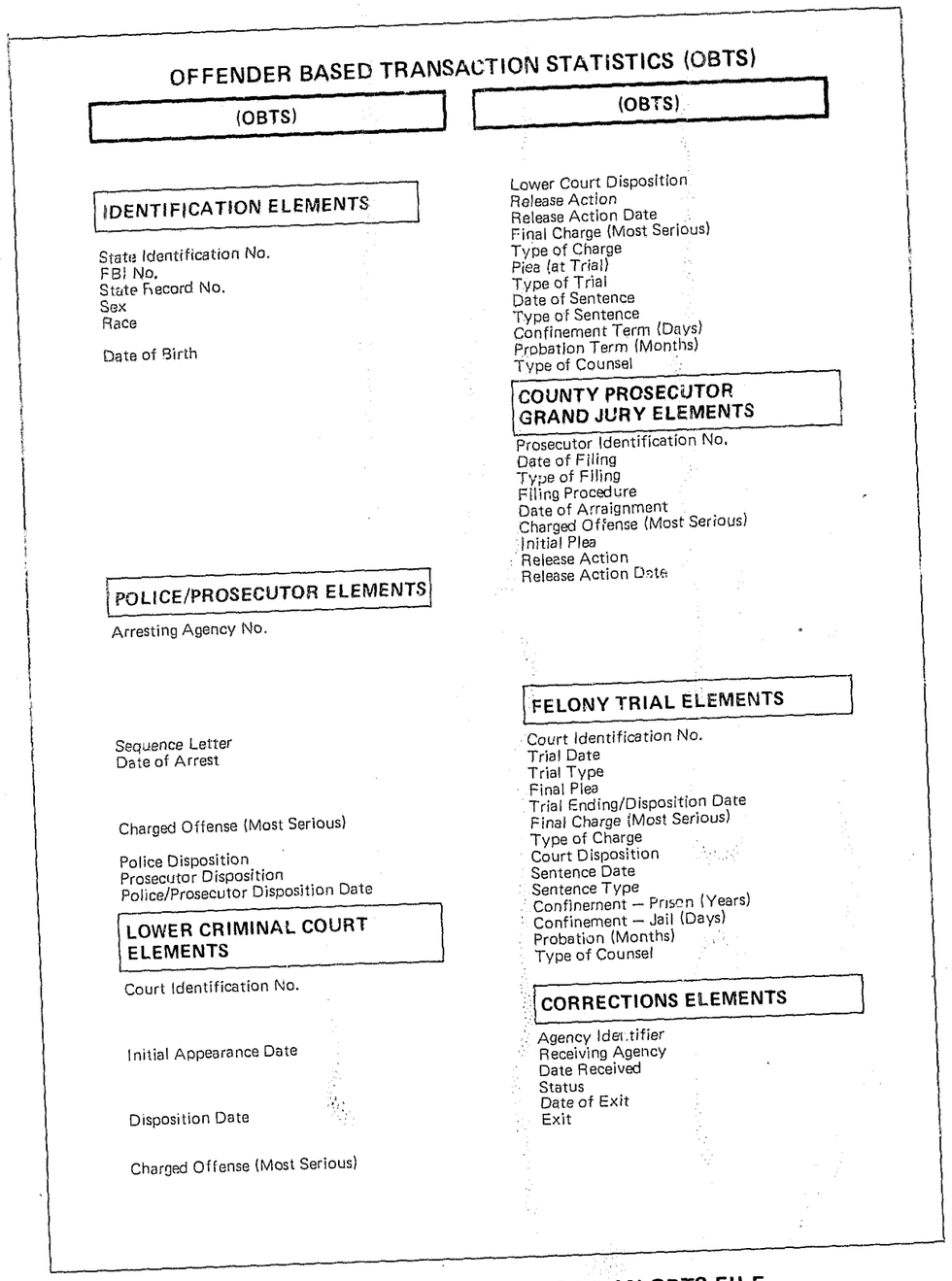


Figure 1. GENERAL STRUCTURE OF AN OBTS FILE

number and fingerprints), additional cycles of interaction can be added to the individual's file with each new reentry to the system.

The processing in the OBTS system is a dual one: that of the defendant/offender, and that of the case. Both can go through several changes. The case can be redefined at each step by the number, type, and seriousness of the charges comprising the case. Similarly, the person can go through a sequence of relabeling from suspect to defendant, to convict, to inmate. The definition of the case and the label attached to the individual will also depend on the varying length of the process. The case can be dropped anywhere before or after filing or prosecution. It can result in a plea bargain, a trial conviction or nonguilty finding, a sentence of probation or incarceration, and any number of alternatives inbetween.

OBTS enables the planner or researcher to supplement the conventional intra-agency analysis with significant interagency observations. For example, one can look at police operations in single cases within a given location (precinct, city, state, etc.) or time period (day, month, year). In addition, one can also follow up on cases initiated and decisions made by the police as they affect the prosecutor's office, the court, and corrections. Individual cases can be traced from their intake at the police department through their handling and disposition.

Although in many states OBTS will be computer based, the feasibility of manual OBTS has been demonstrated in several cities. In any case, the establishment of a good manual system prior to automation is an advisable step.

USES OF OBTS IN ANALYSIS

The considerable amount of data collected by an OBTS system can be analyzed in many ways, from simple descriptive statistics to more complicated measures of correlation and causality.*

Using unidimensional methods (analysis of one variable at a time), questions about the number of cases handled by each agency, the relative distribution of manhours at each phase of the system, the mean number of convictions per month, and similar measures can be computed.

Of special use can be system flow rates, such as:

$$\frac{\text{Number of individuals prosecuted}}{\text{Number of individuals arrested}}$$

$$\frac{\text{Number of individuals convicted}}{\text{Number of individuals prosecuted.}} \text{ or}$$

$$\frac{\text{Number of individuals incarcerated}}{\text{Number of individuals convicted.}}$$

*For an illustration of some of these methods of analysis, see Part II, Chapter 8.

These and similar indices can serve as performance measures of criminal justice agencies, and can point out the "weak spots" with the highest attrition rates (cases dropped prior to trial) in the system. Furthermore, they can help the planner assess the relative "productivity" of one agency compared to others.

Two and three dimensional techniques (analyses of two or three variables at a time) can be used to explore the relationship between selected variables of interest, such as the relative role of guilty pleas as related to the prosecutor's workload, or rates of conviction by length of time in the system for felonies and misdemeanors.

Finally, to more fully understand causal relationships, one can utilize multivariate analyses. For example, the multiple regression method can be used to study the relative impact of system variables--such as actions by the police, the prosecutor, the court, and corrections--on the rearrest of an individual, while controlling for other variables.

Several questions about the characteristics of the system can be conceptualized and answered, using OBTS data:

- What is the direction of change in most serious charges, from arrest to disposition? What is the the relationship between police charges and those pressed by the District Attorney? In what way would plea bargaining be reflected in the charges? In what types of cases are the reductions most significant, etc.?
- What are the most common release types? In what way do release status and amount of bond relate to charges and defendant types? What changes occur in release status throughout the process?
- What is the relative frequency of guilty pleas, jury trials, and bench trials, and how do they relate to offense and offender characteristics?
- What are the conviction rates by offense type, by trial type, and by defendant characteristics? What is the gap between original and final charges, and what are the most typical forms of reduction in charges?
- What is the relationship between sentences and offense and offender characteristics, plea or trial convictions, offenders freed on bail or detained?
- What are the practices of correctional agencies in granting parole, training and treatment programs, types of probation, work release and halfway houses?

- How does recidivism (defined as rearrest, or reconviction) relate to actions taken against that individual as recorded in previous OBTS cycles: Was he convicted in his previous case(s); incarcerated, treated, etc.?
- What is the average (median) time lapse between steps in the process for all cases, misdemeanors or felonies, violent or property offenses, etc.? What part of the system is more time consuming than others?
- How does the total number of cases for a given period, and the resulting caseload, affect the quantity of cases prosecuted, plead, and convicted; and the types of dispositions?
- What is the "attrition rate" of cases, at what phase? What types of crimes are most likely to drop out? (e.g., dismissals). How are these factors related to caseloads at each agency?
- Is there a clear and consistent interagency policy to put more resources in certain types of cases: for example, more police work, special prosecution, and speedier trial directed into the handling of violent crimes, recidivists, major misdemeanors, etc.?

Performance measures of the various law enforcement agencies can be computed, viewed both as absolute figures, and as relative system rates, evaluating the performance of each agency on the basis of input from the previous agencies (i.e., conviction rates based on arrest rates and on prosecution rates). Then, in turn, the effect of that agency's performance on subsequent actions can be reviewed.

OBTS figures can be useful in comparing various system statistics on crimes with surveys on victimization and reported offenses. Such analyses could point to the mortality rates of various offense types before they reach the phase of arrest, to the ways in which the charges brought at each level depart from the actual criminal act as reported by victims. They can also help to estimate which agencies give a more accurate and representative sampling of the universe of crimes--arrest, prosecution, or conviction.

Some of the decision areas and pressing problems of the criminal justice system are clearly implied by these and similar analyses, as will be pointed out in a few examples. Further, according to the special needs of each jurisdiction, additional questions can be analyzed by OBTS data directly geared to deal with policy decisions and trouble spots in the system.

For example, a community might experience a serious jump in the rate of armed robberies. First, with the use of OBTS, it can analyze the existing practices in handling robbery cases. The statistical analysis would give a demographic profile of people arrested on a robbery charge; would describe the rate at which they are prosecuted, plead, and convicted; and would give the final charge for which they are found guilty, the types of sentences, etc. Linked to management and administrative data, it would also compute the relative amount of resources or manpower spent on dealing with robbery cases, as compared to other crimes.

Based on the results of each analysis, decisions about changes can be reached. For example, experienced prosecutors could be assigned to handle the armed robbery cases or a special effort could be made by the police and the prosecutor's office to ensure victim and witness cooperation and a reduction of the number of continuances in these cases. The prosecutors and court might aim at reducing plea bargains and try to convict the offenders, or at least the recidivistic ones, on the original charges. Finally, special programs could be available at the correctional level to rehabilitate these habitual offenders.

As a third step, following the diagnosis of the problem and the implementation of some new decisions to deal with it, the effectiveness of the changes should be evaluated by comparing the before/after picture of the incidence and handling of robbery cases with the use of OBTS data.

Using another example, the pattern of bail decisions could be reviewed using information from OBTS files. With the aid of descriptive and analytic statistics, the possible lack of evenhandedness in granting bail could be discovered. While bail should relate mainly to the defendant's likelihood of showing up for trial, OBTS data might show, for example, that race might act as a determinant of setting and making bail. Further, OBTS data might indicate that pretrial detention increases the defendant's chances of conviction (controlling for the type of case and strength of evidence), probably due to his limited ability to prepare a defense. Once the problem is identified, some decisions can be made to correct the system--like changing release conditions or improved defense services to defendants, etc. As a final step, the success of these changes can be followed up after a given time interval to determine from OBTS if indeed bail decisions are made in a more evenhanded way and without biasing subsequent dispositions.

Another problem often encountered by law enforcement agencies is evaluation of the effectiveness of various programs: the question of "what works." One important measure of success applied by OBTS over time is the rate of recidivism. Recidivism can be defined both as numbers of rearrests and reconvictions, and as type and seriousness of these rearrests. Comparisons of rededivism rates for offenders in diversion programs with those prosecuted, for defendants found guilty with those acquitted, or for convicts sentenced to incarceration with those on probation, could illuminate

the impact of these actions on a person's return to the criminal justice system. Findings from comparisons such as these may suggest changes in strategies that would reduce the rate of recidivism.

OBTS systems, when implemented at the local level, could supply important comparative information to municipal and state planners. The PROMIS system *, first applied in Washington, D.C., is an example of such a working system. Similar to OBTS, PROMIS enables the District's law enforcement agencies to analyze and improve the criminal justice system. Such a system measures the performance of the various agencies at the local level and creates a basis for comparison between municipalities and counties within the state.

With the adoption of OBTS systems by various states, several comparisons can be conducted between jurisdictions at the state level; between urban versus rural areas; between separate agencies, like the courts, in a number of states, or the degree of cooperation among law enforcement agencies in those states; respective rates of recidivism, conviction, incarceration; or other variables of interest.

The success of novel legislation and policy changes in one jurisdiction can be better monitored by comparing its OBTS figures with that of other states. And finally, crime trends, policy changes, and variations in underlying social values and ideologies can be discovered by analyzing OBTS data accumulated over time.

---

* See: "Analysis of the Criminal Justice System with Prosecutors Management Information System"

CHAPTER 10

ANALYSIS OF THE CRIMINAL JUSTICE SYSTEM WITH THE PROSECUTORS MANAGEMENT INFORMATION SYSTEM (PROMIS)

Sidney Brounstein and William Hamilton

INTRODUCTION

A primary deterrent to the consideration of rational and systematic change in the criminal justice system is the absence of factual, empirical data on the performance of the system. Virtually every major national inquiry into the health of the criminal justice system from the Wickersham Commission in 1931 to the National Advisory Commission on Criminal Justice Standards and Goals in 1973 has decried the lack of comprehensive, empirical data that are necessary to elevate the system. The Law Enforcement Assistance Administration took a major step to redress this data deficiency by encouraging all states to develop Offender Based Transaction Statistics Systems (OBTS). In this Chapter, a system called PROMIS, which contains much of the data conceived for inclusion in an OBTS, will be described and some examples of its use will be presented.

THE SYSTEM

Since January 1971 the District of Columbia has had in operation the PROMIS system which has tracked every adult offender prosecuted or considered for prosecution in its Superior Court (the equivalent of a state court of general jurisdiction). The system not only tracks the primary or most serious criminal charge in each case as the OBTS system will do, but also tracks all related charges in each case. Very importantly, PROMIS also contains the reasons for every prosecutorial discretionary decision from intake and screening to final disposition in analyzing the criminal justice system.

Essentially, PROMIS permits a prosecutor's office to accumulate a wealth of information on each case and to receive reports and analyses based on these data so that prosecutors can identify and concentrate on priority areas and exert positive and productive control over their workload, instead of merely reacting to it on a best-guess basis. Not only does this promote effective utilization of prosecutorial time and personnel, but it also serves to attract and retain experienced attorneys. And when local government is aware that existing prosecutorial resources are not wasted by being subjected to operational inefficiencies, requests for additional staff may be looked upon in a more favorable light.

Although PROMIS is basically a tool for day-to-day operational management of the prosecutor's office, it has been designed to provide extensive data for planning and research purposes. It contains relevant information on the workings and decisions of the court process from arrest through final

disposition. The perspective of the criminal justice system provided by the prosecutor is particularly relevant because:

- The prosecutor controls the intake to the court system, deciding what cases brought by the police or other complainants will be "filed" with the criminal court and what cases will be handled and disposed of through other means such as diversionary programs;

- The prosecutor determines how the matters presented to the criminal court are to be "packaged," deciding, for example, whether to file lesser included charges, whether to include all possible counts of a particular charge or just a representative number, and whether to initially join codefendant cases together into a single triable unit;

- The prosecutor can promote expeditious dispositions through the use of plea negotiation strategies and tactics;

- The prosecutor can accelerate case processing times by varying the court processes in certain instances through such means as priority handling of cases of "repeat offenders" or presenting a case directly to the grand jury instead of proceeding through the initial presentment and preliminary hearing.

- The prosecutor can record actions taken in court proceedings, such as decisions on motions or sentencing.

PROMIS contains practically all the data needed for an OBTS from arrest through sentencing. It should be supplemented by other data sources in analyzing the criminal justice system. Some data will also be needed from police, court, and correctional systems to fully meet those requirements. The major point of this Chapter is to demonstrate that PROMIS does provide a powerful nucleus for a local government-based OBTS. For example, its design includes unique identifiers for each defendant, criminal event, court case, and principals in the court case (attorneys, witnesses, police officers, judge). Further information on the data base and its use will be presented later. Before discussing the data base and its applications in analyzing the criminal justice system, it is appropriate to consider data collection alternatives available to state planning agency staff personnel and others interested in such analyses.

Data Collection Alternatives

In developing comprehensive plans, conducting criminal justice systems research, or in related statistical studies, analysts could attempt to collect data from various sources on an ad hoc basis. They sometimes issue contracts to university or private consultants to engage in special data analyses. Such ad hoc data collection and analysis efforts have been found to be expedient in the absence of ongoing operationally based criminal justice information systems. One purpose of this Chapter is to describe, by example, how and why operationally based information systems should provide the data needed for criminal justice systems analysis and planning. Even if computer based systems are not available, manual versions of

operationally based systems are preferable to special data collection and reporting systems as a source of OBTS data.

Operationally Based Systems Versus Reporting Systems or Special Data Collections

The common course of action for criminal justice planners and researchers has been to devise a special data collection effort to develop the information they need for studying a particular problem. Such efforts are always limited by the time and cost constraints of the study. These data collection efforts rarely result in data that are as accurate as those collected as part of an ongoing operation. They tend to be limited with respect to the time frame that can be covered. Later, in the analysis, if any gaps occur or changes in the data collection design appear desirable, it is often impractical to modify the data base, and the research objectives may need to be compromised.

Some planning and evaluation projects acquire data bases from external sources to study a problem that is apparently reflected in the data. As with the special data collection efforts discussed above, the analysis must be constrained by the content of the data base, both as to accuracy and comprehensiveness. The planners generally are forced to accept the data at face value.

Operationally based systems on the other hand, are those that support day-to-day operations. Any planning or research data provided by such systems are by-products of the operational functions. If a planning process is conducted in close coordination with an operationally based system that is being maintained and modified when necessary as users obtain experience with it, it can obtain the following advantages over special or one-time data collection efforts:

- improved accuracy, since data being used in day-to-day operations are subjected to validity checking by automated systems, manual data control personnel, and the users themselves;

- familiarity with the data base on the part of the information systems personnel and the operational "users" so that nonobvious problems in the data and true meanings can be determined;

- ability to structure and maintain longitudinal data following individual persons or cases over time (perhaps several months or years); and

- feasibility of modifying the information system to meet data needs newly determined by planners and researchers.

The data needs of any research or planning program cannot be totally met by any existing operationally based information system because those needs cannot be totally predefined in building the system. The nature of research and planning is such that new problems are always emerging for analysis, and little may be known about the factors that might explain the variation under scrutiny. Until the variables are properly defined, it may may be necessary to collect a large amount of data about each individual "observation."

Further, a sample of "observations" is all that is necessary for planning purposes rather than the processing of all observations as in the case of most operational systems. For example, information on all criminal court cases is essential in an operational system for the criminal court, but only a representative sample of cases is needed to develop statistics for most planning purposes. These are some of the arguments often given to justify special data collection efforts by planners, but the aforementioned advantages of linking planning and research with ongoing operationally based systems are too significant to ignore. The remainder of this Chapter will demonstrate, by example, why this is so. To appreciate the utility of a system like PROMIS for analyzing the criminal justice system, it is important to first develop an understanding of its utility to the prosecutor in solving his daily management problems.

APPLICATIONS OF PROMIS

Struggling to keep pace with a massive influx of cases, hardpressed prosecutors often work assembly-line fashion: each is responsible for cases at a given stage in the proceedings--at screening, arraignment, or trial. No one is in overall control of a case from start to finish. Responsibility and control are fragmented. Cases are lost through cracks in the system: files are misplaced; witnesses fail to appear; numerous continuances result in court dismissals. The habitual, court-wise criminal buries his recidivism in the anonymity of large-scale, assembly-line case processing. He seeks one delay and postponement after another until the government's witnesses are so exasperated and inconvenienced, or their memories of the crime so obscured, that charges are either dropped or dismissed. If the case goes to trial, the prosecutor is often oblivious that there are other cases pending against the accused, or that he is a fugitive, or that the seriousness of his current offense warranted special pretrial preparation of the case.

In addition to the problem of seasoned career criminals who seek to manipulate the system to their advantage, there are internal managerial and operational problems. For example, police officers, expert witnesses, and defense attorneys are scheduled to appear at the same time in different courts on different cases, with the court too often unaware that the conflicts exist until the day of trial. With massive and constantly shifting calendars, case principals are not notified of expected court appearances or of changes and cancellations. Analyses of evidence by chemists, handwriting experts, and other specialists are frequently unavailable on the trial date because of the difficulty of scheduling, coordinating, and monitoring the completion of those activities for a large volume of cases. Furthermore, the chief prosecutor is often in the dark about whether results within the prosecutor's office are caused by subordinates adhering to his policy or departing from it. For example, since the reasons for discretionary decisions by screening assistants are not recorded, the chief prosecutor is unable to determine if refusals to prosecute are consistent with, or contrary to, office policy. Often having only a matter of minutes to review a case before presenting it, a trial prosecutor discovers that essential case documentation is missing or that notations by attorneys who processed the case at previous stages of the proceedings are not clear. He is not sure if essential witnesses are present nor is he reasonably certain

about what aspects of the case their testimony will address. He is also in the dark regarding possible problems with the case--such as those relating to search and seizure or identification--and he is unaware of prior defense-requested continuances and thus is not in a position to determine if another such defense request represents an abuse.

Having been put to the test for several years by the prosecutor's office in Washington, D.C., PROMIS has proven that it can effectively address problems such as those described above and significantly upgrade the performance of urban prosecution agencies.

KEY MANAGERIAL GOALS THAT PROMIS HELPS PROSECUTORS ACHIEVE

The first managerial objective that PROMIS assists with is in helping prosecuting attorneys identify the more serious cases from among the thousands that flow through the Washington prosecutor's office each year. For example, out of dozens of pending assault cases, some may involve defendants who are career criminals and who inflicted serious injury; other assault cases may involve first offenders who are perhaps guilty of only technical violations of the law. PROMIS cuts through these difficulties by assigning ratings to cases on the basis of data obtained by attorneys and arresting officers at the screening stage. This information pertains (1) to the gravity of the crime in terms of the amount of personal injury, property damage or loss, and intimidation involved and (2) to the seriousness of the accused's criminal history based on prior arrests and convictions, aliases used, and the like.

Several days prior to the trial date assigned to a group of cases by the court, PROMIS prints out a copy of the court's calendar for that date but instead of listing the cases in the order the court will call them (oldest first, alphabetical order, in ascending order by docket number), PROMIS ranks them in descending order of their seriousness according to the gravity of the prior record of the accused and the gravity of the crime. High-ranked cases are assigned to a special team of attorneys, called the Special Litigation (Major Violators) Unit, which assures that such cases receive a superior degree of pretrial preparation. When a specially prepared case is called by the court, the team delivers a detailed case work-up to the courtroom prosecutor. In Washington, the conviction rate for cases receiving this special preparation was found to be 25 percent higher than that for those processed routinely.

A second major managerial objective PROMIS addresses is the control and/or elimination of scheduling and logistical impediments that tend to block adjudication of cases on their merits. To this end, PROMIS automatically produces subpoenas, witness and victim telephone lists, a pending-case list for any given witness, and notices for expert witnesses so that all parties concerned can be routinely informed of scheduled appearance dates. PROMIS keeps track of postponements of individual cases and notes, along with the reasons thereof, whether the prosecution, defense, or court is responsible. In addition, PROMIS:

- automatically alerts the prosecutor when the accused has other cases pending against him;

- regularly produces lists of fugitives so that the cognizant law enforcement agencies can systematically seek to apprehend them;

- routinely prints lists of cases pending at various stages of prosecution for more than a specified number of days so that problems of delay can be resolved promptly.

A third principal management objective PROMIS focuses on is the monitoring and enforcing of evenhandedness and consistency in the exercise of prosecutorial discretion. The chief prosecutor in a large, urban agency must inevitably exercise his broad discretionary authority through many assistant prosecutors. His problem, of course, is to assure that the discretion exercised by assistants reflects the consistent and evenhanded application of office policies. Such policies could govern discretion in these areas:

- The decision not to prosecute
- The decision to upgrade, reduce, add to, or subtract from the charges recommended by the arresting officers
- The negotiation and acceptance of pleas
- The decision to allow defendants entry into diversion programs
- The decision to nolle prosequi or dismiss a case
- The initiation, or concurrence in, case postponements.

To monitor and enforce the proper application of discretion in these areas, the related decisions must be visible; that is, they must be recorded and retrievable for subsequent review. Not only must the nature of the discretionary action be recorded (e.g., case rejected for prosecution) but also the reason the action was taken (e.g., case rejected because of illegal search and seizure). Only when reasons for discretionary decisions are known can supervisory prosecutors be in a position to determine whether subordinates' discretionary decisions reflect compliance with office policy. This can be accomplished with PROMIS, because it can generate statistics on the reasons for several different types of prosecutorial actions, ranging from modification of police charges to requests for continuances.

Finally, the fourth major management objective addressed by PROMIS pertains to the analysis of and research on, problems associated with the screening and prosecution of criminal cases. Many of the questions that PROMIS is probing for the prosecutor's office in Washington, D.C., relate to, or interface with, police and correctional concerns as well. The utilization of PROMIS data for planning and research purposes will be discussed at length later.

PROMIS DATA BASE

PROMIS was developed by the United States Attorney's Office for the District of Columbia to encourage new management techniques for the processing of cases.[1] Detailed data are accessible from the PROMIS file so that

the prosecutor may be fully aware of the status of a case. Managers (senior prosecutors) may also use this information to evaluate the performance, demands, needs, and priorities of their sections. In order to aid the prosecutors in their routine evaluations of case priority, the computer assigns scores to both the gravity of the crime and the defendant's criminal history.[2] The object of these scores is to regulate the decisions of prosecutors so that a consistent set of standards is applied in the evaluation of cases. The data base that provides the unique source of information about the criminal court process has been, and continues to be, generated as a by-product of PROMIS. The PROMIS data base contains approximately 50,000 closed criminal cases including about 170 different types of information on each case including:

- Information about the defendant: This includes name, alias, sex, race, date of birth, address, facts about prior arrests and convictions, employment status, and alcohol or drug abuse

- Information about the crime: The date, time and place of the crime, the number of persons involved in the crime, and information about the gravity of the crime in terms of the amount and degree of personal injury, property damage or loss, and intimidation--using the seriousness scale developed by the criminologists, Marvin Wolfgang and Thorsten Sellin[3]

- Information about the arrest: The date, time, and place of the arrest; the type of arrest; and the identity of the arresting officer

- Information about criminal charges: The charges originally placed by the police against the arrestee, the charges actually filed in court against the defendant and the reasons for changes in the charges by the prosecutor, the penal statute for the charge, the F.B.I. Uniform Crime Report Code for the charge, and the Project SEARCH Code for the charge

- Information about court events: The dates of every court event in a case from arraignment through motion hearing, continuance hearing, and final disposition to sentencing; the names of the principals involved in each event including the defense and prosecution attorneys and judge; the outcomes of the events; and the reasons for the outcomes

- Information about victims and other witnesses: The names and addresses of all witnesses, the prosecutor's assessment of whether the witness is essential to the case or not, any indications of reluctance to testify on the part of the witness, and other witness characteristics, such as whether he is related to the victim or defendant.

To make longitudinal analysis work, PROMIS employs a unique offender identification number assigned by the police and based on fingerprints and photographic checks. The unique ID is carried on every case record, following the offender through court processing in PROMIS, and even later in the corrections system. Whenever the offender is rearrested, that same ID is used. This enables tracking the offender over time, identifying and

measuring recidivism, and analyzing relationships with recidivism. The unique offender identification numbers are used in merging historical files of closed cases. Similarly unique numbers are assigned to each court case and criminal event. The criminal event number is used to tie together codefendants and all matters related to a specific criminal incident. Thus, this code is useful in preventing the artificial inflation of crime and court caseload statistics that arise from assigning multiple charges and docket numbers pertaining to a single incident.

PROMIS DATA INPUT

About 80 percent of the data contained in PROMIS is "captured" at the intake and screening stage as the byproduct of the case documentation process. Carbon copies of various forms completed immediately before or during the case screening stage serve as input documents for PROMIS. As a case moves through the subsequent proceedings, additional information about its status is fed to PROMIS. This is achieved through turnaround documents—forms generated by PROMIS in advance of a court event—on which the results of a given proceeding (e.g., preliminary hearing, sentencing, etc.) are recorded and then entered in PROMIS.

Because of the extent of its acquired data, PROMIS dovetails very closely with the informational requirements of other criminal justice agencies and information systems. For example, PROMIS contains the vast majority of data needed by the Computerized Criminal[4] Histories and Offender-Based Transaction Statistics components of LEAA's Comprehensive Data Systems program. And, of course, the numerous facts stored in PROMIS explain why it can contribute so significantly to the achievement of the four major prosecutorial management objectives discussed earlier and why it is capable of generating the many types of reports described below.

REPORTS PRODUCED BY PROMIS

As implemented in Washington, D.C., PROMIS generates on a recurring basis five categories of reports that are of particular value to supervisory and trial prosecutors: misdemeanor calendars, felony calendars, case status reports, workload reports, and special reports. Many of these reports are of assistance to police as well.

Misdemeanor Calendars

These calendars assist management and trial prosecutors to process, in a timely and orderly manner, serious misdemeanor cases scheduled for trial or sentencing. The calendars inform management of the case workload, aid in the preparation of case documentation for trial, and identify cases warranting special pretrial preparation. There are five types of misdemeanor calendars:

1. Five-Day Misdemeanor Calendar. Listing all misdemeanor cases scheduled for trial five days hence, this calendar contains such information as the defendant's and codefendant's name(s), case number, charges, arrest

date, witnesses and their addresses and phone numbers, arresting police officers, number and dates of continuances and reasons for them, prosecutor's and defense attorney's names, and ratings reflecting the gravity of the crime and of the defendant's criminal history. (See Figure 1.)

2. One-Day Misdemeanor Calendar. One day in advance of the trial date, PROMIS produces a calendar similar in content and format to that in Figure 1.

3. Five-Day Misdemeanor Priority Calendar. This advance calendar ranks upcoming cases according to to their seriousness as determined by their PROMIS-computed case rating. Thus the most important cases are identified in a timely manner and receive special pretrial preparation by a team of attorneys called the Special Litigation (Major Violators) Unit.

4. One-Day Misdemeanor Priority Calendar. The purpose, content, and format of this calendar is similar to Number 3 above.

5. Misdemeanor Sentencing Calendar. Issued one day prior to the date of sentencing, this calendar alerts prosecutors so they can prepare documentation for sentencing recommendations. (See Figure 2 for the content and format of this calendar.)

Felony Calendars

PROMIS generates four types of felony calendars: (1) One-Day Preliminary Hearing Calendar, (2) Five-Day Felony Trial Calendar, (3) One-Day Felony Trial Calendar, and (4) One-Day Felony Sentencing Calendar. The purpose, content, and format of the first three felony calendars are similar to the misdemeanor calendar illustrated in Figure 1; the felony sentencing calendar corresponds to Figure 2.

Workload Reports

The following reports allow management to monitor the progress of cases that are specially assigned to individual prosecutors for preparation and trial:

1. Misdemeanor Specially Assigned Cases. This report lists, by attorney, all misdemeanor cases assigned to the Special Litigation (Major Violators) Unit, which gives intensive pretrial preparation to the most important cases. With this report, the chief of the Unit can assign cases on an equitable basis. The report contains the information for each case handled by a given prosecutor: judge, docket number, defendant's name, defense counsel, indictment dates, charges, witness data, and other pending cases against the defendant.

2. Felony Specially Assigned Cases. This PROMIS-generated workload report provides each prosecutor with a listing of all pending felony cases to which he is assigned. The report contains the same type of information as noted for Number 1 above.

DEFENDANT'S NAME	CASE NO.	J/NJ	ARREST	TYP	APO/WITNESSES	NC	DATES	CONTINUANCES	PROSECUTION
1 JXXXXXXX, JXXXXX	03864173	NJ	11/05/73	AF	HXXXXXXX, DXXX R	01	06/30/73	C-CONTINUED TO TRIAL	CXXXXXX, R
					BADGE 0874 UNIT 2D				PENDING
DEFEND: 17.? CRIME 04	CHARGES:		ADW-GUN	BE	RXXXXXXXXX, C				01358673 FEL
RELEASE: PERSONAL RECOG			UNA-MJ		BADGE 2999 UNIT 2D				
DEFENSE: MXXXXXXX,JXXX			CDW-GUN						
CO: RXX, IXXX CXXXX									
2 MXXXXXXX, AXXXXX	04864173	J	11/15/73	AE	SXXX, JXXX R	05	06/30/73	C-CONTINUED TO TRIAL	KXXX, W
					BADGE 0248 UNIT 3D		07/15/73	D-DEFENSE COUNSEL LATE	PENDING
DEFEND: 15.0 CRIME 02	CHARGES:		ADW-GUN	2E	MXXXXXXX, HXXXXX		08/01/73	G-CW NO SHOW	02568573 MIS
RELEASE: CASE BOND			ROBBERY		202-3682598 301-2681594		09/15/73	G-CW UNAVAILABLE	
DEFENSE: HXXXXXXX, AXXX							10/15/73	G-CW NO SHOW	
CO: NONE									

Figure 1. PROMIS FIVE-DAY MISDEMEANOR CALENDAR FOR CASES CONTINUED TO 12/04/73

JUDGE: SMITH

DEFENDANT'S NAME	CASE NO.	INDICTED	DISP DATE	PROSECUTOR CHARGES	FINAL DISPOSITION	PENDING CASE NO	TYP
JXXXXXXX, DXXXXX L	01234573	12/01/73	12/20/73	HXXXXXXX,P BURGLARY II	GUILTY-JURY	02235673	FEL
DEFEND: 10.0 CRIME:10							
PXXXXXXX, DXXXXX L	01356273	12/02/73	12/22/73	MXXXXX, J ROBBERY	PLEA GUILTY THIS CHG	03572673	MIS
DEFEND: 20.0 CRIME:02							
TXXXX, JXXX T JR	02233473	11/05/73	11/30/73	PXXXXX,T ADW-GUN	FOUND GUILTY JURY		
DEFEND: 21.0 CRIME:10							

Figure 2. PROMIS MISDEMEANOR SENTENCING CALENDAR FOR CASES CONTINUED TO 12/12/73

3. Attorney Felony Case Workload by Type of Charge. Listing the types of felony charges by number and as a percentage of total caseload for each prosecutor, the report can assist the manager of a felony trial division in assigning new cases to prosecutors.

Case Status Reports

In addition to calendar and workload reports, PROMIS generates four different summaries on the status of cases pending at various prosecution stages. A series of three reports lists defendants who have had bench warrants issued against them for failure to appear before the court as directed:

1. Misdemeanor Fugitive List. Noting all misdemeanor cases involving outstanding bench warrants for defendants, the report furnishes to police information contained in PROMIS about any given fugitive-defendant.

2. Pre-Indictment Felony Fugitive Listing. The cases of the fugitive-defendants in this listing are pending between screening and the grand jury stage.

3. Post-Indictment Felony Fugitive Listing. This report lists indicted fugitive-defendants whose cases are awaiting felony trial.

These three fugitive listings contain the following information: defendant's name; police fingerprint-based identification number; court docket number; police criminal complaint number; date the bench warrant was issued; the name of the judge issuing the warrant; charges; and defendant's race, sex, date and place of birth, and home address.

4. The fourth case status report lists all cases pending in the grand jury, oldest cases appearing first, and contains the following data for each defendant: case number, defendant's name and police identification number, arrest date and date bound over to the grand jury, release status, defense counsel, arresting officer, charges, and other pending cases.

All of these case status reports are of obvious importance to the prosecutor inasmuch as the cases remain at a standstill until the grand jury acts or the defendant is apprehended.

Special Reports

Comprising this fifth major category of PROMIS reports are Witness Subpoenas, Subpoena Summary Listing, and the Statistical Report.

Special preprinted subpoena forms are generated to advise witnesses where and when to appear for a scheduled misdemeanor trial upon entry into PROMIS of a new or modified trial date. If PROMIS detects an error in the name or address of the witness (e.g., no street number) or if there is insufficient time for the subpoena to reach the witness by mail, this will be noted on a Subpoena Summary Listing. Thus alerted, the Washington, D.C., prosecutor's Witness Notification Unit, staffed primarily by paralegals,

telephones or otherwise contacts witnesses to whom subpoenas have not been issued.

The Statistical Report provides management with an overview of the case workload for a given period with regard to:

- Screening– Misdemeanor and felony cases considered, charged, rejected, or reduced;
- Preliminary Hearings – Cases bound over, dismissed, aborted through a nolle prosequi action, or reduced;
- Grand Jury– Cases indicted, ignored by the grand jury, dismissed, or referred to misdemeanor prosecution;
- Dispositions– Dispositions are separated into four groups– felony trials, unassigned misdemeanors, Special Litigation (Major Violators) and total misdemeanors. Within each of these four groups, the total cases are separated into types of outcomes (e.g., guilty, not guilty, dismissed, aborted through nolle prosequi);
- Delays– Average delays for cases disposed between arrest and indictment, between indictment and disposition for felonies, and between arrest and disposition for misdemeanors;
- Pending Cases– The number of pending cases at various points in the prosecutorial process;
- Fugitives– The number of bench warrants issued and quashed during the period and the number of pending fugitives at the end of the period.

THE PROMIS RESEARCH PROJECT

Recognizing the value of the PROMIS data base, the National Institute of Law Enforcement and Criminal Justice[4] funded a research grant to conduct analyses in six topics that are of national interest. Before characterizing the data needs formulated for this research project, the six topics will be outlined. These are:

- Prosecution Performance. The evaluation of prosecution performance is going beyond gross conviction rate statistics to examine such factors as the rates and associated reasons for declinations to prosecute and other court or prosecution terminations of prosecutions prior to trial. Also, patterns of plea negotiations will be examined so that conviction statistics are conditioned by the nature and gravity of the crime charged and the crime pled to, as well as the nature and gravity of the defendant's prior criminal record. Finally the performance of a public prosecutor will be scrutinized from the standpoint of consistency and evenhandness of discretionary decisions.

- Police/Prosecutor Relations. Traditional indicators of police system performance, such as clearance by arrest rates are inadequate. By extending the measures of police performance (e.g., arrests surviving first judicial review), the quality of police arrests and other measures of police performance can be improved. The project will attempt to learn about the relationships between police and prosecutor policies and decisions, and the performance measures.
- Patterns of Criminal and Related Community Behavior. Expanding upon the work of other investigators, the relationship of criminal behavior patterns to criminal justice agencies is being explored. The effect of situational variables (place of occurrence, familiar distance between adversaries, availability of weapons, etc.) on the decisions of police, prosecutors and jurists will be measured. Likewise, the geographic distribution of victimization, crimes reported and crimes prosecuted will be analyzed in studying patterns of crime, behavior of offenders, and citizens.
- Plea Bargaining. Plea bargaining maintains its position in the criminal justice process despite the clearly recognized jeopardy in which it places the rights of the accused. This is largely a result of the fact that plea bargaining is viewed as a necessary expedient. Before plea bargaining is excised from the system of justice, empirical evidence on its functioning will be examined. It is important to know whether it is generally employed in an evenhanded manner and whether the bargains are generally reasonable and fair. It is also necessary to assess the likely effects of curtailing or modifying the practice of plea bargaining.
- Judicial Decision Making. The evaluation of judicial decision making in large, urban court systems has generally concentrated on individual sentencing records. To obtain a true picture of judicial decision making, the evaluation will address composite empirical data on all major judicial functions including the administration of continuances, the administration of pretrial hearings, trials, and sentencing.
- Speedy Trial: Court Delay and Defendant Flow. The National Advisory Commission on Criminal Justice Standards and Goals has recommended maximum limits of 60 and 30 days respectively for felony and misdemeanor delay between arrest and trial. However, before arbitrary limits are imposed on court delay, it would be well to examine empirical data on the causes of delay, on the nature and extent of delay, and on the effects that delay has on the rights of the accused and on the public safety of the community (as expressed in the relationship of recidivism for example).

Within these major topics, some examples of special analyses being undertaken are: evaluations of diversion programs, recidivism, gun cases, and the female offender.

PROMIS Research Methodology

Before characterizing the data needs of the PROMIS Research project, it will be helpful to outline briefly the methodology being employed. A primary task of the research project is applying multivariate statistical analysis techniques to a research data base extracted on a random sample basis from the historical files (closed cases) of PROMIS. The statistical analysis is concerned with testing theoretical models hypothesized to explain causal relationships between a dependent variable (say an effectiveness measure such as conviction rate) and a set of independent variables, such as case and offender attributes, time in queue for courtrooms, and case processing actions. Some of the independent variables may represent discretionary actions (policy variables) on the part of the criminal justice agencies, and the idea is to determine what changes these agencies can make to improve effectiveness.

A simulation model of the court system is being built to incorporate relationships and test changes suggested by the multivariate analysis. The model generates cases with attributes in accordance with patterns detected in an empirical analysis of the closed cases for a base year. "Cases" and offenders flow through a network of court proceedings and decision nodes. Equations developed through the multivariate analysis will be used to set "branching ratios," i.e., the proportions of cases flowing out of a decision node along the alternate outcome paths. For example, based on the case attributes, the model will simulate the proportions of offenders being released and detained in jail as a result of bail decisions. Other examples of branching points are (1) cases postponed for various reasons, and cases going to trial; and (2) trial acquittals and guilty verdicts. Court resources will be considered. Monte Carlo techniques will be used to analyze queues and statistics will be gathered to measure system performance.

Changes in policies and procedures on the part of police, prosecutor, and court components that appear to offer potential for improving criminal justice effectiveness will then be field tested under actual operating conditions using controlled experimental designs. Evaluations of results will be made in coordination with a criminal justice coordinating committee.

PROMIS Research Data Needs

The data needed by the PROMIS research project may be characterized as follows:

1. longitudinal analyses (tracking over several points in time)
   - case processing events from arrest through various court proceedings and final court disposition
   - offender based tracking including all pending cases against the same defendant, as well as the ability to track closed cases and prior criminal history of the same defendant.
2. cross sectional analysis (single time frame)
   - offender, victim, witness attributes and relationships

   - criminal event attributes and seriousness scores
   - geographic distributions

SOME EXAMPLES OF PROMIS STATISTICAL ANALYSES

Creation of Variables for Research Purposes

In research studies, one often attempts to analyze relationships among variables. It may be necessary to construct a meaningful variable from several data elements, depending on the nature of the study. An example is provided by the witness cooperation study, recently conducted by the U.S. Attorney's Office in Washington, D.C.[5]

Under PROMIS, the reasons that a case is rejected for prosecution or that a trial is delayed (by reason of continuance) to another time are listed by the prosecutor. Among these reasons for negative action in a case are several that involve witness cooperation in the prosecution of a case. For example, there are five witness-related reasons why a case may be dismissed at its initial screening by an Assistant U.S. Attorney. These are: the complaining witness signs off in writing (the complaining witness does not want to prosecute), the complaining witness does not show at the screening (and does not notify the officer as to why), the complaining witness refuses to come to the screening, the witness is not consistent in his story (raising the issue of credibility), and the witness garbles his story (seems to be confused). In sum, these reasons may be defined as composing a "witness problem area" for the prosecutor. These reasons may be defined as indicators of witness "noncooperation" at this stage in the case prosecution. Now it becomes possible to argue that witness behavior may be defined with respect to noncooperation. If we assume no (or few) cases are prosecuted in which the witness does not cooperate at the screening, we have created a variable. That is, we have created a quantitative or characteristic description of some phenomenon that possesses more than one value or category. A variable allows for measurement.

The ability to create variables that make conceptual sense also is a condition necessary for the researcher to conduct an analysis of a given phenomenon. Using the variable created above--witness cooperation--the analytical question of why some people cooperate and why other people refuse to cooperate in the prosecution of a case can be asked. With the development of a management information system we get the information needed to make operational definitions which in turn allow for the creation of variables. These give us a basis on which we can clearly conceptualize and analytically measure the problem of witness noncooperation.

Once variables are created, however, we must be sure that we are in fact measuring what we claim to be measuring. Continuing with our example of witness cooperation, this point is easily made. So far we have considered cooperation only <u>from the point of view of the prosecutor</u>, not the witness. From the point of view of the witness, cooperation may be perceived quite differently. For example, in the study of witnesses to common law crimes in the District of Columbia, it was found that in a field

interview several months after the fact, 78 percent of those labeled "noncooperative" by the prosecutor realized that they were witnesses, and of this group 84 percent stated that they were willing to serve as witnesses.

Research Strategies

An interesting example of how research methodology can be used in information system design occurred when consideration was being given to redesigning a form for use in PROMIS. In forms design, one is often faced with evaluating the cost of asking more questions (in the extreme, the respondent may refuse to fill out too long a form) against the value of the information provided by the questions. One way of estimating the "value" of the information is to determine whether the variables being measured by the answers to the questions under consideration have any effect on the outcome of the cases.

The new form under consideration was called the PROMIS Worksheet. It collects data used to compute "scores" for the seriousness of the crime and the recidivistic tendencies of the defendant, both of which may be subsequently used by the prosecutor as a basis for resource allocation. The form collects data on offender, victim, and witness relationships. In particular, it was redesigned to capture more data on witness characteristics in the light of research findings in a previous study of witness cooperation.

In the process of redesigning the PROMIS Worksheet, one of the issues that arose was whether a detailed set of questions should be asked of the second most essential lay witness (in addition to asking these questions of the primary victim and most essential lay witness). While these questions may produce useful information, each questionnaire item that is completed adds to the processing time and cost per case. In addition, it is possible that as the number of data elements grows, so will the recording error and omission rates. In other words, would it be better to ask fewer questions of more witnesses or many questions of fewer witnesses?

In view of the several types of costs in obtaining information, it was deemed desirable to try to determine what is gained with each question that is asked. For example, does the fact that the most essential witness has a physical or health problem affect the rate at which property crimes are successfully prosecuted? Is it useful even to have a new most essential lay witness, much less ask questions of him? Issues of this sort can be addressed by applying appropriate statistical methods to the PROMIS data base.

A first attempt at building a model of the value of witnesses to the prosecutor was the following: The probability of prosecutor success, for a given type of case (e.g., stolen vehicle, aggravated assault), is a priori related to the number of witnesses in the case. A measure of prosecutor "success" was the conviction rate achieved in the cases studied; i.e., the proportion of guilty pleas and guilty verdicts to total cases accepted for prosecution. It is necessary to introduce "control" variables into such a model; i.e., variables that are entered to ensure that we do not mistakenly attribute to a particular independent variable X, systematic effects upon

the dependent variable Y, that are in fact attributable to other variables that are also determinants of X.

After structuring sufficiently controlled regression models using alternatively as the dependent variable "no papering" rate (cases filed by the police, but rejected at intake by the prosecutor) and "conviction rate," it was possible to explore the effects of such independent variables as the number of witnesses and information gathered in questions about each witness. The results obtained seemed to indicate that each additional witness contributes noticeably to the conviction rate, even after as many as four or five witnesses have been named in the case.

By partitioning the total cases observed into the categories of "felonies and misdemeanors" and into the categories "crimes against persons" and "property crimes," it was possible to establish whether the general pattern also held for selected major categories of crime. The findings suggested that the marginal contribution of each witness is about the same for felonies as for misdemeanors, but greater for property crimes (i.e., arson, extortion, burglary, larceny, stolen vehicle, forgery, fraud, embezzlement, stolen property, or property damage) than for crimes against persons (i.e., homicide, kidnapping, sexual assault, robbery, or assault).

Next, the issue "What information should be obtained from witnesses?" was examined. Fifteen of the twenty-one explanatory variables which were considered resulted from interrogating victims and witnesses, and only one of these fifteen was significant above the .10 level. The result was fewer significant variables than we would normally obtain at the .10 level simply by chance from fifteen variables that were absolutely unrelated to the conviction rate. This casts some doubt on whether even the one variable found significant is, in fact, systematically related to the conviction rate.

Thus, the results indicated that it was more important to obtain information on an additional witness rather than to ask more questions of fewer witnesses. The form was kept intact, but implemented on a sampling basis rather than for all cases. This was practical since the form was useful primarily for research purposes rather than operations.

Use of Data Base as Basis for Sample Survey

To conduct a field survey in the witness cooperation study, it was decided to use the PROMIS data base as the sample frame to select cases. To draw the sample from PROMIS, a file which contained only those cases eligible for inclusion in the sample was extracted from the PROMIS master file. This file contained 6,266 cases and of these, 3,170 had been rejected for prosecution either at screening or later in nolle prosequi actions or dismissals prior to trial by the U.S. Attorney's Office. Of the cases rejected, 38 percent had been rejected for witness noncooperation. This subfile was then used to select a probability sample that would guarantee the inclusion of noncooperators for comparison with cooperators.

Cross Tabulations

Even though research studies generally rely on statistical analyses and deal in aggregates, it is important to record measurements for individuals (names can be blocked out for privacy and security purposes as these are not needed for research). This capability is necessary to be able to develop cross tabulations and regression models. For example, one might want to study what happens to black male defendants (within a certain age group, who are unemployed at the time of arrest) with respect to bail decisions, pretrial detention, case processing time, and outcome. Such an analysis requires unique identification of individual offenders and the ability to relate their characteristics and cases.

Passing the Entire Research File

Since operational data systems are characterized by retrieving and processing one record at a time, on-line real time applications are popular in such systems. In research or planning applications, one generally needs to read the entire data file to develop the required statistics. The research data base itself may be constructed by means of a random sample of records drawn from the operational data base. Alternatively, the researcher may draw his sample simultaneously with reading the entire file to generate statistics.

Geocoding Applications

The PROMIS data base contains street codes and street addresses for the location of the offense, the arrest, the residence of the witnesses, and the residence of the defendant. For the analysis of patterns of criminal behavior and citizen involvement, the research project is planning geographically oriented studies of victimization, offense reporting, arrest, and prosecution patterns. What types of citizens are more likely to report crimes and cooperate in prosecution? What characterizes the situations of the different types of crime?

To accomplish such studies, it is planned to link data records from a victimization survey conducted in Washington, D.C. by the Bureau of the Census for LEAA, a citizen opinion survey on crime conducted by the D.C. Office of Criminal Justice Plans and Analysis, the police offense reports tape, a crime index file, census tapes, social indicator files, and PROMIS. Each of these files contains data which can be accessed at different geographic levels; for this, geocoding may be necessary to make the records compatible. For example, the level of detail to which the victimization survey results can be made available will be census tract level at best due to the size of the sample. In order to relate PROMIS data records on crimes prosecuted in a particular census tract, it will be necessary to use a geocoding program such as the Bureau of Census' ADMATCH[6] to match certain street addresses in PROMIS records against an address coding guide prepared by the Metropolitan Washington Council of Governments. Once a match is obtained, the corresponding census tract code can be accessed from the Address Coding Guide and appended to the PROMIS record for subsequent analysis.

Improvements in PROMIS

Recently, a new version of PROMIS has been implemented. Whenever a systems change of this magnitude is made, it will have an effect on research projects that require historical data spanning several years. It is necessary for the researchers and planners to analyze any changes in data elements and make sure they can establish equivalencies between old and new definitions and codes. Similarly, it is important to understand new data items and discontinued data items. Any statistical interpretation of longitudinal data covering data collected under the old and new systems must consider system differences.

The new Promis design resulted largely from research applications. Under the old PROMIS, each charge in the case was placed in a separate trailer record. Actions, such as continuances, also formed trailer records related to each charge. There were a varying number of charges per case and a variable number of actions per charge. Thus, it was a little cumbersome to determine the current status of the case. The file organization was acceptable for operational purposes since the processing programs were tailored to the file organization. For research purposes, it is desirable to be able to apply standard statistical packages, such as Statistical Programs for the Social Sciences (SPSS), Harvard's Data-Text, or UCLA's Biomed systems. Moreover, such data record formats are difficult to process. For example, the statistical packages cannot handle a varying number of trailer records. The new PROMIS provides a case header record which contains case status and summary data such as current continued date, case disposition, number of continuances, and number of witnesses. Many research data needs can be met by extracting those case summary records. There are still some variable length records in the file (charges, witnesses) but a special program has been written to analyze the file and develop fixed length formats for application of the statistical packages.

Another problem under the old operational system was that the court assigned a new case number to a felony case after indictment. For operational purposes, it sufficed to track by the new case number. For statistical purposes, it was desired to show this as one case under the new number in order to allow tracking all events in that case from arrest through final court disposition. This merging of the data into one case record has been accomplished under the new PROMIS. In addition, numeric codes are used in the new PROMIS in lieu of alphanumeric codes to facilitate the application of standard statistical analysis packages.

Conclusions

What can be concluded from the experiences in utilizing PROMIS for research and planning purposes? The conclusions may be summarized as follows:

- The PROMIS data base has been useful as a sample frame for designing field surveys and in-depth research studies. Data from the operational data base has been effectively linked with survey data.

- When certain data items required by planners and researchers were weak in reliability or missing, forms were modified and procecures installed for gathering the required data.

- Multiple regression techniques were effectively used to analyze the criminal-justice system and to test forms design alternatives by showing how certain data items contributed to criminal justice system performance.

- Using the data for research purposes has pinpointed data inaccuracies and inconsistencies and needed improvements in file structure. It also has sharpened the user's understanding of the importance of certain data and procedures, such as longitudinal data and those data measuring witness cooperation.

- The planning agency for the District of Columbia has utilized statistics drawn from the PROMIS data base in developing the annual Comprehensive Plan for criminal justice programs, for progress evaluation during the year, and for special research studies.

Effective communication among planning, research, information systems and operational personnel has benefited the research, the planning, the information system, and office policy and operations to a greater extent than could be accomplished by isolated efforts. Only through use of an information system can that system be validated and refined.

NOTES

1 For a complete discussion of the development of PROMIS, see William A. Hamilton and Charles R. Work, "The Prosecutor's Role in the Urban Court System: The Case for Management Consciousness," Journal of Criminal Law and Criminology. Vol. 64, No. 2 (June 1973), 183-189.

2 The gravity of the crime is scored according to a modified version of the Sellin-Wolfgang seriousness index (see Thorsten Sellin and Marvin E. Wolfgang, The Measurement of Delinquency [New York: John Wiley, 1964]). The defendant's criminal history is evaluated according to a scoring technique that reflects the "base expectancy" scoring system developed by Donald Gottfredson (see Donald Gottfredson and Jack A. Bonds, A Manual for Intake Base Expectancy Scoring, April 1, 1961 (Form CDC - BEGLIA), Sacramento, California: California Department of Corrections, Research Division).

3 Ibid.

4 The PROMIS Research project is funded under Grant No. 74-NI-99-0008-G by the National Institute of Law Enforcement and Criminal Justice, Law Enforcement Assistance Administration, U.S. Department of Justice, Washington, D.C.

5 "A Study of Witness Cooperation in the District of Columbia," Dr. F. Cannavale, Jr., Institute for Law and Social Research, 1125 Fifteenth Street, N.W., Washington, D.C., forthcoming.

6 U.S. Department of Commerce, Census Use Study, ADMATCH User's Manual, (Washington, D.C.: Bureau of the Census, 1970).

CHAPTER 11

STATISTICAL TECHNIQUES AND THEIR LIMITATIONS IN THE ANALYSIS OF CRIMINAL JUSTICE DATA

Brian Forst

INTRODUCTION

Many crucial questions in the criminal justice system that have been addressed previously by means of conjecture and anecdote are now being confronted factually and systematically. Some examples:

- What effect does the pretrial release decision have on whether a defendant is convicted?
- What effect does the use of a diversionary program have on whether a defendant recidivates?
- To what extent does delay in court processing increase the likelihood that a witness problem will develop in a case?
- How many homicides are deterred, on the average, with each execution of a convicted murderer?

The fundamentals of the scientific method--construct a theory, collect pertinent data, subject the theory to empirical validation using the data, and compare the findings with previous work--still form the basis for addressing such questions. However, the specific tools used in carrying out this research have progressed dramatically. Many of these improvements have taken place since the 1950's, following the advent of high-speed, large-capacity computer technology. Complicated data manipulations that were totally infeasible using a calculator, scratch pads, and sharp pencils are now commonplace.

These developments in quantitative research have placed new burdens on policy setters in the criminal justice system. Police officials, lawyers, court administrators, judges, corrections officials, and lawmakers are being confronted more frequently than ever with elegant theories formulated in cryptic mathematical language and tested on the basis of formidable sounding statistical criteria (e.g., the ratio of the regression coefficient to its standard error to test for partial effect, the coefficient of determination adjusted for degrees of freedom to test for goodness-of-fit, the likelihood ratio chi-square to test for statistical interactions, and so on). In this Chapter we attempt to cut through such esoteric jargon and describe in simple terms the problems involved in analyzing data, the basic techniques for dealing with each problem, and the limitations of these techniques.

UNIVARIATE AND MULTIVARIATE STATISTICS

Consider the issue of delay in the time it takes for the court to process cases. We can describe this phenomenon either in the absence of any other factor using univariate statistics, or we can concern ourselves with the relationships between the variable "case processing time" and other factors using multivariate statistics. The state of the art of univariate statistics, unlike that of multivariate statistics, remains today essentially as it was a century ago.

Univariate Statistics

Univariate statistics are useful in providing measures that describe the way that cases are distributed over the time it takes the court to process them. Figure 1 depicts such a distribution (for convenience, based on imaginary rather than real data):

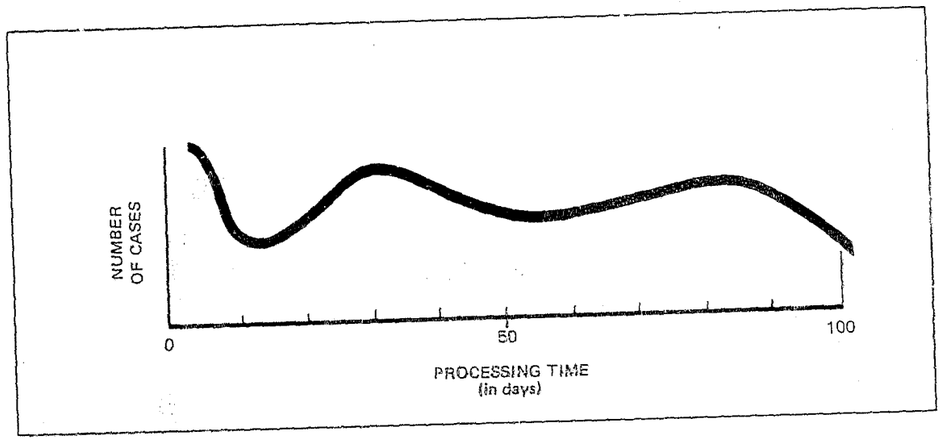


Figure 1. DISTRIBUTION OF CASES OVER TIME

As with most pictures, Figure 1 describes case processing time better perhaps, than would a thousand words that only characterize the picture. It is, nonetheless, convenient to be able to describe the distribution of case processing time in terms of certain statistical measures. One might say, for example, that more cases leave the system on the first day than on any other day--more briefly, that the "mode" is one day; or that half the cases are processed within the first 45 days--more briefly, that the "median" is 45 days; or that the "mean" (i.e., average) number of days to process a case is 50; or that the distribution is "polymodal," with peaks at days one (due to rejections of cases at screening), 30

(the mode number of days to misdemeanor trials), and 80 (the mode number of days to felony trials); or that the "range," a measure of the total dispersion or spread of the distribution, is 150 days.

Such measures are useful for making definitive quantitative comparisons of case processing times in one jurisdiction with those in another. These statistics also serve as important components of multivariate analysis, which will occupy the remainder of this Chapter.

Multivariate Statistics

Of more policy relevance than the univariate characteristics of case processing time are the multivariate relationships between case processing time and other factors. For example, if the mean case processing time for felony cases increased from 70 to a 100 days, what would happen to the percentage of cases that go to trial? Or, if witnesses were paid a sum of money to appear in court so that continuances were reduced by, say, 30 percent, what would happen to the mean case processing time? We can see that it is useful to know both how case processing time affects other factors (e.g., witness cooperativeness, convictions, and recidivism), and what factors influence case processing time (e.g., experience of the prosecutor, defense counsel, and judge; existence of tangible evidence; and degree to which the defendant is a "career criminal"). In the former case, we would be looking at case processing time as a cause. In the latter case, we would be looking at case processing time as an effect. Such issues as these can be addressed using tools of multivariate statistical analysis.

EXPERIMENTAL AND NONEXPERIMENTAL DATA

One way to learn about causal relationships is to conduct a controlled experiment. Because of ethical and legal considerations, however, controlled experiments have extremely limited appropriateness in the criminal justice system. But we can gain insights into causal relationships by applying multivariate techniques to nonexperimental data (i.e., data that accumulate in the normal course of operations). Nonexperimental data are becoming more abundant in the criminal justice system, particularly due to the growth of automated record keeping.

The challenge in trying to draw inferences about causality from nonexperimental data stems from the fact that there are, typically, countless hypothetical causal relationships one could dream up that could have produced the resulting data. It is by no means certain that we will identify the true set of causal relationships. Even if we do identify this set for consideration, we may still incorrectly conclude that some other set is the true set, due to measurement errors, the omission of important factors from the data, or random errors that result from having too few observations for the analysis.

CAUSAL RELATIONSHIPS AND STATISTICAL ASSOCIATIONS BETWEEN TWO FACTORS

Statistical techniques are well suited for revealing the existence of an association between two factors. Measures such as the "correlation coefficient" are commonly used to reflect the strength of this association. Unfortunately, the simple statistical relationship between two variables is so often misleading that it is generally unwise to take it very seriously as an indication of a true causal relationship.

An example based on real data will serve to highlight the danger of drawing inferences about the causal relationship between two factors based on a simple statistical association between the two. It has been found that felony arrests are 8.7 percent more likely to be accepted for prosecution when the victim is a business, ignoring every other factor. Is it appropriate to conclude from this statistic that the prosecutor is favoring businesses over private citizens? No, since it is easier for the prosecutor to obtain convictions for felony arrests when stolen property is recovered or when there are witnesses, and those cases involving victims that are businesses are more likely to have recovered property and witnesses than other cases. In fact, when these and several other additional factors are included in the analysis, the observed statistical association between the factors "victim is a business" and "case is accepted for prosecution" disappears. Hence, it is reasonable to conclude that the simple two-variable relationship observed above is spurious.

A statistical association between two variables is called a simple relationship when it ignores every other factor, and a partial relationship when it takes explicit account of other factors. For the example just described, the simple relationship between the factors "case is accepted for prosecution" and "victim is a business" was significantly positive, but the partial relationship observed was not.

Another example of a spurious statistical relationship that has been found in criminal justice data is one between the defendant's attorney and the outcome of the case. When every other factor was ignored, it appeared that cases handled by public defenders were more likely to wind up as a conviction against the defendant than cases handled by other defense lawyers. Does this mean that public defenders do not serve their clients as well as other defense counselors? Not at all. In fact, in the jurisdiction studied, the opposite appears nearer to the truth. The public defender was found to take a disproportionate number of cases which involved serious crimes (i.e., involving weapons, personal injury, etc.) and defendants with extensive arrest records. The findings indicate that these types of cases are more difficult to defend. After taking account of these and other pertinent factors, the statistical indicators suggested that public defenders are *more* likely to get their clients "off of the hook" than other counselors. In the jargon, while the simple relationship observed between the factors "conviction" and "public defender" was weakly positive, the partial relationship observed was significantly negative.

TECHNIQUES FOR TAKING ACCOUNT OF OTHER FACTORS

How do we take account of other factors in the analysis of non-experimental data? The simplest way is through cross classification, based on the factors that cause partial relationships to differ from simple relationships. For example, in investigating whether the prosecutor favors businesses, one would compare cases involving business victims and other types of cases for differences in the percentage of arrests accepted for prosecution, *after* dividing the cases into four categories: (1) those involving witnesses and recovered property, (2) those involving witnesses but no recovered property, (3) those involving recovered property but no witnesses, and (4) those involving neither witnesses nor recovered property. Similarly, in the example of analyzing the effectiveness of public defenders, one might compare conviction rates between the cases they handle and other cases after cross classifying by the seriousness of the case and the extensiveness of the defendant's arrest record.

Comparing percentages between various categories, in the manner suggested in the preceding paragraph, is the basic ingredient of a statistical technique known as contingency table analysis. This technique provides an opportunity to infer whether differences between percentages of selected characteristics are real or simply due to an insufficient number of observations, after taking into account other factors that may affect the difference, as noted above. The test for inferring whether the observed differences is sufficiently large to qualify as one that is real rather than random is based on the chi square statistic.

A commonly used alternative to contingency table analysis is multiple regression analysis, or simply "regression analysis." In regression analysis, the factors that cause partial relationships to differ from simple relationships can be accounted for either by using these factors as the basis for grouping the observations into categories to be analyzed separately, or by incorporating them as independent variables (i.e., the explanatory factors) in the analysis. Regression analysis provides a basis for determining a number called the regression coefficient, which is an estimate of the effect that a change in one factor has on another factor, after other factors have been acounted for statistically.

We can get a general sense of how regression analysis works by returning to our time in system example. A simplistic regression analysis of the effect of case processing time on the likelihood that a case will go to trial might produce a regression coefficient of -.01, which would suggest that each additional month of delay, *on average*, reduces the probability of trial by one percent. We regard this example as simplistic because we expect that the relationship between case processing time and likelihood of trial is really nonlinear; that is, while the one month increase from two months to three months to process a case might cause the probability

of a trial to decline from .15 to .10, an increase from seven months to eight months might increase the probability from .08 to .09. More powerful (and more complicated) versions of regression analysis are capable of dealing with such problems.

In both contingency table analysis and regression analysis, the "other factors accounted for" are called control variables, so named because it is the absence of a controlled experiment that requires another means of statistical control.

An important problem in controlling for other factors consists of determining which factors are the most important to control for. This problem is generally addressed in two ways; theoretically (Which variables seem to be the most important to take account of?) and empirically (Which set of control variables provides the most accurate estimates?). Since the empirical approach is the more objective of the two, it is generally the preferred approach, although bizarre empirical results that defy a reasonable theoretical explanation should not necessarily be taken as correct. Both contingency table analysis and regression analysis provide objective methods for determining the "best" set of control variables. These go beyond the scope of this Chapter, but the reader can consult sources cited at the end of the Chapter for specifics. In any case, the mere existence of control variables in an analysis of causal relationships, without a reasonable justifiction for that particular selection of control variables out of a larger set of variables, is a very bad practice that should arouse the reader's suspicion.

POINT ESTIMATION AND HYPOTHESIS TESTING

There are two basic modes of statistical inference--point estimation and hypothesis testing. In analyzing relationships among two factors, the former is used to estimate the degree of relationship between the factors. The previous discussion about estimating the average effect of a change in case processing time on the likelihood of trial is an example of point estimation.

Hypothesis testing, on the other hand, is used in analyzing causal effects to infer from a sample whether a relationship exists between two factors. In hypothesis testing, we might say, for example, that if case processing time and the likelihood that a case goes to trial were, in fact, absolutely unrelated, the probability of obtaining a statistical association at least as close as we have obtained is .03. If we regarded .03 as being sufficiently small, we would reject the hypothesis that the two variables are unrelated. The determination of what constitutes a"sufficiently" small probability to cause a rejection of the hypothesis is totally arbitrary, although the numbers .01 and .05 have become popular, for reasons that are not completely clear. If we accept the proposition that any two

variables are at least remotely related, however infinitesimally so, then given an arbitrarily high level of significance, a large enough sample will always cause us to decide that a relationship exists!

Point estimation and hypothesis testing are most useful when they are combined. Either used alone is seldom of much value. The point estimate says nothing about the reliability or precision of the estimate; that is, it says nothing about the chances that a different random sample of the same size will produce a similar result. The hypothesis test, on the other hand, says nothing about the value that is estimated. For example, it is of little value to know only that the observed difference between the percentage of cases won by public defenders and the percentage won by other defense lawyers is significant at, say, .02; presumably, one would be more interested in knowing what the difference is.

COMPLICATED RELATIONSHIPS

If we consider again the matter of analyzing the effect of changes in case processing time on the likelihood that the case will go to trial, we see yet another problem: the causal relationship between these two variables is likely to run in the *opposite* direction as well. Changes in the likelihood of trial ought to affect case processing time, too. In particular, the weakest cases, with small likelihood of trial, would seem unlikely to be in the system for long. Very strong cases may be more likely to leave the system as pleas of guilt prior to trial. It would seem to be the fairly strong cases that survive longer in the court.

It is appropriate, but rare, to take account of such confounding effects in analyzing a particular causal relationship. If one were primarily interested in one phase of a complicated relationship between two or more factors, a misleading conclusion could be drawn if the analysis fails to separate the major components of the full set of causal relationships.

There are two basic types of complications: circularity and recursivity. Circular relationships, such as that between case processing time and the likelihood of trial, both start and end with each factor involved in the circle. Recursive relationships involve a chain of causally related factors, with one affecting the next in a noncircular manner. An example of a recursive relationship is that between the experience of an arresting police officer, the amount and quality of evidence obtained in a case, and the likelihood of conviction. Presumably, the first of these affects the second, the second affects the third, and any other direct relationship among these three variables is likely to be negligible. Circular relationships generally pose greater analytic problems than recursive ones.

It should not be a surprise that the techniques for dealing with such problems are as complicated as the relationships themselves. These techniques have been developed mostly since 1960 by economists who have greatly advanced the techniques of regression analysis and by sociologists (in particular, Leo A. Goodman, University of Chicago) who have made contingency table analysis more powerful. Further discussion of these techniques is beyond the scope of this Chapter.

HOW TO CHOOSE BETWEEN STATISTICAL TECHNIQUES

The array of multivariate statistical methods is bewildering. Not only are there various techniques within contingency table analysis and regression analysis; there are also related multivariate techniques such as factor analysis, the principal components method, multiple classification analysis, cluster analysis, discriminant analysis, and canonical correlation analysis.

Incredibly enough, the choice of one particular technique over another is often as arbitrarily made as the choice of a level for rejecting an hypothesis. Unlike the choice of a level of significance, however, there *are* objective criteria for selecting among alternative techniques of multivariate analysis. The important questions to be answered are

- How accurately does each technique predict or describe the phenomenon in question?
- How sensitive (or "robust") are the results to various modifications within each technique?

One way to assess the accuracy of each technique is to use a split sample validation procedure. Under this procedure, part of the observations are used to estimate the relationships among factors, and the remainder are used to measure how closely each technique (and each variation or model within each technique) predicts the true values of the phenomenon being described. A standard measure of predictive accuracy is a statistic called the "mean square error."

"Robustness" is determined by examining each reasonable modification of the analysis to see how sensitive the results are to each modification. Modifications typically consist of the use of linear and nonlinear assumptions, the use of alternative sets of control variables, and the inclusion and exclusion of "outliers" (i.e., observations that are so unusual as to be of questionable validity). When the results are sensitive to such modifications, it is appropriate to investigate which modification is the most accurate and to doubt whether *any* is accurate.

It is, perhaps, most common to select a technique on the basis of one's familiarity with it or the realism of the assumptions on which the technique is framed. Familiarity is obviously an inferior criterion for selecting a particular technique, for much the same reason that one's familiarity with a hammer would be a poor rationale for using only a hammer to build a house. Realism of the assumptions of a technique may be an equally inferior criterion for selecting or rejecting a technique. If one technique that happens to be based on supposedly restrictive or unrealistic assumptions consistently predicts or describes a phenomenon more accurately and efficiently than every other technique, it is difficult to see why its results should be ignored.

SOURCES OF FURTHER INFORMATION

The following books are recommended for further reading both for general background information and specific topics, as indicated in the titles:

Hubert M. Blalock, Jr., Social Statistics (New York: McGraw-Hill, 1972).

James A. Davis, "Multivariate Contingency Tables: An Exegesis of Goodman's Recent Papers," in Sociological Methodology, edited by Herbert Costner. (San Francisco: Jossey-Bass, 1974), pp. 189-231.

N.R. Draper and H. Smith, Applied Regression Analysis (New York: John Wiley & Sons, 1966).

J. Freund, Modern Elementary Statistics (Englewood Cliffs, N.J.: Prentice-Hall, 1967).

William J. Goode and Paul K. Hatt, Methods in Social Research (New York: McGraw-Hill, 1952).

Donald F. Morrison, Multivariate Statistical Methods (New York: McGraw-Hill, 1967).

Sidney Siegel, Nonparametric Statistics (New York: McGraw-Hill, 1956), especially pages 195-202 for a discussion of the use of the chi square statistic in analyzing relationships among factors in contingency tables.

Henri Theil, Principles of Econometrics (New York: John Wiley & Sons, 1971), especially pages 429-539 for a discussion of the analysis of circular and recursive relationships among variables.

Hans Zeisel, Say It with Figures (New York: Harper & Row, 1968).

CHAPTER 12

DATA RELIABILITY AND DATA PURIFICATION

Joyce Deroy

INTRODUCTION

Before applying sophisticated analytical techniques to any data base, it is important to understand what the data are measuring and whether the data are sufficiently accurate to be useful in analysis. One way of detecting erroneous data is through automated data editing and verification techniques. Any errors detected through this process can be listed for manual correction by checking against the contents of the physical source data forms, such as those in court case jackets. This is most effectively accomplished as part of an ongoing operational process where errors are corrected during the current period's processing cycle. If this is not possible, either through the lack of extensive automated verification techniques or clerical personnel shortages, then the task has to be accomplished with a sampling of historical files. Thus it is important for research purposes to maintain the historical computer files and the closed court case files in a form where both can be sampled in conjunction with each other.

TYPES OF ERRORS

The experiences in using PROMIS (see Chapter 9) are useful for illustrative purposes. PROMIS was originally designed in 1970 to satisfy primarily operational needs, but secondarily research needs. As a result of experience in both operational and statistical applications, PROMIS was redesigned, reprogrammed and implemented. The new PROMIS contains much more extensive automated data editing and error detection procedures. A number of forms and action codes were streamlined. In addition, several record format improvements were made to simplify analysis.

Naturally, for longitudinal analysis, it is important to track offenders and cases over time. A research project established to utilize historical PROMIS records had to assure that the codes used corresponded to comparable actions. Data records processed under the old PROMIS had to be transformed into a comparable format with the new PROMIS. Also, the data editing and verification conditions had to be made consistent.

A computer program was prepared to convert the "old" PROMIS records to simulated input transactions in the new PROMIS data format. These transactions were then used as input to the new PROMIS data edit and verification program. This program structured master file records in the new format for "correct" transactions, and rejected and listed erroneous transactions for manual error correction. The types of errors detected

by the new PROMIS include omissions, invalid codes and inconsistencies, such as the date of the trial preceding the indictment. Errors, which appeared to be conveniently correctable by computer, were corrected in a subsequent run. Statistics were also run off showing the error rates and the pattern of errors to assess data reliability for developing descriptive statistics.

In some projects, data errors will be randomly distributed over the observations that form the basis for policy recommendations. This type of error tends to obscure measurement of the true degree of association that exists between the factors under investigation. In other instances, the errors are systematically distributed in some way over the observations. This type of error produces biases or distortion in the inferred relationships. In either case, the determination must be made balancing the cost of data purification and the cost of implementing an erroneous research finding.

One example of erroneous data that is difficult to detect and correct, but quite important to verify, is the street address of witnesses. In a study of witness cooperation, it was found that about 30 percent of the recorded witnesses could not be located due to erroneous addresses recorded in the PROMIS data base.* It was important to attempt to isolate whether these errors resulted from:

- false addresses deliberately given the police by witnesses,
- erroneous recording by the police,
- erroneous recording by the prosecutor, or
- clerical errors in transcribing from source data forms to machine-readable inputs.

Each of the above types of errors implies a quite different problem. Time constraints precluded a thorough analysis during the study, but a cursory review indicated quite a few obviously false addresses deliberately given. One common example of this was "1600 Pennsylvania Avenue." It was of concern to note the lack of attention given to accuracy of this data item by police and prosecutor alike. For purposes of that study, a sample analysis showed the errors were equally distributed among witnesses labeled cooperative and witnesses labeled noncooperative by the prosecutor so it did not appear to seriously bias the sample. However,

*A Study of Witness Cooperation in the District of Columbia, Dr. Frank Cannavale, Jr., Institute for Law and Social Research, 1125 Fifteenth Street, N.W., Washington, D.C., forthcoming.

such errors can be crucial in locating witnesses for trial appearances, and both police and prosecution management were accordingly advised.

AN EXAMPLE OF A RELIABILITY ANALYSIS

To estimate the reliability of the PROMIS data 100 cases were randomly selected from the more than 15,000 screened* by the prosecutor in 1973. An attempt was made to locate documents prepared by four separate criminal justice agencies (police, prosecutor, court and bail agency) on these 100 cases. The purpose of such an exhaustive search was to duplicate to some extent sources of a particular data element so that in addition to reporting the incidence of agreement and disagreement between the data in PROMIS and the source document, the reliability of the source documents themselves could be assessed. It was not possible to locate documents prepared by each of the four agencies for every case though often a second or third pass through the files uncovered a document missed previously.

When as much data as could reasonably be located on the 100 cases had been gathered and recorded, the task of comparing each of the 120 data elements for each case as recorded in PROMIS to the corresponding item on one or more of the source documents was begun. (It should be mentioned that between recording and comparing, all fields which were coded for entry into the computer system were recoded so that coding errors would also be detected.) Figure 1 is a sample of the form designed for recording each of the possible outcomes of the comparison for each of the 120 fields. A separate form was completed for each case since tallying responses on a master sheet would not have permitted recovery if the totals for any field failed to equal 100. In addition, this method facilitated checking by a supervisor for consistency across coders.

Reliability estimates were produced by tallying for each field the results of the comparison for each of the 100 cases. Separate conclusions could be made about the incidence of omissions, unknowns, agreement and disagreement of each data element compared to the source document.

In some cases the researchers were interested in the reliability of a set of data: for example, defendant characteristics or arrest information rather than the reliability of each data element contributing

*The prosecutor screens or reviews each arrest made by the police and elects to file charges in approximately 75 percent of the arrests.

Case Number 16873
PDID 57002
No Papered Case YES

"-0" on PROMIS sheet means blank
Leading zeroes are suppressed on PROMIS sheet

Field Name	Notes	Complete below if entry other than blank/unk appears on more than 1 source document: 1 All sources agree	2 Disagreement among sources	Complete 1 of 3 categories below. Check only 1 box in applicable category (in the 1st category, 1st and 2nd box may be checked.): PROMIS other than blank/unk: 3 Agrees at least once	4 Disagrees at least once	5 All entries blank/unk	6 No document to compare	PROMIS unknown: 7 1 entry not blank/unk	8 All entries blank/unk	9 No document to compare	PROMIS blank: 10 1 entry not blank/unk	11 All entries blank/unk	12 No document to compare
Current case number				✓									
PDID number				✓									
M/F Code or Bkdn date				✓									
Arrest date				✓									
Arrest time				✓									
Arrest address				✓									
Offense date				✓									
Offense time				✓									
Offense address				✓									
Defendant address				✓									
Defendant sex				✓									
Defendant race				✓									
Defendant DOB	Compare year only			✓									
No. of codefendants				✓									
Urine test results											✓		
Arraign Appear								✓					
Bail type granted													✓
Bond amount													✓
Defense counsel code													✓
Defense counsel type													✓
Screening AUSA				✓									
Reviewing AUSA				✓									
Papering date							✓						
Prop/Evid recovered							✓						
Arresting PO badge				✓									
Assisting PO badge				✓									
Arresting PO unit				✓									
Assisting PO unit				✓									
Number of witnesses	PO only for no papered case			✓									
Final disposition				✓									
Trial type	Jury or Nonjury -look out for cont. array												✓
Number of continuance				✓									
Disposition date							✓						
Sentence date	Look at Cont. Array												✓
Special assignment	1 in PROMIS=Spec. assgn												✓
Action date 1													✓
Action reason 1													✓
Action date 2													✓
Final disp/NP reason	Pre-breakdown charges show final disp 390 or 391												
Charge 5													
Final disp/NP reason													
Charge 6													
Final disp/NP reason													
1 (1=J, 2=K, 3=L, 4=U)	Look at Charge also												
2													
3													

Figure 1. A SAMPLE FORM FOR RECORDING POSSIBLE OUTCOMES

to that set. For this purpose the statistics for groups of related data elements were grouped so that estimates of reliability could be stated in more general terms. Figure 2 shows the reliability of defendant identifiers as an example.

DEFENDANT IDENTIFIERS:
(PDID, SEX, RACE, YEAR OF BIRTH)

RELIABILITY OF SOURCE DOCUMENTS

Instances Where 2 or More Source Documents Were Located = **82**

Agreements = **77**
Disagreements = **5**

OPPORTUNITIES TO COMPARE PROMIS AND SOURCE DOCUMENTS			(385)
Agreements	98.4%	(379)	
Disagreements	1.0%	(4)	
Agree and Disagree	.5%	(2)	
	100 %		
UNKNOWN IN PROMIS			(2)
Source Document Indicated Data Known		(0)	
Unknown in All Source Documents		(2)	
	100 %		
BLANK IN PROMIS			(0)
All Source Documents Blank		(0)	
Source Document Contains Data		(0)	
	100 %		
NO COMPARISON DOCUMENT			(13)
			(400)

Figure 2. RELIABILITY OF DEFENDANT IDENTIFIERS

Few data elements were omitted frequently. Some were so seldom able to be located on the source documents for comparison that their reliability could not be assessed in the 100 case sample. Use of such data elements in the research project was suspended pending further investigation.

In some cases the reliability of the data could be greatly improved by loosening the definition of agreement, i.e., by counting a slight variation from the source document as an agreement. For example, if the date of a court event was allowed to vary one day from the source document and still be counted as an agreement, the percentage of agreement (versus disagreement) rose from 95.6 percent to 98.2 percent.

In other cases the reliability of a data element could be greatly increased by collapsing the categories defined by the variables. One instance of this was defense attorney type. PROMIS had attempted to collect three different values for this variable: public defender, CJA (an attorney appointed to represent an indigent defendant), and privately retained. In the initial pass the reliability of the field given the three values was found to be extremely low. It was discovered that the individuals coding this information for entry into PROMIS did not have available the data to distinguish between CJA and privately retained, though the public defender identification was easily discernable. When the variable was redefined as having only two values: "public defender" and "not public defender," the instance of agreement (where there was an entry in PROMIS and a source document for comparison) became 100 percent.

The major benefit of this reliability analysis was to demonstrate the quality of most of the data fields and assume that only reliable data were used in the statistical analyses. One side benefit of the reliability study was that insufficiencies in the manual and computer system for collection and recording of data were highlighted. The knowledge gained from closely examining the system was turned back into it so that data collection techniques which were imperfect were improved for future use.

CHAPTER 13

A MODEL TO AID IN PLANNING FOR THE TOTAL CRIMINAL JUSTICE SYSTEM

Alfred Blumstein

INTRODUCTION: THE PLANNING PROCESS

In this paper, we examine the criminal justice system planning process, describe and discuss one tool (the Justice System Interactive Model) that can be a valuable adjunct to the planning process, and indicate how it has been used in certain jurisdictions. Other jurisdictions can then determine how they might make use of it.

Planning is an analytical process in which an organization attempts systematically to make rational choices for the future. The emphasis is on the process by which those choices are made, rather than on the choices themselves. In this case the choices are "rational" in the sense that they address the objectives of the criminal justice system, as opposed to other objectives that may well influence behavior in the criminal justice system (for example, the desire of political figures to be re-elected and the desire of civil servants to preserve their jobs). Even though these objectives are real, (and are certainly rational from the perspective of the people who are concerned with them), rational planning in the criminal justice system involves concern for questions like crime reduction, enhancement of liberty, and reduction in the total social cost of crime control.

The second principal characteristic of the planning process is the orientation to the future, making choices now for implementation in a future--and therefore uncertain--world. The planning process must face up to that uncertainty. In the criminal justice field, such uncertainty derives partly from the inability to predict the variety of social and political changes that will influence behavior within the criminal justice system, and partly from the effect of its own actions on criminal behavior. Furthermore, the mutually adaptive behavior between crime and the actions of the criminal justice system results in each of them influencing each other in ways that are only minimally predictable, while general changes in the political and socioeconomic climate will affect both crime and the criminal justice system.

Be that as it may, the presence of uncertainty does not preclude the use of analytical techniques and techniques of statistical estimation any more than the uncertainty about the roll of a pair of dice precludes the experienced gambler as well as the statistician from being able to project that the roll of a "7" is six times more likely to occur than the roll of a "12," even though the next roll might well be a "12." Because the processes in the criminal justice system, however, are far more complex, no analyst can make anything like comparable projections, and so considerable judgment must be required in obtaining and using any such projections.

The third principal emphasis in the planning process is on the need to make choices, an operation which involves a combination of forecasting, prediction of impact, and attribution of the cost benefits of a given action. Thus, the planning process inherently involves the value considerations of the community. It would be even more useful, however, if that valuation process were applied to factual information about the consequences of alternative courses of action once those consequences are illuminated.

Such a planning process of "making rational choices for the future" involves the following steps:

1. Describe the current system;
2. Project the future environment;
3. Develop alternatives among which to choose;
4. Analyze the impact of the alternatives ("pre-evaluation");
5. Allocate resources to the choices and implement them;
6. Evaluate the impact ("post-evaluation");
7. Repeat the process on a regular and continuing basis.

The need for these seven steps is reasonably logical. First, in order to consider sensibly any revisions or changes in the current system, it is necessary to have some characterization of it. It is necessary to know what the present crime situation is, how the current criminal justice system works, what resources are applied, where and how they are applied, and how the budgets are distributed over the parts of the criminal justice system and over the various crime types.

Because the planning process is concerned about the future, the next logical step would be to project the description of the current system to the future. This involves forecasting anticipated crime rates, as well as projecting the future behavior of the criminal justice system. The projection of crime rates might be undertaken by a variety of means, including simply assuming that current crime rates would persist into the future, extrapolating the last few years of data, using more sophisticated time-series analysis techniques, or using more sophisticated models involving the crime rates per demographic group (e.g., age, race, sex combinations) and projecting the demographic mix. Each of these methods has its virtues, the simpler ones being the easier to understand and to implement, the more elaborate ones imposing a considerably greater demand for data and for computational power, although sometimes providing considerably greater predictive power in return.

For the features of the criminal justice system other than crime rates and future behavior of the system itself, it would be desirable to find parameters that are reasonably stable over time. Thus, for example, the number of cases processed will vary from year to year, and so the number

of guilty or innocent dispositions will vary accordingly. In many settings, however, the proportion found guilty or innocent may well be far more stable than the total costs and workloads, (which may depend much more on total number of units processed). Here again, the easiest projection would be to assume that next year's branching ratios, costs, workloads, etc., are the same as this year's. If a longer time series on these parameters were available, then one might project that time series in ways similar to the projection of crimes. In other cases, however, the introduction of new programs, or new court decisions, or new legislation will force a qualitative change in some of those parameters, and that qualitative change must be taken into account in the planning process. In some cases there may be empirical evidence to indicate how a new program affects the value of a system parameter (e.g., experience in other jurisdictions or evaluative reports). In other cases, the best estimate is simply some judgmental guess.

Since the planning process involves choices, one must first define alternatives among which to choose. The development of such alternatives is inherently a creative process by which new ways of doing things are proposed by staff members of the existing agencies (although they often seem to have a compulsion to continue doing things as they have been); or by staff from the planning agencies through scanning innovations that have been used successfully in other jurisdictions; or by other means as diverse as citizens' suggestions, brainstorming sessions, etc.

The choice among these alternatives must be made before they are implemented in the local jurisdictions, and so a "pre-evaluation" process must be undertaken. In that pre-evaluation, a judgmental process is again called for, but that judgmental process can be enhanced significantly by the use of appropriate analytical tools. The next section contains the discussion of one such tool, the JUSSIM model, which can aid significantly in assessing the impact of a particular change (e.g., a pre-trial diversion program) on different parts of the system. Because this pre-evaluation stage occurs prior to implementation, it is inherently an analytical or abstract process, and so can better employ some form of mathematical "model" that represents abstractly the behavior of the criminal justice system, rather than an experiment imposed on the operating criminal justice system. The pre-evaluation process provides an initial assessment of the impact of a change both in the part of the criminal justice system to which it is applied, and also in other parts of the system. Such an assessment should include not only the "downstream" parts (e.g., the effect of a change in police arrest practice on court workloads), but also the feedback or recidivist effects. Normally, once a planning agency has made such an impact assessment, it then allocates resources to new programs--taking into account the estimate of these impacts--and weighing impacts along with the more direct costs and benefits associated with alternative choices.

Having made the choice and implemented it on at least a pilot basis, or possibly throughout a jurisdiction, the planning process can then undertake an experimental "post-evaluation" of its impact. This evaluation is modeled on the classic random, controlled experiment. But such experiments in the criminal justice system are limited by the difficulty of their

being truly controlled, by the inevitable selection effects that creep into the experiment, and by the multiplicity of influences that prevail in a social setting--all of which influence crime rates or other measures of effectiveness. Despite these problems, the evaluative information represents an important part of the planning process and should be fed back into improvement of all the previous stages for the next planning cycle.

A JUSSIM MODEL FOR THE ANALYSIS OF IMPACTS (JUSSIM I)

As described above, a central part of the planning process is the analysis of the impact of alternative choices, and as suggested this can best be accomplished with some form of model prior to implementation. The general method by which this process is done is reflected in Figure 1. A proposed system change is considered; it is translated into estimates of parameter effects which are then put through some kind of model (which may be a mathematical model or some judgmental process) to generate estimates of the system effects or "secondary effects" resulting from that change. After such a calculation and after reviewing those system effects, a user might conclude that the initial parameter estimates were inappropriate, and this reconsideration might result in estimation of different parameter values. Or he might conclude that the model does not adequately reflect some of the complexities he is certain do occur, and so he may well change the model. After several cycles through this process, he may be satisfied with the estimate of those system effects and go through the same process for another system change.

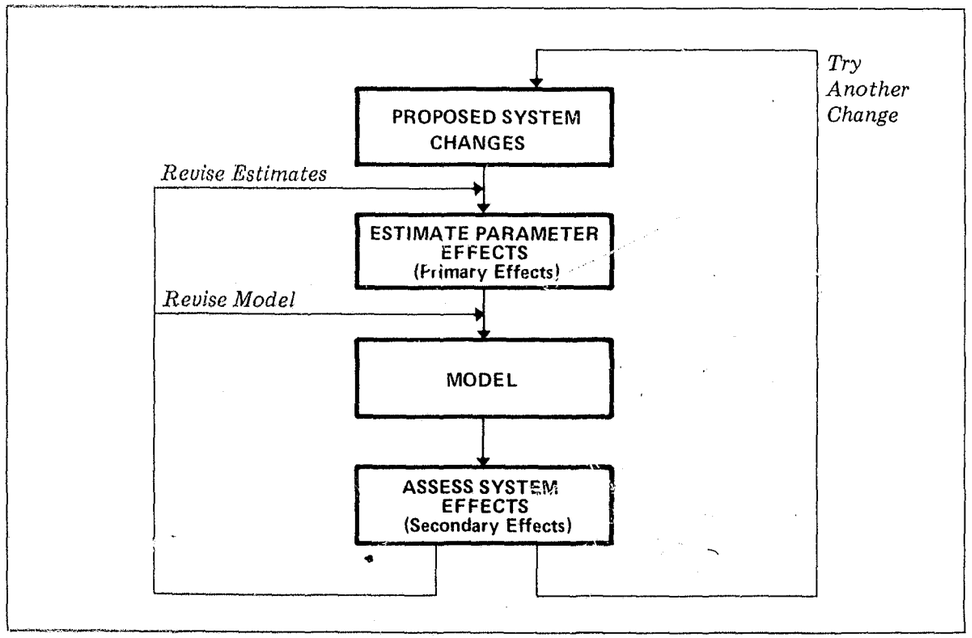


Figure 1. GENERAL PROCESS OF IMPACT ANALYSIS

This is a totally general characterization of any analysis of a proposed system change. The critical aspect, however, is the use of a model that provides the basis for translating a programmatic change in the criminal justice system to an estimate of the impact of that change on performance measures of the system. In a sense, the model may be viewed as a "black box" in which the parameter estimates are the knobs (reflecting measures like arrest policy, punishment policy, resource allocation, etc.) and the system effects are the dials (representing measures of output effectiveness such as crime rate, costs, etc.). This conceptual "black box" simulation of the criminal justice system enables the planner to twist the knobs, note the dial readings that result, and then to use that experience to help him choose the next knob settings to explore.

The crucial stage of this process is the formulation and use of the model. One such model that has been used in a number of jurisdictions is the Justice System Interative Model (JUSSIM). The model uses a computer in an "interactive" mode, with a user sitting at a terminal, calling a stored data base characterizing his criminal justice system, and interacting in a conversational way with the computer program, with no special requirements for technical training or computer programming skills, since the entire process goes on in plain English.

Single-Stage Analysis

The operation of the JUSSIM model can best be explained by examination of a single stage in the flow diagram of cases through a criminal justice system. Consider, for example, the jury trial stage for a single crime type (say, robbery). Let us assume for the moment that 100 robbery cases come to the jury trial stage, a number calculated based on flows from the earlier stages.

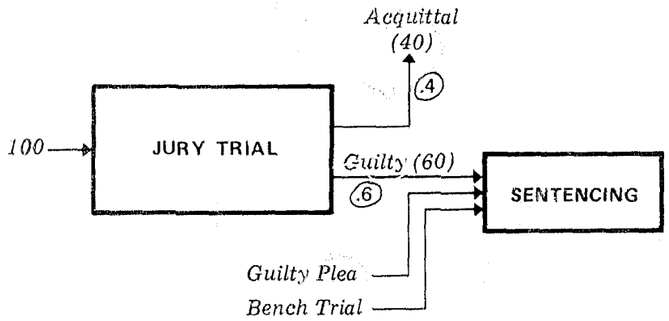
As shown in Figure 2, there are two output flow paths from the jury trial: "acquittal" and "guilty". If the branching ratios from jury trial to these two paths are .4 (probability of acquittal) and .6 (probability of conviction) respectively, then the jury trial stage feeds 60 offenders to the sentencing stage (which also receives offenders coming from guilty plea and bench trial). This same branching and collection process at the earlier stages provided the basis for calculating the 100 cases coming to jury trial. In this simple computation, the branching ratios (the probabilities of conviction and acquittal) are required input data. All input data are shown circled in Figure 2.

We are now interested in calculating the workloads, costs and resource requirements associated with the two principal resources used to process cases at jury trial: judges and prosecutors. Let us assume that the average unit workload for judges in robbery jury trials is 6 hours, and that the prosecutor (with more preparation time) must spend an average of 20 hours per case. These numbers also are necessary input data, and so are also circled.

Focusing on the judges now, their workload in handling robbery jury trials, at 6 hours per case, for 100 cases, is then simply 600 judge-hours. If a judgeship-hour (including his support staff and facilities) costs

$100 (also an input datum), then the cost of the 600 judge-hours is simply $60,000. If a judge is available 1,000 hours per year for processing cases, then the resource requirement is the workload (600 judge-hours) divided by this annual availability, or 0.6 of a full-time judge is required to process robbery jury trials. Similar computations are made for prosecutors, and these are also shown in Figure 2.

The robbery jury trial cost, then, is the sum of the resource costs associated with that stage, $60,000 for judges and $40,000 for prosecutors, for a total of $100,000. Similarly, the court costs associated with robbery can be computed as the sum of the resource costs for all stages in the court system. Then, the total court costs would simply be the sum of those costs for all the various types of crime.



	Judges	Prosecutors
Unit Workload	(6) hrs.	(20) hrs.
Workload	100 x 6 = 600 judge hrs.	100 x 20 = 2000 prosecutor hrs.
Unit Cost	($100)/hr.	($20)/hr.
Cost	600 x 100 = $60,000	2000 x 20 = $40,000
Annual Availability	(1000) hrs/judge-year	(2000) hrs/prosecutor-year
Resource Requirement	600/1000 = .6 judge	2000/2000 = 1.0 prosecutor

Stage Cost: $60,000 + $40,000 = $100,000.

◯ Circled numbers represent required input data.

Figure 2. AN ILLUSTRATIVE JUSSIM STAGE: Robbery Jury Trials

Basic Operation

The operation of the JUSSIM model begins with a "base case" reflecting the current operation of the system. All the data on the base case parameters must be collected and stored. The user of JUSSIM--the criminal justice system planner himself--then creates a "test case" by making changes in any of the base case parameters at the computer terminal. The program then reports to him the changes in flows, costs, workloads, and resource requirements resulting from the changes he introduced. He thus uses the model as a very flexible design tool by making contemplated changes, rapidly obtaining an assessment of the effects of those changes, and then trying another change suggested by reconsideration of the previous try. If the designer does not like the implications of a proposed change, then he can reject it immediately and try another.

Operating the model to assess the consequences of a system change, a user necessarily has to make assumptions about the detailed primary consequences of the change as they are reflected in changes in the system's parameters. For those changes that initially appear desirable to make, he may then want to explore the assumed consequences more carefully. Several system planners can explore the same system changes, each using his own assumptions. If a threshold of acceptance or rejection lies outside the range of consequences calculated by this group, then acceptance or rejection is clear, and further exploration is not necessary. If it lies within the range, then closer examination is required to assess the validity of the various assumptions made.

One of the virtues of this process is that it forces the planners into a debate over their assumptions and estimates rather than on the generalized goodness of a possible change. Thus, the model serves the same function as any other model, it lowers the level of argument to issues that are more fundamental and empirically testable than they would be without the model.

Model Structure

To generalize the previous example, we describe here the operation of the model in terms of its inputs, its outputs, and the relationships between them.

Inputs

The basic inputs for JUSSIM are enumerated below:

1. A listing of the crime types considered
2. An enumeration of system stages
3. For each stage, a listing of the stages it feeds and specifications of the branching ratios characterizing the proportion of flow along each of those flow paths for each crime type
4. A listing of resources, and their associated
   - A. annual unit availability (say, hours per year per resource)
   - B. cost per unit time (say, dollars per hour)
   - C. the maximum amount ("capacity") of the resouce that might be made available

5. A tabulation of the unit workloads, or times for processing a unit of flow, at any stage or flow path by each resource applied for each crime type

6. A reference flow, typically the number of reported crimes by crime type (or the number of arrests) that sets an absolute level of flow throughout the rest of the system when the branching ratios are specified.

The branching ratios, the unit processing times, and the reference flow are all specified for each crime type.

Outputs

The outputs of the JUSSIM model are presented to the user in whatever order and organization he specifies. The potential outputs include the following variables:

1. Flow through each processing stage

2. Costs associated with each stage for any aggregation of stages grouped into specified "subsystems," including the complete aggregation into a single, "total system"

3. Resource costs, indicating the costs associated with each of the resources.

4. Resource workloads, the number of manhours per year imposed on each of the resources

5. Resources required, the numbers of each of the specified resources that would be required to handle the workload.

All of these output measures are functions of crime type, and so each can be presented for each individual crime type, or as a single value summed over all the crime types.

Basic Relationships

Assuming knowledge of all the input parameters, one can then calculate the output measures by the following relationships.

- Knowing the basic input flow (say, reported crimes) and the branching ratios, the flow at each stage is calculated as the sum from each of the stages feeding it. These are the flows in the preceding stage, multiplied by the corresponding branching ratio.

- The resource workload is the flow at each stage or flow path where the resource is applied, multipled by its unit workload there. The total resource workload is then calculated as the sum of these resource workloads over all the stages and flow paths.

- The number of resources required per year is the resource workload divided by the annual availability of an individual unit of resource.

- The processing costs at a stage is the flow through the stage, multiplied by the workloads and the unit costs for each resource applied at that stage. The subsystem costs associated with a "subsystem" (or any aggregation of stages) is simply the sum of all the costs associated with its constituent stages.

Operation of a Run

In the operation of a JUSSIM run, the user's basic role is to create a "test case" to compare with a "base case" already stored. The base case is a complete set of data describing a criminal justice system and is developed individually for each jurisdiction. Once he has drawn a base case from a data file, the user then generates a test case as changes to the base case. Sitting at a terminal, the user is asked a sequence of questions about what changes he wants to make. Each of these questions is an entry gate to a "phase", which leads the user by a sequence of hierarchical steps directly to the parameters he wants to change. A separate phase is provided for changing each of the following parameters:

1. Branching ratios

2. Unit workloads

3. Annual unit resource availability

4. Resource unit costs

5. Resource capacity constraints

6. Reference flow.

Once a phase is entered, further detailed questions permit the user to specify precisely which parameters he wants to change and the crime type(s) for which he wants to make the change. All the questions are in clear language, and the answers regarding stage numbers, resource numbers, crime groups, workload numbers, and other code numbers are based on the code numbers in the base-case data file. The program then displays the base-case value for the parameters the user identifies and asks him to type in the new or test-case values.

In dealing with the multiple parallel channels of flow for all the crime types*, most users, at least initially, do not want to sit through the complete detailed output for each crime type. In a separate phase, JUSSIM permits them to choose one of a number of standard complete partitions of the crime types (e.g., into felonies and misdemeanors, Part I

* The FBI's Uniform Crime Reports (UCR) system organizes its data into 29 separate crime types. JUSSIM permits 30 crime types.

and Part II crimes, etc.) or to specify their own crime groupings. Thus, a user may choose to have the complete details on one or a few individual crime types but combine the remainder into a single group. The standard groupings are specified in the data file.

In another phase, the user specifies the output tables to be displayed. The output tables present calculated results on flows, costs, workloads, and resource requirements. The data are listed for the base case, the test case, and the absolute and percentage change in going to the test case. These results can be presented by crime group or summed over all crime groups. In the early stages of an exploration, a user will presumably want to conserve time and will examine results only for the total system or for a few critical subsystems. At the end of an exploration, the user is more likely to want more detail.

At the end of a run, the user is asked if he wishes to re-run the model. In addition to using this to explore a new issue, he will do this iteratively if he is dissatisfied with the implications of some of the assumptions he has made and would like to reconsider some of them. In a re-run, he is given considerable flexibility in respecifying his new base case (e.g., calling a data file, using his latest test case), and then begins again to create a new test case.

THE FEEDBACK MODEL (JUSSIM II)

The JUSSIM linear model discussed so far looks only at the downstream consequences of a change in the criminal justice system. It cannot explicitly calculate the effects on the future load on the criminal justice system of new programs that might change either the probability of rearrest of released individuals or the time between release and rearrest. For example, on the one hand, a pre-trial probation program might reduce the probability of arrest for some of the treated individuals. (This could occur if the threat of prosecution for the deferred charges acted as a deterrent, or if the treatment provided during the probation were more successful in rehabilitation than any of the alternative treatments that the defender might have received.) On the other hand, these released individuals would return to the street more quickly than they might otherwise, and so the opportunity for earlier rearrest is increased.

Because so many programmatic and procedural changes under consideration by criminal justice planners involve these kinds of considerations, an extention of the JUSSIM model (designated as "JUSSIM II") incorporates these feedback flowpaths into the model.

A simple flowchart of the feedback version of JUSSIM is shown in Figure 3. The new feedback flowpaths can now be characterized simply as a branching ratio, i.e., the average proportion of individuals released from some point of the criminal justice system who will ever be rearrested. We must also associate an average time until rearrest for each of the feedback flow paths. Such a concern for elapsed time was not necessary in the linear model.

Also, since a vital aspect of the JUSSIM structure is its crime-specific characterization of systems operation, the feedback JUSSIM model must specify the new offense for which recidivists are rearrested. This requires the inclusion of a "crime-switch matrix" in the model, a matrix which distributes the flow of individuals released from the criminal justice system after arrest for one crime type into a flow of individuals rearrested for another (or possibly the same) crime type.

In addition to the flow of recidivists, the feedback model also requires specification of the input flow of virgin arrests: the individuals arrested for the first time who may subsequently reappear as recidivists. This is a distinction rarely made in the conceptualizations of the criminal justice system, and less often in available data.

Thus, the principal new elements required to extend the linear JUSSIM model to incorporate the feedback flow of recidivists are:

1. The probability of rearrest by crime type of first arrest
2. The average time until rearrest by crime type of first arrest
3. A crime-switch matrix linking the prior and subsequent crime types for recidivists
4. The number of virgin arrests by crime type.

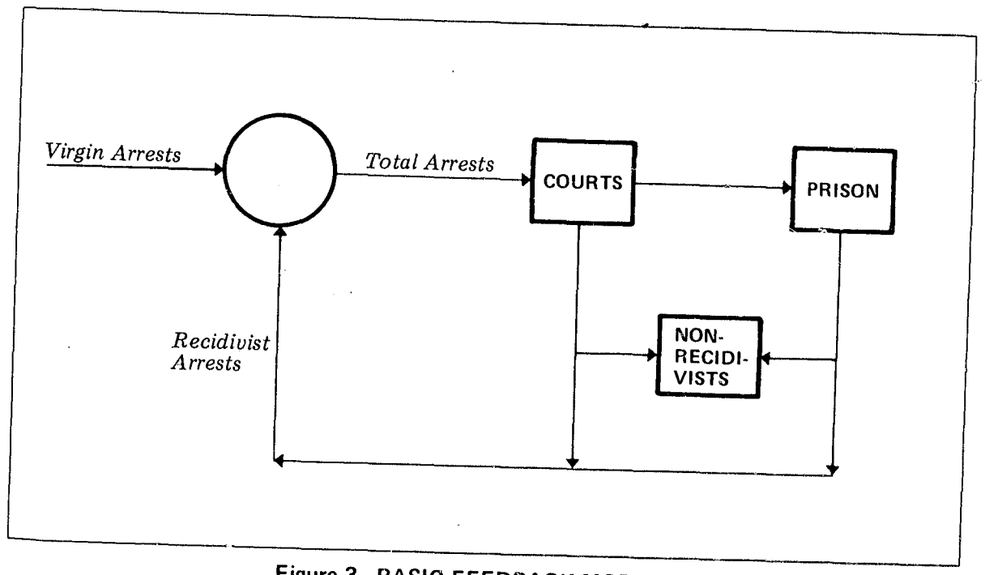


Figure 3. BASIC FEEDBACK MODEL

In addition to these extensions to the basic model, the feedback JUSSIM II model requires significant changes in the use of the model. The linear JUSSIM model is a one-period model, typically one year. Because of the significant time delays in the recidivism process, the feedback version must be a multi-year model. Thus, a new user option has been added to the feedback version. The user can request a printout of selected results at the end of each intermediate year in the multi-year run, and he can examine these results before their impact (in the form of a new queue of recidivists) has been registered on subsequent years. This is in keeping with the interactive character of the JUSSIM planning tool that encourages the user to reexamine his assumptions if the results either do not seem plausible or satisfactory.

The most significant change in the operation of the model, however, are the new data demands placed on the user. Obviously the user must now come to grips with the problem of identifying virgin arrests, recidivism probabilities, recidivism time constants, and elements in the crime-switch matrix.[1]

This feedback version, JUSSIM II, is currently operating, but it has not been implemented as widely as JUSSIM I. This is partly because of its greater data demands and partly because its use requires somewhat more sophistication on the part of the user as a result of its more complex structure.

IMPLEMENTATION

As expected, the response to the interactive JUSSIM model was far more positive than to a previous batch-processed model from which it was derived. A number of states (e.g., California and Alaska), cities (e.g., Pittsburgh, Philadelphia, and Denver), and countries (e.g., Canada and Sweden) have organized data collection efforts to describe their systems as a base case for the JUSSIM model and plan to use a version of the model for their own planning. The Alaska model focuses on the corrections system and therefore represents the police and court subsystems as only a small number of stages, and therefore also provides greater detail (e.g., a separate stage for each institution) for the corrections subsystem. In some cases, staff changes in these organizations resulted in discontinuation in the use of these models.

An important use of the model is as a teaching tool for system planners. The Urban Systems Institute of Carnegie-Mellon University has used it in one-week short courses on systems analysis for criminal justice planning. The planners readily learned to operate the terminal, to respond to the program's interrogations, to translate their project ideas into judgments about model-parameter changes, and to operate the model effectively as a design tool. The model then plays the important role of forcing the planner to consider the entire system and the interdependency of its parts in his planning. One retired police officer commented, "This was the first time I've thought much about what happens to the people I arrested."

In Pittsburgh, three masters degree students at Carnegie-Mellon's School of Urban and Public Affairs developed a base-case data file of the Allegheny County criminal justice system over a three-month summer internship with the Allegheny County Regional Planning Council of the Pennsylvania Governor's Justice Commission. A follow-on student team, as part of a second-year project course, used that data base to perform a variety of analyses regarding potential improvements in the system.[2] New projects considered included bail reform, diversion of less serious cases from the higher court to the minor judiciary, and diversion from a juvenile detention home. Upon graduation, one of the students became a member of the Planning Staff of the Council. There he continued to work at integrating the model into the Council's planning process. Upon his departure after one year, the model was continuously maintained, used, updated, and extended by a staff member with no special technical training.

SOME TYPICAL USES OF JUSSIM

Once one has a JUSSIM model, there are a variety of kinds of uses to which it might be put. For example, it could be used to estimate the downstream implications of an incremental decision at any stage by calculating the average or expected cost per additional arrest (or reported crime, charge, conviction, etc.) taking into account the fact that some of the arrestees will be dismissed without charge, some will be acquitted in the courts, and some will penetrate through to a correctional institution. An average cost would reflect these various probabilities of penetration and the costs associated with each degree of penetration.

Users are often interested in the distribution of costs (or workload) associated with the various portions of the criminal justice system, since knowing those distributions in a jurisdiction might suggest a reallocation. Such a reallocation consideration might be influenced by comparing their status to national distributions.[3]

An important value of the JUSSIM model is the fact that it provides the "entry price" in the development of a statistical data base for a wide variety of research projects. Any single project may not be able to justify the cost of collecting the total base of information it needs, but the aggregate needs represented by a range of users who might then make use of that collected base could usually more than justify that expense. Even more important than the expense is the time and effort on the part of various potential users that would be required to collect their information, thereby effectively deterring any serious quantitative inquiry into the operation of the system. With the ready availability of a planning data base in a well-structured format such as JUSSIM, any potential user can inquire directly to obtain an estimate of a particular parameter or of any calculated value, or to obtain a quick estimate of the impact of one or another potential system change.

The JUSSIM model is particularly helpful in evaluating a wide variety of potential system changes. The user introduces these changes one at a time or in combination, and receives a quick estimate of the downstream implication of such changes. Different users, particularly those on different sides of a particular policy question, can come togehter to

develop estimates of the favorable and unfavorable values of parameter changes associated with a system revision. The cost, resource, and flow implications of those alternative estimates can then be calculated with the JUSSIM program. In many cases, the original basis of disagreement may disappear. Each of the advocates may develop new insights into some of the implications of the change he was endorsing or opposing; or the basis of disagreement may boil down to a particular parameter value, and that value might then be developed more carefully empirically, either by a test project in that jurisdiction or by obtaining estimates that may have been derived in another jurisdiction.

The JUSSIM program also has significant potential value for conducting cross-jurisdictional comparisons. Since the data are in a consistent, comparable format across jurisdictions--particularly within a state which has the same criminal justice structure--the State Planning Agencies (SPA) can generate data files for the different jurisdictions or regions within its purview and can conduct cross-jurisdictional comparisons by searching for jurisdictions which have deviant parameters.

As it develops new standards and goals procedures, each SPA will be seeking to find a means of assessing the cost and workload implications of the various standards and goals being considered by state or national commissions, or promulgated by them. One of the shortcomings identified in the reports of the National Advisory Commission on Criminal Justice Standards and Goals has been the absence of some of the cost implications associated with the standards promulgated. With a JUSSIM model for the state and its component jurisdictions, the SPA as well as the Standards and Goals Commission can begin to address some of those implications, particularly for those standards or goals that can be translated into branching ratio, workload, or unit cost parameters. A study currently underway by the American Bar Association Commission on Correctional Standards is using the JUSSIM model for precisely that purpose: to assess the cost implications of the correctional standards in the National Advisory Commission's report.

See the Appendix at the end of this chapter for a case study of a particular use of JUSSIM in Pennsylvania.

JUSSIM LIMITATIONS

It should be recognized that all models are finite and limited. The limits come from the inherent necessity to abstract reality rather than to represent it literally in any computational model. Given this, one could try to build an extremely elaborate and complex model which pushes toward an extensive representation of reality. Experience in such modeling efforts has shown that almost always such models fail to be very useful because of their elaborateness and complexity. Such elaborateness requires more data than can be reasonably obtained; the size of the model implies large costs per run of operation; and enormous complexity clouds the bases for an observed calculated result. Experience with modeling of other systems has made it clear that it is important to limit the size of any model.

This philosophy has been reflected in the JUSSIM model by limiting its size to 32,000 words of computer storage. Doing so makes the computer program accessible to most commercial time sharing computing services. More fundamentally, however, it restricts the complexity by restricting the number of crime types, stages, flow paths, resources, workloads, etc., that can be incorporated. If a user wants to increase the complexity of one kind of structure (to create more stages, say), then he must give up on another (decrease the number of resources being considered or the number of crime types, for example). Of course, a user with greater computer capacity who is prepared to deal with the greater complexity can readily go in and revise the dimensions of those parameters if he is prepared to deal with the more complex model.

In its operation, JUSSIM I (and the downstream portion of JUSSIM II) ignores the passage of time. In particular, this calculates the downstream impact of a crime in the year in which it was committed. Thus, a three-year sentence for a robbery committed in 1975 is counted in 1975, even though part of the time is served in 1976 and later. This would still give a good estimate of corrections cost if the system is not changing radically over time, since people sentenced in 1974 and earlier serve time in 1975, thereby roughly compensating for the 1975 crimes whose impact occurs in later years.

This structure makes the model simpler and much of the computation considerably easier and less expensive. But the structure does not allow the JUSSIM model to deal with queueing or delay problems, which explicitly focus on the passage of time and which are often of considerable concern to courts. There are a variety of "job shop simulation" programs which can be used much more effectively for this purpose. These include the General Purpose Simulation Systems (GPSS) models, and various other simulation programs that specifically address these delay issues. When he focuses on queueing issues, it would be far more efficient for a planner to use a separate model addressed to that purpose rather than to try to complicate a single model with all the complexity necessary to provide this wide range of flexibility.

REFERENCES

For those interested in more detail on the JUSSIM model than can be presented here, a number of references are available. The original flow model from which the JUSSIM model was derived was developed by Blumstein and Larson[1] as an outgrowth of the work of the Science and Technology Task Force [2] (especially Chapter 5) of the President's Commission on Law Enforcement and Administration of Justice. That original model existed in batch processing mode, which made it relatively inaccessible to most non-technical users. The JUSSIM model was re-programmed in an interactive mode by William Glass and is described in Belkin, Blumstein and Glass [3] and Blumstein [4]. Its particular implementation in Allegheny County was described by Cohen et al.[5] and in the juvenile justice system there by Blumstein and Stafford [6]. Other implementations have been described by Gordon Cassidy [7] for Canada, and by others. Sections of some of these references have been used in this paper.

User's Manuals for JUSSIM I (the downstream model) [8] and for JUSSIM II (the feedback model) [9] are available from the Urban Systems Institute at Carnegie-Mellon University. Copies of the JUSSIM programs, written in dialect-free Fortran IV and dimensioned to a computer with 32k words of core storage, are also available. The programs are well documented internally and contain a basic data file so that they can be operated immediately. Interested agencies can write for order forms to the Urban Systems Institute, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213.

Parts of the materials in this paper are derived from various segments of these earlier documents.

(1) Blumstein, A., and R. Larson. "Models of a Total Criminal Justice System," Operations Research, Vol. 17, No. 2 (March 1969), pp. 199-232.

(2) Task Force Report: Science & Technology, Report of the Task Force on Science and Technology of the President's Commission on Law Enforcement and Administration of Justice, June 1967.

(3) Belkin, Jacob, Alfred Blumstein, William Glass, and Michel Lettre. "JUSSIM, an Interactive Computer Program and Its Uses in Criminal Justice Planning," Proceedings, International Symposium on Criminal Justice Information & Statistics Systems, Gary Cooper, ed., Project SEARCH, October 1972, pp. 467-477.

(4) Blumstein, Alfred. "Management Science to Aid the Manager: An Example from the Criminal Justice System," Sloan Management Review, (Fall 1973), pp. 35-48.

(5) Cohen, Jacqueline, K. Fields, M. Lettre, R. Stafford, and Claire Walker. "Implementation of the JUSSIM Model in a Criminal Justice Planning Agency," Journal of Research in Crime and Delinquency, (June, 1973).

(6) Blumstein, A. and Richard Stafford. "Application of the JUSSIM Model to a Juvenile Justice System," Proceedings of the National Symposium on Computer Applications in the Juvenile Justice System, December 6-8, 1973, Atlanta, Georgia, pp. 60-84.

(7) Cassidy, R. Gordon, R. George Hopkinson, and William Laycock. "A Preliminary Description of the Canadian Criminal Justice System," Report of the Ministry of State for Urban Affairs and Ministry of the Solicitor General, Ottawa, Canada, October 1973.

(8) Belkin, Jacob, Alfred Blumstein and William Glass. "JUSSIM, An Interactive Computer Program for Analysis of Criminal Justice Systems," Urban Systems Institute, Carnegie-Mellon University, July 1971. (JUSSIM I User's Manual).

(9) Belkin, Jacob, Alfred Blumstein, and William Glass. "JUSSIM II, An Interactive Feedback Model for Criminal Justice Planning," Urban Systems Institute, Carnegie-Mellon University, January 1974 (JUSSIM II User's Manual).

NOTES

[1] Some preliminary estimations of values of these parameters are explored in: J. Belkin, A. Blumstein, and W. Glass, "Recidivism as a Feedback Process: An Analytical Model and Empirical Validation," Journal of Criminal Justice, I, No. 1 (March 1973).

[2] See: (1) Kenneth Fields, Michel A. Lettre, and Richard Stafford, "A Description of the Allegheny County Criminal Justice System," Urban Systems Institute, Carnegie-Mellon University, February 1972; (2) Jacqueline Cohen, M. Lettre, and R. Stafford, "Analysis of the Allegheny County Criminal Justice System: Present Operations and Alternative Programs," Urban Systems Institute, Carnegie-Mellon University, February 1972; (3) Jacqueline Cohen et al., "Implementation of the JUSSIM Model in a Criminal Justice Planning Agency," Journal of Research in Crime and Delinquency, June 1973.

[3] As reflected, for example, in "Expenditure and Employment Data for the Criminal Justice System: 1972-73," National Criminal Justice Information and Statistics Service Report #SD-EE No. 5, February 1975; Government Printing Office, Washington, D.C.

FOOTNOTE TO BLUMSTEIN'S "A MODEL TO AID IN PLANNING FOR THE CRIMINAL JUSTICE SYSTEM"

David L. Dougherty

The "approximate" nature of the JUSSIM model is further reinforced by the data collection process necessary in most jurisdictions to establish the base case for JUSSIM.

In Prince George's County, Maryland, where JUSSIM was established in 1974, considerable efforts were required to (1) obtain data from the entire system and (2) make the data comparable among agencies so that the model could operate. In Maryland, as in most states, some criminal justice agencies are entirely local and maintain their own statistics for management and planning, while others are regional or statewide and maintain statistics largely on other than local bases. In Prince George's County, the data collection problem was compounded by the existence of several municipal police departments with concurrent jurisdiction with a county-wide department which handled most criminal cases and all investigations.

For the development of the base case, it was necessary to obtain data from three levels of government. Municipalities provided reported crime and arrest data for a limited number of crimes. The County provided the majority of crime and arrest data from the police department, warrant service data from the shefiff's department, and Circuit Court (highest level court) processing data from both the Circuit Court and the State's Attorney. State agencies provided data from the District Court (where all defendants were initially charged and misdemeanants tried). The District Court's statewide data processing center provided information on pretrial detention. State agencies provided information concerning Public Defender operations on the county level, probation on the county level, and incarceration. All juvenile data was provided by the state Office of Juvenile Affairs.

In most cases, data was not automated and in other cases available automated data was reorganized by manual procedures for use by the model.

Given the variations in methods of collection, storage, and reporting of data by the individual agencies involved, a necessary first step was the development of an offense matrix through which to organize all crime classifications into standard groups which could be used by JUSSIM. UCR classification was used as the basis for groupings although some groups consisted of more than one UCR type. For example, the District Court did not distinguish between aggravated and simple assault; thus a single category of assault was used. Nor was there a distinction in the District Court data between auto theft and larceny; thus a single category of larceny was used. Further inconsistencies existed between UCR classifications and corrections data. Some additional combining was done because of the low volume of offenses of a particular type. The final number of groups established was 25 although the model can handle 30 groups. Further refinements of agency data collection processes could provide more detail among crime types in the future.

Having established a set of offense groups, data was collected based on that set, a process which called for special programming in the case of

automated data, or hand tallies when data was not automated or when special programming was not feasible.

Police data from UCRs was reorganized into the new crime types established and hand tallied. Special programming which required a limited amount of manual reorganization was used to generate District Court data for the county. An additional program was used with District Court data to provide detention information.

Circuit Court data was not automated on the county or state level and was thus hand tallied from docket books in the Court Clerk's Office into the categories established. Public Defender information was also hand tallied from attorney reports available in the state office of the Public Defender.

Data concerning probation and parole and incarceration was available from automated reports but required additional manual organization for use.

Due to the time necessary for a defendant to move through the system from arrest to disposition and complete sentence, data for the periods covered by the base case did not, in all cases, involve the same defendants; thus it was necessary to further justify the data between agencies. This was done by establishing the number of crimes reported as the base figure and calculating the branching ratios according to actual data available, which was largely a manual task.

It thus becomes essential to review the input as well as the model's output with line agencies to assure that the required manipulations of the data do not vary too far from reality while recognizing that the model does not provide an actual picture of the operation of the system.

APPENDIX. A CASE STUDY--ALLEGHENY COUNTY, PENNSYLVANIA*

The practical utility of the JUSSIM model can perhaps best be demonstrated by citing an example of its implementation in a local CJS planning agency. Allegheny County is located in western Pennsylvania and includes the city of Pittsburgh. The total population of the county is about 1,600,000. In 1969, a Regional Planning Council was established there under the auspices of the Pennsylvania Governor's Justice Commission. This council was charged with the responsibility of developing plans to improve the CJS in Allegheny County (AC). The council itself is composed of various officials from different parts of the CJS and some "citizen members." There is also a full-time professional planning staff.

The JUSSIM model was applied to AC and the county's CJS was described in a data file for the JUSSIM computer program. The research involved collecting and refining the necessary data and exploring various proposals for improving the county's CJS. This discussion includes only a sample of the full analyses provided to the planning council.[1]

The Model of the AC-CJS

A simplified flow diagram of the AC criminal justice system is presented in Figure 1. This chart indicates the processing of all crimes committed in the county which come to the attention of the CJS. The process is initiated by a crime report or police arrest and is terminated in the corrections subsystem. The blocks on the diagram represent processing stages in the system and the arrows indicate the flow paths. The quantities indicated on the chart are

1. Kenneth, Fields, Michel A. Lettre, and Richard Stafford, "A Description of the Allegheny County Criminal Justice System" and Jacqueline Cohen, Michel Lettre, and Richard Stafford, "Analysis of the Allegheny County Criminal Justice System: Present Operations and Alternative Programs," Urban Systems Institute Reports, School of Urban and Public Affairs, Carnegie-Mellon University, Pittsburgh, 1972.

*The material in this section is quoted from "Implementation of the JUSSIM Model in a Criminal Justice Planning Agency," by Jacqueline Cohen, K. Fields, M. Lettre, R. Stafford and Claire Walker; Journal of Research in Crime and Delinquency, June 1973.

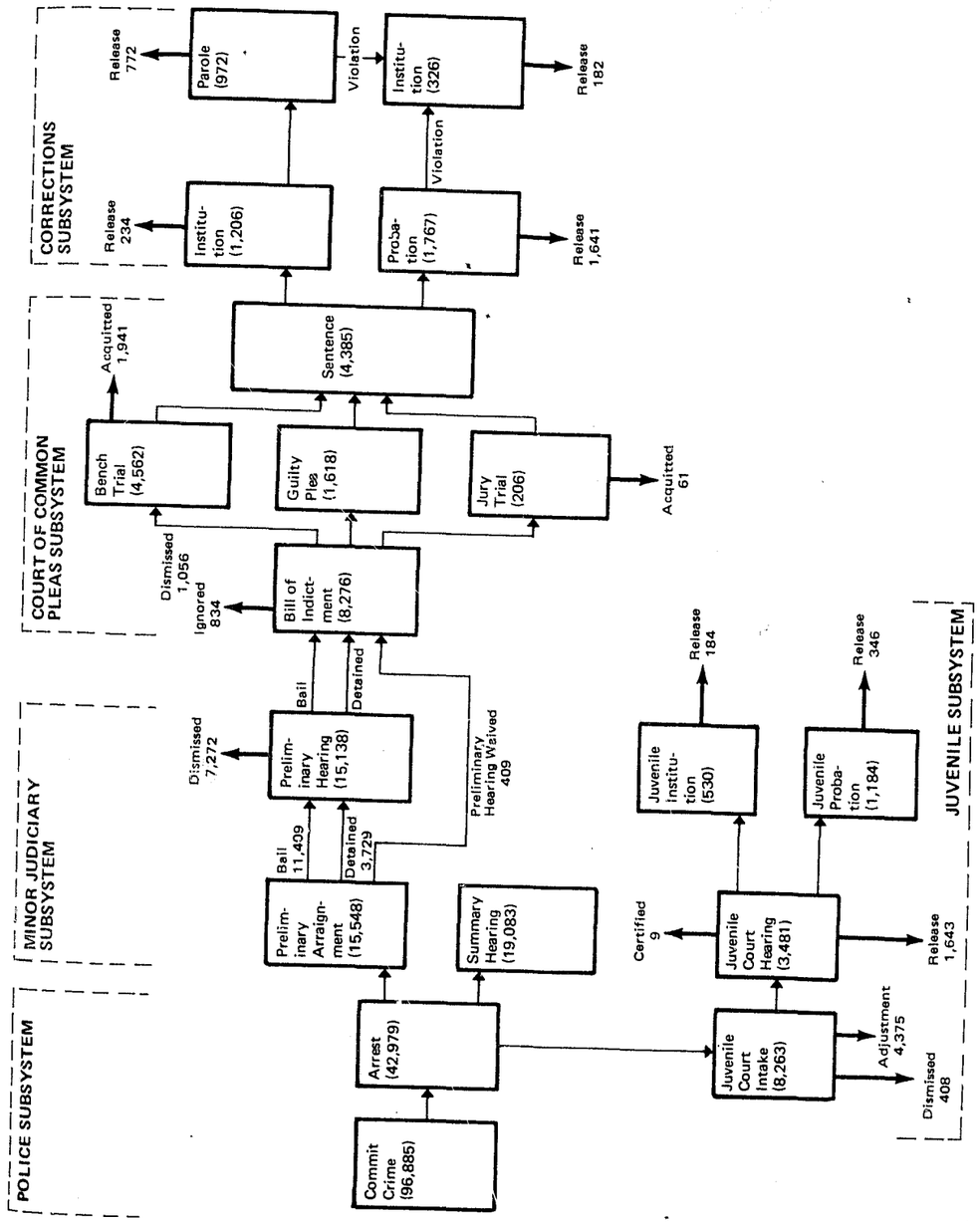


Figure 1. FLOW DIAGRAM OF THE ALLEGHENY COUNTY CRIMINAL JUSTICE SYSTEM

the estimated annual flows through each processing stage for the county.

The AC-CJS is made up of five major subsystems: the police, the minor judiciary, the Court of Common Pleas, adult correction, and the juvenile subsystem. A unit of work typically enters the system in the police subsystem in the form of an arrested individual. After arrest, juveniles are referred to the Juvenile Court and enter the juvenile subsystem while adults are referred to hearings in the minor judiciary subsystem.

All hearings in the minor judiciary subsystem are conducted by magistrates. Those adults arrested for minor summary offenses receive summary hearings before a magistrate. At the summary hearing the magistrate determines the guilt or innocence of the defendent and sentences guilty offenders. Adults charged with the more serious indictable offenses are arraigned before a magistrate and bail and counsel decisions are made. At a subsequent preliminary hearing, the magistrate decides whether there is a prima facie case against the defendant. The defendant may be dismissed at this point or held for the Court of Common Pleas.

If the defendant is held for court, a bill of indictment is prepared and sent to the grand jury. Indictments upheld by the grand jury are sent to the court. Before trial the defendant is arraigned in court where the charges may be dismissed (indictment quashed or demurrer sustained) or the D.A. may choose not to prosecute. If the case goes to trial the defendant may plead guilty or elect to be tried by a judge or jury.

Convicted offenders may be sentenced to probation or to a state institution. The persons in institutions can then be paroled before completing their sentences. If the conditions of probation or parole are violated, the individual may be sent to a state institution. When the term of the sentence is completed the convicted offender is unconditionally released back into the community.

The juvenile subsystem is generally less formal than the adult subsystem. Almost all cases referred to Juvenile Court are first reviewed by a probation officer. After this review, the probation officer may deal nonpunitively with the case himself or hold the case for a hearing in Juvenile Court. At the hearing the judge can release the juvenile or refer him to a correctional agency.

Various data were needed to quantify this structural description of the AC-CJS. These included the flows through each stage, the branching ratios, the workloads, the resource availabilities, and the unit costs. These data were collected from a variety of different sources. Exact figures were often obtained from the office files, annual reports, and budgets of the various agencies. Lacking these sources, many of the officials and employees of the system were able to supply reasonable estimates. These were the primary sources of data; the secondary sources included estimates from other jurisdictions[2] and

Table 1. SUMMARY OF DATA ON RESOURCES UTILIZED IN THE ALLEGHENY COUNTY CRIMINAL JUSTICE SYSTEM

Resource Units	Cost/Unit	Total Units Available in a Year Per Individual Resource	Stages Where Resources Are Applied	Source of Data
County Detective Hours	$ 7.16	1,700	2, 3	Primary
County Patrolmen Hours	6.75	1,700	3	Primary
Pittsburgh Detective Hours	8.42	1,700	2, 3	Primary
Pittsburgh Patrolmen Hours	7.61	1,700	3	Primary
District Magistrate Hours	16.47	542	4, 6, 7, 8	Secondary
Pittsburgh Police Magistrate Hours	55.38	542	6, 7, 8	Primary
Detention Days	7.00	365	7, 8, 13	Primary
Common Pleas Judge Days	860.50	211	11-14	Primary
Public Defender Days	118.00	211	11-14	Secondary
District Attorney Prosecutor Days	100.00	211	11-14	Secondary
District Attorney Indictment Hours	13.11	1,610	9, 10	Secondary
Grand Jury Hours	129.39	1,266	10	Secondary
Behavior Clinic Days	387.00	230	—	Primary
Pre-Sentence Investigation Days	62.82	230	—	Secondary
County Probation Officer Years	69.00	35	16	Secondary
State Parole Officer Years	480.00	35	19	Secondary
State Institution Years	4,438.00	1	17, 18	Primary
Juvenile Intake Officer Days	76.67	230	20	Secondary
Juvenile Judge Days	480.20	211	21	Primary
Juvenile Prosecution Days	76.37	115	21	Secondary
Juvenile Probation Officer Years	231.88	35	20, 23	Secondary
Juvenile Detention Days	18.95	365	20, 21	Primary
Juvenile Institution Years	9,621.40	1	22	Primary

2. When primary data were not available for the analysis of Allegheny County, estimates from Los Angeles and Connecticut were often used. The Los Angeles information is reported in Alfred Blumstein and Richard Larson, "Models of a Total Criminal Justice System," in Systems Approach and the City, M.D. Mesarovic and R. Reisman, eds. (New York: American Elsevier, 1972). pp. 209-51. The Connecticut data are found in Jacob Belkin and Alfred Blumstein, "Methodology for the Analysis of Total Criminal Justice Systems," Urban Systems Institute Working Paper, School of Urban and Public Affairs, Carnegie-Mellon University, Pittsburgh, 1970.

various combinations of primary and secondary sources. Table 1 summarizes much of this basic input data for the resources in Allegheny County. Each resource is listed along with its unit cost and annual availability. Wherever possible, the unit costs include the costs of administration and backup personnel. The table also indicates the processing stages where the resource is utlized and the source of the data.

Analysis of the AC-CJS

With this basic data on CJS operations, the JUSSIM model can be used to generate a quantitative description of the present system. This description includes the distribution of system and subsystem costs, the degree of flow penetration, and the resource requirements, thus providing the planning council with an empirical basis for examining the operations of the various agencies in the CJS.

For example, Table 2 shows the distribution of the total adult system costs among the major subsystems. As noted, the correction subsystem accounts for almost two-thirds (63.9 percent) of total adult CJS costs.[3] Depending on the crime type, the major cost

Table 2. DISTRIBUTION OF TOTAL ADULT COST AMONG THE MAJOR SUBSYSTEMS FOR SELECTED CRIME TYPES

	Percent of Total Adult Cost for Police	Percent of Total Adult Cost for Court	Percent of Total Adult Cost for Correction	Total Adult Cost
All Crime Types	14.3	21.7	63.9	$17,085,900.
Murder and Non-negligent Manslaughter	1.3	8.1	90.5	2,237,700.
Forcible Rape	5.8	22.5	71.6	386,500.
Robbery	4.3	12.1	83.6	2,774,800.
Aggravated Assault	13.6	33.1	53.3	933,000.
Burglary	12.3	22.2	65.5	3,033,600.
Auto Theft	45.1	40.9	14.0	344,400.
Simple Assault	6.1	34.0	59.9	295,000.
Stolen Property	10.6	18.4	71.0	310,000.
Commercial Vice	16.5	14.6	68.9	318,600.
Narcotics	10.0	19.7	70.3	2,143,300.
Drunk Driving	3.0	70.9	26.1	391,100.
Disorderly Conduct	31.0	56.2	12.8	113,400.

3. It should be noted that Table 2 represents the distribution of only those system costs directly attributable to criminal reports and arrests. For example, the police cost here includes only the cost of dealing directly with specific crimes (e.g., responding to calls, making investigations and arrests). The cost of maintaining patrols and providing other public services like traffic control are not included.

component can be police (e.g., auto theft), courts (e.g., drunk driving), or correction, (e.g., robbery). Table 2 also shows the large system costs associated with robbery and burglary which account for one-third of the total system cost.

The court subsystem in Allegheny County includes the minor judiciary, detention, and the Court of Common Pleas. A summary of the flows and costs of this court subsystem is shown in Table 3. The first column indicates the percentage of those arrestees appearing at the minor judiciary who are subsequently held for court for selected crime types. In all cases those not held for court are either disposed of at a summary hearing or dismissed at the preliminary hearing.

Column 2 of Table 3 shows the input to the court expressed as a percentage of the total arrests for each selected crime type. These percentages indicate the probability of penetrating to the Court of Common Pleas after arrest, and they reflect the effect of both the minor judiciary and the Juvenile Court in limiting the flow of arrested individuals into the Court of Common Pleas. It is clear that with the exceptions of murder and narcotics cases, the majority of arrested individuals never reach trial in the Court of Common Pleas.

The remainder of the table presents the distribution of total court costs among the magistrates, detention, and the Court of Common Pleas. The Court of Common Pleas is the most expensive component of this subsystem. Its costs are generally eight times

Table 3. FLOWS AND COSTS IN THE ADULT COURT SUBSYSTEM FOR SELECTED CRIME TYPES

Crime Type	Percent of Cases at Magistrates' Hearings Held for Court	Percent of Arrestees Disposed of in Court of Common Pleas	Total Court Cost	Percent of Total Court Costs		
				Minor Judiciary	Detention	Court of Common Pleas
All Crime Types	22.7%	19.2%	$3,712,600	8.0%	25.5%	66.5%
Murder and Non-negligent Manslaughter	100.0	100.0	182,300	4.7	47.1	48.2
Robbery	64.5	40.9	334,400	2.6	25.4	72.0
Aggravated Assault	39.8	39.9	308,600	5.7	12.9	81.4
Burglary	69.2	36.2	672,500	2.9	26.2	70.9
Larceny	31.5	14.0	279,100	4.2	29.8	66.0
Stolen Property	53.5	42.1	57,100	7.0	29.8	63.2
Weapons	55.3	48.1	45,000	8.0	26.9	65.1
Narcotics	64.2	55.7	421,800	6.8	12.9	80.3
Gambling	35.8	44.9	54,100	14.6	0.0	85.4

the costs of the minor judiciary, even though the magistrates handle more cases than the Court. Detention costs also constitute a sizable portion (26 percent) of the total court cost.

Tables 4 and 5 illustrate the chances of penetrating to the correction subsystem after entering the CJS. Table 4 presents the proportion of arrested adults sentenced to probation or incarceration for several representative crime types. The chances are relatively high for the serious property crimes (.54 for robbery, .51 for burglary) and quite low for vice crimes (e.g., .01 for gambling). Table 5 presents the percentage of juveniles committed to a juvenile correctional institution after referral to Juvenile Court for several crime types. Surprisingly, the probability of being institutionalized appears to be highest for the strictly juvenile offenses of truancy and ungovernable behavior--perhaps because the court often suppresses more serious charges so that the juvenile will not have a record of criminal charges against him. If this is the case, then the juvenile is really being institutionalized for more serious criminal acts, although formally charged with a minor juvenile offense.

The data on the detention or bail status of arrestees were also analyzed. They showed that the probability of dismissal at the preliminary hearing is higher for arrestees on bail than for those held in detention. Over all crime types, the probability of dismissal for arrestees on bail is .521, while it is only .356 for arrestees held in detention. There are at least four hypotheses that could explain the phenomenon:

Table 4. PROPORTION OF ARRESTED ADULTS SENTENCED TO PROBATION OR INCARCERATION FOR VARIOUS CRIME TYPES

Crime Type	Adult Arrests	Number of Adults Sentenced to Probation or Incarceration	Percent of Arrested Adults Being Sentenced to Probation or Incarceration
Murder and Non-negligent Manslaughter	86	58	67%
Robbery	464	249	54
Aggravated Assault	1,011	204	20
Burglary	1,103	560	51
Larceny	1,038	189	18
Stolen Property	373	88	24
Weapons	383	49	13
Narcotics	1,687	623	37
Gambling	774	8	1
All Others	25,413	945	4

1. Those arrestees out on bail have both the money and the time to secure counsel for their defense, thereby increasing the probability of dismissal.

2. The police magistrates make a bail/detention decision at the preliminary arraignment that places in detention those arrestees who are more likely to have committed the crime and more likely to jump bail if it is made available to them.

3. Arrested persons who know they are guilty are more likely to accept detention, with the expectation of getting a credit on the sentence for the time served in detention.

4. There is a stigma associated with detention (whether it is caused by the arrestee's indigence or his potential risk out on bail) that increases the probability that a magistrate will hold an individual for trial.

While the model does not suggest the reason for the occurrence of this phenomenon, the process of data collection and analysis reported here fosters identification of this kind of phenomenon. The explanation must be sought through a more detailed analysis of individual cases.

This type of descriptive analysis is often the basis for new insights into the functioning of the system. It also documents the extent of

Table 5. INSTITUTIONALIZATION OF JUVENILES FOR VARIOUS CRIME TYPES

Crime Type	Total Juveniles Referred to Juvenile Court	Number of Juveniles Committed to an Institution	Percent of Juveniles Committed to an Institution
All Crime Types	8,263	530	6.4
Aggravated Assault	262	32	12.2
Burglary	823	61	7.4
Larceny	930	29	3.1
Auto Theft	789	96	12.2
Simple Assault	558	12	2.2
Vandalism	607	9	1.5
Narcotics	502	29	5.8
Disorderly Conduct	1,155	12	1.0
Runaway	637	38	6.0
Truancy	230	46	20.0
Ungovernable	533	102	19.1
All Others	1,237	65	5.3

known problems in system operations. In either case the results of the analysis frequently suggest changes that might improve the operations in question. These changes can then be "tested" with the JUSSIM model. By using the extensive data base and the analytical methods provided in JUSSIM, planners can anticipate and evaluate the indirect and system-wide effects of any change and, with this kind of information, can compare the merits of various programs and make choices among them.

The problem of bail reform has been selected to demonstrate JUSSIM's usefulness as a planning tool. In 1970, approximately 3,800 persons from Allegheny County were detained for some period of time, and another 9,000 were released on surety bond after paying a premium to a bondsman. There are a great many disadvantages associated with this bail/

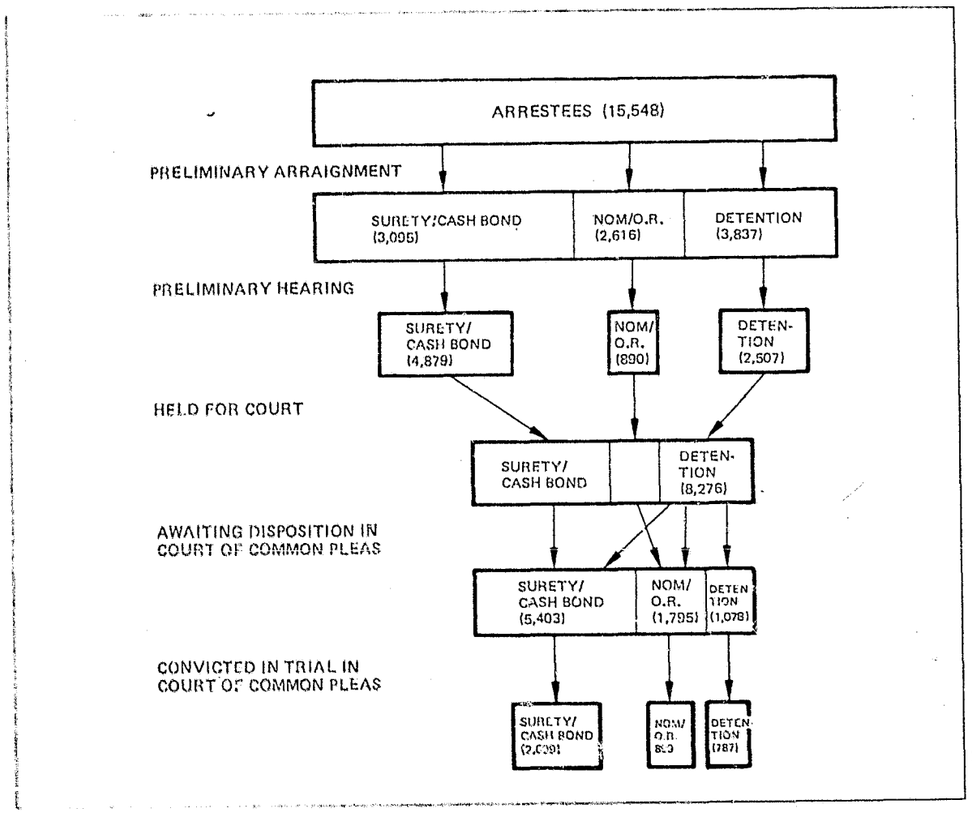


Figure 2. ESTIMATED FLOW OF DEFENDANTS IN ALLEGHENY COUNTY IN 1970

The problem of bail reform has been selected to demonstrate JUSSIM's usefulness as a planning tool. In 1970, approximately 3,800 persons from Allegheny County were detained for some period of time, and another 9,000 were released on surety bond after paying a premium to a bondsman. There are a great many disadvantages associated with this bail/ detention system. For those detained in jail these include the loss of liberty, the loss of one's income and perhaps one's job, the disruption of family life, and the stigma of guilt often associated with being in jail; for those released on bail, there is the cost of the bond premium. While it is often difficult to measure all the social costs associated with these disadvantages, some of the economic costs can be estimated. The JUSSIM flow and cost data are particularly useful in this analysis.

Figure 2 is a graphic representation of the estimated defendant flow from preliminary arraignment to trial in Allegheny County. There are three defendant statuses at each step in the process: (1) released on surety or cash bond, (2) released on nominal bond or own recognizance, or (3) detained. At present there are four points where an arrested individual can be released from detention: (1) at the preliminary arraignment by posting bail or being released on nominal bond, (2) at the preliminary hearing through the dismissal of charges or release on nominal bond or bail, (3) sometime before trial by posting bail or on nominal bond after filing a special petition to the court, and (4) after trial through an acquittal.

Of the 15,548 defendants at preliminary arraignment, 16.8 percent (or 2,616) are released on nominal bond or their own recognizance (O.R.). 58.5 percent (or 9,095) are released on surety or cash bond, and 24.7 percent (or 3,837) are held in detention until their preliminary hearing. Only 8,276 of the original 15,548 defendants are held for court after their preliminary hearing. Of those held for court 30.3 percent (or 2,507) are detained at Allegheny County Jail. 10.8 percent (or 890) are released on nominal bond or O.R. and 59.0 percent (or 4,879) are released on surety or cash bond. This released population of 5,769 is ultimately augmented by another 1,429 defendants who are released from jail on nominal or surety bond at some time before their disposition in the Court of Common Pleas. This additional release of defendants changes the defendant status at the time of trial to 21.7 percent, 65.3 percent, and 13.0 percent on nominal bond, surety bond, and in

detention, respectively. This means that of the original 3,837 defendants detained by magistrates after preliminary arraignment, 2,759 (or 71.9 percent) are later released either through the dismissal of charges or on some form of bond or recognizance before trial. In addition, there are an estimated 291 persons detained from the time of preliminary arraignment right through trial in the Court of Common Pleas who are subsequently acquitted at trial.

If the current procedures for gaining information about defendants were improved, many of the defendants now released after spending some time in detention could be released much earlier. This earlier release of defendants would represent a substantial reduction in the detention population and in the costs of detention. These potential savings in detention costs can be estimated from the flow and cost data in the JUSSIM model. The results are presented in Table 6.

If all the defendants currently released from detention were released immediately after their preliminary arraignment, the annual detention population would be reduced by almost 80 percent and the savings in detention costs would be over $475,000. It is, however, doubtful that even the best bail release program would be able to select these 3,050 defendants from the 3,837 presently detained. Nevertheless, a reasonably achievable bound on the detention population after preliminary arraignment

Table 6. POTENTIAL DETENTION COST SAVING IF ALL CURRENTLY RELEASED DEFENDANTS WERE RELEASED AT PRELIMINARY ARRAIGNMENT

Current Point of Release	Number of Defendants Released	Average Number of Detention Days Per Detentioner	Detention Cost Savings (at $7.00 per Day)
a. Dismissed at Preliminary Hearing	1,330	6	$ 55,860.
b. Bailed Before Trial	1,429	30	300,090.
c. Acquitted in Trial	291	60	122,220.
Total	3,050		$478,170.

might be the current percentage of defendants in detention at the time of trial--namely, 13 percent (see Figure 2). This would decrease the present detention population by 1.816 defendants. Assuming that these defendants come proportionately from the three populations in Table 6, the resulting savings in annual detention costs would be about $284,000.

The same flow data can also be used to estimate the annual private expenditure on bail premiums. At present a total of approximately 6,920 of the 9,619 defendants released on surety bond are subsequently dismissed either at preliminary hearing or at trial (e.g., indictment quashed, nolle prossed, or acquitted). Using the district attorney's bail schedule we estimated that these persons pay an average of $102 each in nonrefunded bond premiums, or a total of $705,840 annually.[4]

Almost all of this private cost of release for dismissed defendants can be avoided by changes in the present bail procedure. There are essentially two strategies likely to affect the defendant expenditure on bond premiums. Insuring the early discovery of information about the defendants can increase the number of defendants released on nominal bond or on their own recognizance. The court could also implement a cash bond system where the defendant pays an 8 percent premium to the court, to be refunded (less a small service charge) if he is dismissed. If he is convicted, the premium goes toward the payment of court costs. The combined private savings in bail premiums and public savings in detention costs from changes in the bail procedures would be almost $1,050,000 annually.

The implementation of these bail reform strategies would require additional manpower expenditures by the court. Defendants must be interviewed to obtain the information needed to decide on their release on nominal or cash bond. This information must then be verified. Also, there must be some follow-up for defendants released on nominal bond to assure their appearance in court. This will include written notification and telephone contacts about court appearances.

4. The average premium of $102 per defendant was estimated by taking 8 percent of the weighted average of the current bail schedule established by the D.A.'s office.

There are several alternative bail programs that could be implemented. Among them is a program to extend the use of nominal bond. This program would primarily affect indigent defendants. Alternatively, a court-operated cash deposit system could be implemented for those defendants presently eligible for surety bond. And finally, the program to extend the use of nominal bond could be combined with the court-operated cash deposit system. In addition to the redeeming social merits of these programs, each also has some economic merits which can be explored with the JUSSIM model.

The potential savings and costs associated with each program can be determined from existing flow and cost data with the anticipated workload data for the programs. A typical bail-release program would require at most the following average workloads for a bailable defendant: .5 hours for interview, 1.0 hours for verification, and 1.5 hours for follow-up. It can also be assumed that the annual cost for each new resource in the bail program will be $12,500 and that each resource is available 1,840 hours per year (230 days/year X 8 hours/day). When these assumptions are combined with the present flow data, the expected annual cost of a new program can be calculated.

Table 7 presents a comparison of the costs and savings, both public and private, associated with each of the three bail programs. The highest ratio of total savings to costs is found for the court-operated cash deposit system (alternative 2), although all the savings here are private with no public savings. This program would save the cost of nonrefunded bond premiums for many defendants currently released on surety bond. However, it would not aid those indigent defendants who will still be unable to raise the cash deposit required. The combined program (alternative 3) has the dual advantage of affecting both indigent defendants and those on surety bond while also having the highest net total savings of any of the programs ($850,500 compared to $354,500 and $537,000). This information was made available to the planning council to aid in its deliberations about bail reform in the county.

Table 7. POTENTIAL SAVINGS AND PROGRAM COSTS OF NEW BAIL RELEASE PROGRAMS

Program	Defendants Affected	Defendants to be interviewed	Potential Public and Private Savings	Expected Costs
1. PROGRAM TO AFFECT INDIGENTS	Earlier Release of Currently Released Detentioners	2,759	$ 356,000. (det. costs)	
	Detained through Trial	1,078	–	
	Extended Nominal Bond to 40% of Bailable Population at Arraignment	11,711	211,000.*	
	TOTAL	15,548	$ 567,000.	$212,500. (for 17 resources)
2. COURT-OPERATED CASH DEPOSIT SYSTEM FOR SURETY DEFENDENTS	Dismissed at Preliminary Hearing or Trial	6,920	$ 637,000.*	
	Convicted	2,699	–	
	TOTAL	9,619	$ 637,000.	$100,000. (for 8 resources)
3. EXTEND NOMINAL BOND AND CASH DEPOSIT SYSTEM	Earlier Release of Currently Released Detentioners	2,759	$ 356,000. (det. costs)	
		(262)	24,000.*	
	Detained through Trial	1,078	–	
	Extended Nominal Bond to 40% of Bailable Population at Arraignment	4,684	$ 211,000.*	
	Surety Defendants Dismissed at Preliminary Hearing or Trial	5,129	472,000.*	
	Convicted Surety Defendants	1,898	–	
	TOTAL	15,548	$1,063,000.	$212,500. (for 17 resources)

* for prem. costs

Part IV.

USE OF POPULATION DATA

CHAPTER 14

DATA USE TOOLS AND TECHNIQUES AVAILABLE TO CRIMINAL JUSTICE PROFESSIONALS

Ronald E. Crellin, George L. Farnsworth,
and Leo Schuerman

INTRODUCTION

The full potential of available data use tools and technologies, particularly as they relate to small areas, has yet to be fully realized by professionals working in criminal justice planning. The purpose of this chapter is:

- To describe the available computer based tools and techniques for processing, structuring, linking, analyzing and displaying small area data (census tracts and specialized subcity administrative districts).

- To provide a selected reference to these products so that potential users have a useful guide to their accessibility and use.

GEOGRAPHY

Geography is an important element associated with most items of data. Any official involved with the distribution of services, public or private, must also be concerned with the distribution of the potential client population in relation to the supply and demand for these services.

Having current and accurate data on a timely basis by appropriate geographic areas is critical to the decisionmaking process. One of the problems in obtaining comprehensive data profiles of geographic areas of interest and cross tabulating administrative data with socio-economic data from the census is the incompatibility of geographic areas for which data is collected and reported.

In the criminal justice field in particular, it is critical to know precisely at a point in time

- what happens
  - to whom
    - at what location
      - with what impact

and how this compares to what happened at some prior time and to predict what is likely to occur at some future point in time. Only then can effective decisions be made about what to do, what resources to provide, and what combination of services to deliver.

National planners find access to national and state level data relatively easy. Data at the substate level, and particularly at the city and subcity levels, is becoming increasingly important, however, because of the need to distribute funds through county, regional, and local agencies.

Regions for which census data are of interest may include:

- The nation
- A state
- Standard Metropolitan Statistical Area (SMSA): one of 243 geographic units in the nation, including the entire area in and around a city that forms an integrated economic and social system.
- A county (or in Lousiana a parish)
- A city
- A census tract: small areas into which cities and their adjacent areas have been divided for statistical purposes
- A city block group: a region within a city that includes more than one city block and forms a unit
- A city block: a well-defined rectangular or irregular piece of land bounded by streets, roads, streams, railroad tracks or other features

Other units of interest may include a school district or police precinct. Definitions of these and other related statistical units are contained in the Glossary at the conclusion of this Chapter.

Although census data is aggregated to the census tract or block level, local data is frequently aggregated to individual special purpose areas such as health district, school district or police precinct. Such geographic disagreements cause problems for potential data users. Theoretically, data could be hand coded to any geographic areas the user requires.

Practically, with the large volume of data that planners are now required to handle, computer processing is the best way that comparisons can be done economically, efficiently and accurately.

To solve the problem of aggregating data from different sources to common geographic areas, the computer must know which addresses are in which census tract, school district or police precinct.

Just as decisionmakers need maps to do manual geographic analysis, the computer needs a computer-readable map to perform these operations automatically.

DUAL INDEPENDENT MAPS ENCODING (DIME) FILE

Computer-readable maps are already available for over 200 Standard Metropolitan Statistical Areas (SMSA's) throughout the country. What is needed are:

- Computers
- Computer Programs
- Resources
- Trained Manpower

plus an understanding of how to apply this new technology to criminal justice planning.

Computers are accessible to most public agencies, either inhouse or through remote terminals. Computer programs are available in the public domain. Resources can be obtained from various sources for sound projects. Trained manpower with the necessary combination of skills and experience is difficult to find, at the local level.

The underpinning that allowed the development of the DIME technology is a computerized street map called a Geographic Base File (GBF)/DIME (Dual Independent Map Encoding) file. Auxiliary computer programs match addresses with areas such as census tracts identified in the DIME file.

DIME files have been created for more than 200 SMSA's. A list of these SMSA's is found in the appendix at end of this Chapter.

To aid potential users of these data products and technologies in the solution of service delivery problems in the criminal justice field, the following section will include a brief description of the computer programs for sorting, structuring, linking and displaying data from a variety of sources.

The GBF/DIME-File can be a useful tool for local governments and other organizations concerned with small area data. This reference file, developed by the Census Bureau, provides a means to efficiently relate local data to census data or to other local data through address information.

GBF/DIME stands for Geographic Base File/Dual Independent Map Encoding and refers to the computerized files created by the DIME process. This approach combines address information with information sufficient and necessary to describe the urban street network. By considering each street as a series of lines and each intersection as a node, the entire region covered by the file can be viewed as a series of interrelated nodes, lines and enclosed areas. Other features, such as streams or jurisdiction boundaries, may also be defined in terms of segments and nodes.

In creating a GBF/DIME-File, each distinctive element represented on a map is examined and, at points where streets or other features intersect, end, or curve sharply, a node number is assigned, as seen in Figure 1.

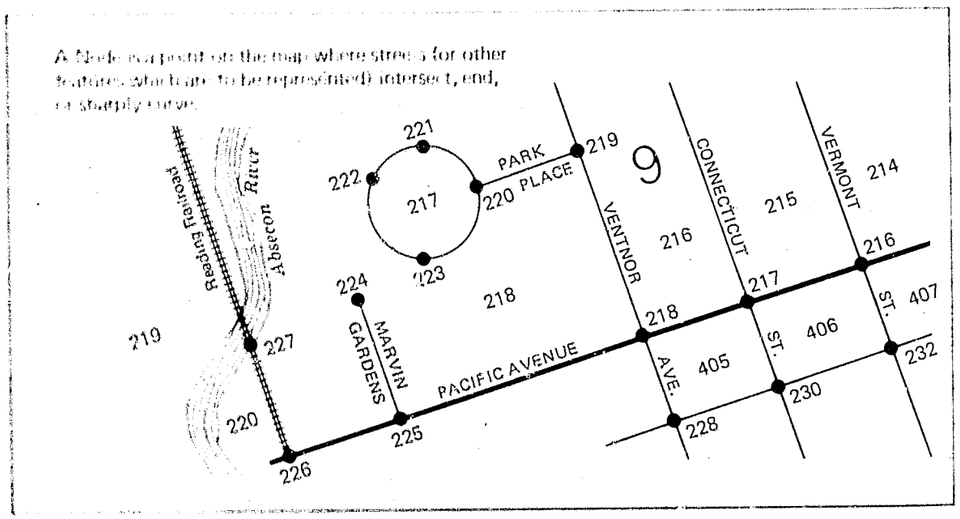


Figure 1. NUMBERING THE NODES

For each segment of a street (i.e., the length between two nodes) the DIME file contains the following information: The "from" node, the "to" node, the street type (e.g. avenue, drive, way, etc.) the address range on the right-hand side of the street and the address range on the left-hand side.

Moreover, each block and census tract is uniquely numbered. These geographic areas (on the left- and right-hand sides of the street segment being defined) are also included for each record in the GBF/DIME-File.

Figure 2 shows some of the information contained in the GBF/DIME-File for the 100 block of Atlantic Avenue.

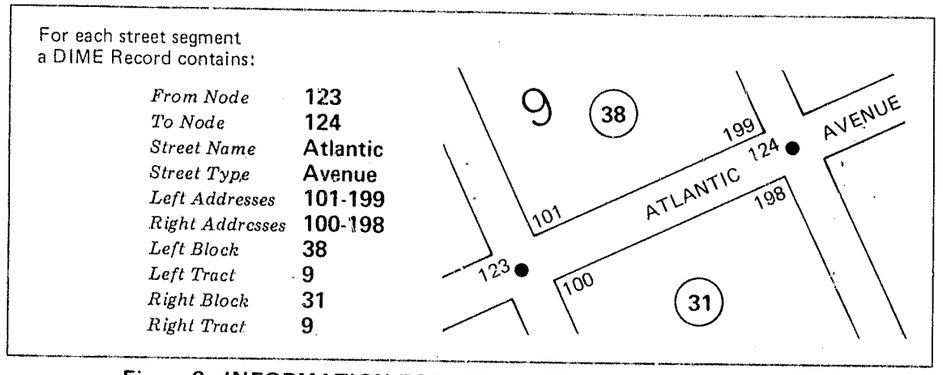


Figure 2. INFORMATION FOR 100 BLOCK OF ATLANTIC AVENUE

The term DIME then refers to the fact that the basic file is created by coding two independent matrices: a) the nodes at the end of the line segments and, b) the areas enclosed by the nodes and line segments.

The computer constructs these two independent networks--one of line boundaries and one of areas--and matches them (Figure 3).

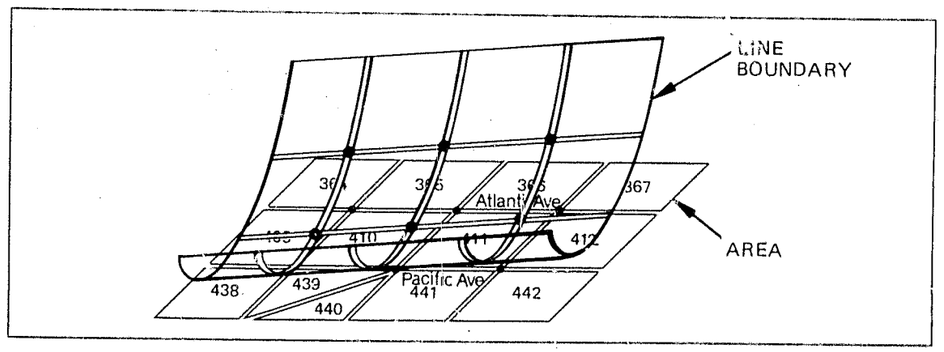


Figure 3. MATCHED NETWORKS OF LINE BOUNDARIES AND AREAS

This will ensure that the resultant network is completely represented and all the land is accounted for.

Once a map is digitized, the computer can use this information to plot a complete replica of the source map.

Dime Applications

DIME, which grew out of the research efforts of the Census Bureau, has been found to have applications for local users as well. State and local governments, and private data users have discovered that the GBF/DIME-File could be used for geocoding and integrating their own files and records. Moreover, local data could be efficiently related to census data at various geographic levels and information could be easily retrieved by the computer.

DIME has been used by local agencies in determing the best location for child day-care centers, in restructuring police beats, in implementing carpool programs, in designing school bus routes. Numerous applications of DIME technology have also been reported by specialists in transportation, public health, urban and regional planning, law enforcement and criminal justice communications, marketing, social welfare, political administration and other fields.

Available Computer Programs

Many computer programs for use with DIME were developed by the Bureau of the Census. However, the Program for Data Research (PDR), of The Social Science Research Institute at the University of Southern California continues program maintenance and documentation and makes programs available and provides assistance to interested users. A list of these programs and their applications follows:

ADMATCH — ADMATCH is a computer program package designed for uses in assigning geographic codes to local records using a DIME, ACG, or similar geographic base file. The system includes a preprocessor, which standardizes street addresses subject to variability, and a matcher, which links geographic base files to local data files. ADMATCH is used for address related geocoding. This involves adding coordinates, census designations, or local geographic codes to local data files containing house addresses.

The ADMATCH systems, both OS and DOS were written in IBM System/360 Assembler Language and are intended for users of IBM System/360 computers. The ADMATCH/OS system is still available but has largely been superceded by ZIPSTAN/UNIMATCH listed below.

A COBOL version of ADMATCH, while not prepared by PDR staff, is available but is somewhat less efficient than the ALC versions.

C-MAP — C-MAP is a relatively simple program written in Basic FORTRAN for producing choropleth maps. Input is via control cards plus a scan-line file which may be hand coded or automatically generated from DIME files using the DACS and SCANGEN programs. C-MAP produces maps up to 5 printer sheets in width and any desired length. While not containing all of the options of SYMAP, it is much more efficient both in terms of core utilization and machine time.

CARPOL — The CARPOL computer program is designed as a large-scale approach to automated carpooling. The program generates lists of potential fellow-riders from which a candidate can create his carpool group.

The program has a primary search radius based on geographic areas and a secondary search based on non-geographic criteria such as common workdays, hours, and driver/rider specifications. While developed for use with the DIME geographic base file, geographic units other than census tracts - zip-codes, traffic zones, telephone exchange areas - may be used.

CARPOL is available in FORTRAN IV and can be run on most computers.

CATS — The CATS program allows users to generate reports from the First Count Summary Tape. The system provides for data record selection, item selection, aggregation, data manipulation, and bar chart generation.

The CATS program i written in COBOL and can be run on any machine with a suitable COBOL compiler.

CENSLCT — This program searches a Census Summary Tape and extracts records for a geographic-coded area as specified by the user. Up to five copies can be made of a new tape and the user may specify a different block size for each copy.

The CENSLCT program is written in ANSI COBOL and will run on any computer having a suitable COBOL compiler.

COIN — The COIN program will add geographic (x/y) coordinates to a DIME file. These coordinates must be present if the DIME file is to be used for calculating areas or locating centroids of user-specified areas contained in a DIME file. In addition, coordinates are essential for network analysis and almost all types of computer mapping.

The COIN program is written in FORTRAN for an IBM System/360 computer.

CRAM — Basically, CRAM is a generalized system for determining the service area for a set of facility locations. In its most general form, service areas are constrained by facility capacities, and travel times.

The system uses a DIME file as its network base and in addition requiries demand and facility capacity inputs.

It is anticipated that the full CRAM system will consist of five (5) separate FORTRAN programs and two standard sort passes.

DACS — The DACS program is a flexible computer program for computing areas and centroids of user-specified blocks, census tracts or other areas defined in a DIME file. For example, centroid locations are required for input to GRIDS and some other computer mapping packages. Options have been programmed to generate SYMAP A-OUTLINES adjacency listings, and areas in acres, square feet or square miles. (Also see the description of the SCANGEN program and the C-MAP program.)

DACS is written in FORTRAN IV.

DCODE — The DCODE program adds special local area codes to the DIME file. These special local areas may be regions or districts not standard to the DIME file that are desired for geographic coding and mapping. The program reads a user-prepared dictionary and the input DIME file and writes a new copy of the file with codes included. In addition, its edit procedure checks for coding errors in the input dictionary.

The DCODE program is written in FORTRAN and will run on any computer with a suitable FORTRAN compiler.

DIME — The DIME - File Creation and Editing Package was designed for any city or rural area who wish to create their own DIME file.

The package is divided into two major sections. The first section deals primarily with the clerical procedures for the initial DIME file creation. The second section describes the computer programs used in converting and editing the manually encoded data to a master file for machine use.

The computer programs provide the capability for creating an automated DIME file, topological editing, address editing, and inserting corrections. In addition, digitizing instructions and a coordinate insertion computer program are included.

DIME con't. — The topological and address editing computer programs may also be used by SMSA's to edit standard census files.

The three basic programs which make up the DIME - File Creation and Editing Package were written in FORTRAN for an IBM System/360 computer.

DUPE-PAL — DUPEPAL, written in FORTRAN IV, contains two programs.

DUPENODES lists by census tract the node/maps which occur on a DIME file and flags all detected duplicate nodes.

PALS generates adjacency lists for an area code from a DIME file.

FIRST COUNT REPORTER — The FIRST COUNT REPORTER is a series of three (3) programs designed to produce information from the First Count Census Summary Tapes.

1) Reporter

First Count Reporter is a program designed to produce a standardized report from the First Count Summary Tapes. Input may be a Summary Tape or a CENSLOT-produced tape. Data may be organized by standard geographic codes or by specifically defined areas. If special area codes are used, the CODER program must rewrite the input tape.

2) Coder

Local area codes may be added to each record on a converted tape making it possible to organize data to the user's specifications for the final report. Local areas must be defined in terms of Census areas.

3) TCP (Tape Conversion Program)

The Tape Conversion Program converts the Census Bureau's standard First Count Tape from character to binary format.

The First Count Reporter programs are written in FORTRAN. They can be run on any computer having a suitable FORTRAN compiler.

FOUR-USE — The PDR staff has developed a system of programs for accessing the Fourth Count Summary tapes. These programs are designed to handle Files A, B and C, both for the population and housing counts. Each program is intended to perform a specific set of tasks that result in an output file that can stand alone or serve as input into another program. The aim of the program is to reduce data handling

problems encountered by the general user in the extraction of data from the Census Fourth Count of population and housing. The basic design of the program is user oriented. The production is a set of programs centered around the FORTRAN IV language with one program written in ANSI COBOL. By building a set of limited task programs, the need for a large capacity computer is reduced.

Included in the FOURUSE Package are the CENSLCT 4, VARSL VARSLCT 4, Driver 4, Reporter 4 and Coder 4, COMPACT 4, and CATS 4 programs.

GRIDS — GRIDS is a generalized computer graphics system for users of almost any computer system. It will produce various types of printer maps using a large variety of files quickly and easily. The objective is to enable the non-programmer to produce simple maps quickly and easily with little preparation. At the same time, the program can provide complex mapping capabilities for the more sophisticated users.

GRIDS is written in ANSI FORTRAN IV and will run on any computer system with a suitable FORTRAN compiler regardless of computer word size or operating system.

INDEXER — INDEXER is a program written in FORTRAN IV which creates and prints statistical indices using a simple series of commands, requiring no knowledge of programming by the user.

SCANGEN — SCANGEN is written in standard FORTRAN IV and is designed to allow a user to convert a SYMAP A-CONFORMOLINE deck (either clerically prepared or output from the DACS program) into a scanlines file as input to C-MAP.

SECS — SECS is a program designed to detect errors in the digitizing of DIME files. All possible pairs of segments within each census tract are examined to determine whether intersections occur. Intersections constitute errors since the DIME file segments are, by definition, connected only at their node point.

The SECS program is written in ANSI COBOL and will run on any computer with a standard compiler. This program has recently been enhanced to product a scaled map of nodes and coordinate locations as an aid to error detection.

TALLY — TALLY is a generalized computer program written in FORTRAN IV which creates geographic summary records from any administrative or other data file.

TIDE — TIDE stands for Terminal Interactive DIME Environment and consists of series of computer programs designed to allow those who wish to gain an understanding of DIME technology "hands-on" interactive computer instruction in how to create, edit and utilize DIME files for problem solving.

TIDE is written largely in FORTRAN IV with selected interactive extensions.

UNIMATCH — UNIMATCH is an advanced generalized record linkage system which will perform most geocoding applications including simulation of ADMATCH and an improved address matching algorithm.

UNIMATCH has many matching capabilities not available in ADMATCH such as the ability to handle building names, street intersections and non-address matching. Data files containing statistical information of any type can be linked, as long as a common information structure between the files can be identified and isolated.

UNIMATCH consists of three separate programs, a compiler, an assembler and an executor. The UNIMATCH system also provides utility functions such as field conversion, dictionary searches, record imputation, etc.

The system enables a user to define his own procedure for matching two files having one or more identifiers in common. In addition, the updating of a file and field conversion operations can be performed.

UNIMATCH is written in IBM 360 Assembly Language.

UPDIME — UPDIME is a system for correcting, extending and updating DIME files. The system also has the capability for editing. It consists of a computer program and is supported by a set of independent publications which are used in updating the Metropolitan Map Series. The system uses coding transaction cards which reflect those geographic changes that are to be used as data input to the UPDIME program. While requiring a larger machine capacity than the Census Bureau's CUE program, UPDIME performs topological and other edits on the results of the supplied corrections and provides faster turn around time and better error detection.

The UPDIME program can be used to keep DIME files current as changes in the urban geography occur continually. These changes include the addition of streets and boundary line changes.

UPDIME is written in FORTRAN IV.

ZIP-STAN — ZIPSTAN is an improved and redesigned system for standardizing addresses, designed to replace the current ADMATCH preprocessor for use with UNIMATCH. It generates a standardized match key from unformatted street address fields. It is also useful as a general file copying and reformatting utility program.

ZIPSTAN is written in IBM 360 Assembly Language.

The staff of PDR is currently working on an interactive system for use with DIME files. This system will include the ability to abstract, edit and update the files, as well as integrate local data in the DIME file structure.

SMALL AREA DATA USE PRODUCTS: DOCUMENTATION AND COMPUTER PROGRAM PACKAGES

Documents*

Census Use Study Report Series

This report series documents the activities and results of the New Haven Census Use Study.

Doc. No.

1. General Description. March 1970, 26pp. An overview of the development and operations of the New Haven Census Use Study.

2. Computer Mapping. August 1969, 44pp. A report on the mapping of census and local data using several computer mapping techniques.

3. Data Tabulation Activities. March 1970, 23pp. A report on the contents and uses of special tabulations provided to local agencies from the 1967 special census of New Haven, Connecticut.

4. The DIME Geocoding System. July 1970, 46pp. A report on the development of the DIME geographic base file, including a description of the file and the edit system, uses of the file and methods for creating a DIME file.

5. Data Interests of Local Agencies. April 1970, 92pp. A description of a series of surveys undertaken to explore the needs of local agencies for small-area data.

6. Family Health Survey. September 1969, 41pp. A report on a sample survey taken to augment data from the special census of New Haven with information on various elements of family health.

*The publications described are obtainable from the U.S. Census Bureau in Washington, D.C. 20233

7. Health Information System. October 1969, 67pp. This report documents the development of a maternal and child health care information system utilizing census and local data.

8. Data Uses in Health Planning. June 1970, 48pp. This report outlines the uses of data in health planning based on the general research conducted at the Census Use Study.

9. Data Uses in Urban Planning. February 1970, 28pp. A description of the general findings of the Census Use Study as they apply to the field of urban planning.

10. Data Uses in School Administration. April 1970, 30pp. A report describing the uses of data in school administration based upon activities conducted at the Census Use Study with local school administrators.

11. Area Travel Survey. September 1970, 43pp. A description of a sample survey conducted to augment the New Haven special census data with basic data for use in transportation planning.

12. Health Information System II. March 1971. 306pp. A supplementary report on the development and implementation of a computer-based health information system including materials, documentation, methodology and analysis.

14. Geocoding with ADMATCH - A Los Angeles Experience. January 1971, 23pp. A report describing the use of ADMATCH with a variety of local geographic base files such as ACG, DIME, and street tract indexes.

General Publications Series

These publications document the projects and activies of the Census Use Study in a number of fields of small-area data applications.

A Geographic Base File for Urban Data Systems. May 1969, 24pp. (Reprinted with permission of System Development Corporation). A description of the Dual Independent Map Encoding approach to the creation of a geographic base file.

Unified Statistical Evaluation Study - Report No. 2. (Preliminary). October 1972, 35pp. This report presents graphic illustrations of selected socioeconomic characteristics of Marion County, Indiana, taken from the Census Fourth Count Summary tape.

Handbook for Manpower Planners - Part I. September 1973. 102pp. A publication designed to assist manpower planners in the use of data from the 1970 Decennial Census.

Handbook for Manpower Planners - Part II. November 1974. This publication presents published and non-published sources of manpower data and suggests data applications to the field of human resource planning.

The First International DIME Colloquium: Conference Proceedings. May 1973, 115pp. This report presents the conference proceedings of the First International DIME Colloquium, held in Washington, D.C., August 27-29, 1972.

DIME Workshops: An interim Report. May 1973, 126 pp. A description of the DIME (Dual Independent Map Encoding) Workshops designed to assist local personnel in the use of the Census Bureau's GBF/DIME files and related computer techniques for small-area data analysis.

Data Uses in the Private Sector: Proceedings of the Executive Seminar. March 1973, 120pp. Included in this publication on the proceedings of the first executive seminar are reports dealing with the application of the DIME and small-area data technology to such areas as site selection, marketing, and trade area analysis

The Uses of GBF/DIME. September 1974, 170pp. A compendium of GBF/DIME applications to such fields as education, health, human resources, public safety and transportation, as documented by users throughout the country.

Data - Information - Decision: The Indianapolis Experience. October 1974. This illustrated publication documents the use of census data by decisionmakers in Indianapolis.

Technology Made Easy Series

This series of publications is designed to present complex, technical material and concepts in a form that can be readily understood by the local user.

1. ADMATCH Adventures. A non-technical, illustrated explanation of the ADMATCH computer program, which assigns geographic codes to data records containing address information.

2. CARPOL: An Overview. This simplified and illustrated version of the CARPOL computer program provides the non-technical administrator or carpooler with an introduction to the program and the technology behind it.

3. Day Care Center. An illustrated description of a system designed to provide information to assist administrators in allocating resources for social services.

4. DIME Comi and Stories. A cartoon explanation of the usefulness of DIME and the basis of Dual Independent Map Encoding.

Southern California Regional Information Study (SCRIS) Report Series (Selected)

This series deals with essential concepts and components of a modern information system. SCRIS, originally established as a joint effort of the Bureau of the Census and the County of Los Angeles, has attempted to transfer the experience gained in New Haven to a larger urban area.

1. Computer Graphics. A joint study of SCRIS and the Los Angeles City Planning Department, January 1970, 22pp. An example of the effective use of a geographic base file for referencing data, the application of ADMATCH for organizing data, and the use of the computer as a device for displaying information.

2. ACG/DIME Updating System: A First Look. March 1970, 26pp. A description of the initial thinking on a fundamental work item in the SCRIS program, the development of update and maintenance procedures for the geographic base file.

4. ACG/DIME Updating System: An Interim Report. October 1970, 55pp. An amplification of report No. 2 describing a proposed system and plans for updating the Los Angeles ACG/DIME file including mapping and clerical procedures as well as computer processing of update transactions.

8. ACG/DIME Updating System: The Long Beach Experience. October 1971, 68pp. The report contains extensive descriptions of procedures covering the mapping, clerical coding, recordkeeping, and computer processing actually used for a portion of the Long Beach file. Estimates of time, space, and cost requirements are also presented.

10. Southern California Regional Information Study Research Notes: 1970-1971. December 1971, 88pp. This report presents a compendium of Research Notes describing programs and projects of SCRIS. It contains both descriptions and full operating instructions for the use of a large number of programs including: (1) Coordinate Edit Program; (2) DIME-ADMATCH Record Conversion Program-DARC; (3) ACG-ADMATCH Record Conversion Program-AARC; (4) DIME Dictionary Coding Program-DCODE; (5) Coordinate Insertion Program-COIN: (6) Shortest Path Program; (7) DIME Area Centroid System-DACS; (8) Census Summary Tape Selector CENSLCT: (9) First County Reporter; (10) Columnar Aggregation and Tabulation System-CATS; 911) Report on Test of Hand-Digitizing; (12) CRIDS Catalogued Procedure.

Social and Health Indicators System Reports

The purpose of the Social and Health Indicators System, originally sponsored by the Office of Economic Opportunity (OEO) and now supported by the Department of Health, Education and Welfare, Health Resources Administration, is to provide a mechanism for monitoring the health status and social and economic well-being of populations receiving (or eligible to receive) health and social services in a given area.

The reports, in general, detail each of the three stages of the program:

Stage I: The initial phase of the program determines whether relevant and accurate data is readily available in adequate detail, timeliness, and appropriate (small-area) geographic level for the selected sites.

Stage II: Data is reconstructed on an annual basis for the 1965-1970 period and yields an historical data base to serve as the foundation for a social and health indicators system.

Stage III: The final report phase of the program involves refinement, supplement and extension of the 1965-1970 data base and the selection of key social and health indicators.

The following reports have been issued:

1. Atlanta: Part I, April 1974, 61pp.
2. Atlanta: Part II, October 1973, 212pp.
3. Los Angeles, June 1973, 320 pp.
4. Mound Bayou, Mississippi: Part I, June 1972, 65pp.
5. Mound Bayou, Mississippi: Part II, June 1972, 282pp.
6. Phoenix: Part I, October 1973, 59pp.

Computer Program Manuals

1. DOS ADMATCH: An Address Matching System. A computer program package designed for use in assigning geographic codes to local records using a DIME or similar geographic base file. It includes a users' manual and computer programs (written in IBM 360 Assembly language).

2. OS ADMATCH: An Address Matching System. A new version of the DOS programs specifically redesigned for more rapid processing of large files on medium-size and large computers. It includes a users' manual and computer programs (written in IBM 360 Assembly language).

Social and Health Indicators System Reports

The purpose of the Social and Health Indicators System, originally sponsored by the Office of Economic Opportunity (OEO) and now supported by the Department of Health, Education and Welfare, Health Resources Administration, is to provide a mechanism for monitoring the health status and social and economic well-being of populations receiving (or eligible to receive) health and social services in a given area.

The reports, in general, detail each of the three stages of the program:

Stage I: The initial phase of the program determines whether relevant and accurate data is readily available in adequate detail, timeliness, and appropriate (small-area) geographic level for the selected sites.

Stage II: Data is reconstructed on an annual basis for the 1965-1970 period and yields an historical data base to serve as the foundation for a social and health indicators system.

Stage III: The final report phase of the program involves refinement, supplement and extension of the 1965-1970 data base and the selection of key social and health indicators.

The following reports have been issued:

1. Atlanta: Part I, April 1974, 61pp.
2. Atlanta: Part II, October 1973, 212pp.
3. Los Angeles, June 1973, 320 pp.
4. Mound Bayou, Mississippi: Part I, June 1972, 65pp.
5. Mound Bayou, Mississippi: Part II, June 1972, 282pp.
6. Phoenix: Part I, October 1973, 59pp.

Computer Program Manuals

1. DOS ADMATCH: An Address Matching System. A computer program package designed for use in assigning geographic codes to local records using a DIME or similar geographic base file. It includes a users' manual and computer programs (written in IBM 360 Assembly language).

2. OS ADMATCH: An Address Matching System. A new version of the DOS programs specifically redesigned for more rapid processing of large files on medium-size and large computers. It includes a users' manual and computer programs (written in IBM 360 Assembly language).

3. GRIDS: A Computer Mapping System. A Computer program package for use on both small- and large-scale computers. Within a grid pattern structure it produces density, shading and value maps. It includes a users' manual and computer programs (written in Basic Fortran Iv).

4. DIME: A Geographic Base File System. A computer program package for creating a DIME Geographic base file. It includes clerical instructions, a computer manual and programs (written in Basic Fortran IV).

5. CARPOL. A computer program package for large-scale carpooling using DIME technology. It is designed to present applicants with lists of potential carpoolers who live nearby and work similar shifts. Its documentation is accompanied by administrative suggestions for executives and technical data for staff members not familiar with computer technology. It includes a users' manual and computer program (written in Basic Fortran IV).

6. UNIMATCH 1: A Record Linkage System. A computer program prackage for matching data files. It has the capability of handling building names, street intersections and non-address matching. It includes a users' manual and computer program (written in IBM Assembly language).

GLOSSARY

BLOCKS: Blocks are usually well-defined rectangular pieces of land, bounded by streets or roads. However, they may be irregular in shape or partially bounded by railroad tracks, streams, or other features. Blocks may not cross other boundaries such as city limits. Statistics for blocks are tabulated for all cities of 50,000 or more and for the urbanized areas of these cities. In addition, the Bureau also collected and tabulated data for blocks on a contractual basis in over 900 areas outside urbanized areas. A list of these contract block areas is available upon request from the Census Bureau.

BLOCK GROUPS: A Block group is a combination of contiguous city blocks and generally has a population of about 1,000. They are subdivisions of census tracts and are defined without regard to the boundaries of political or administrative areas, such as cities or minor civil divisions. When a block group straddles one or more of these boundaries, data for those parts in different areas are tabulated separately. For the purpose of providing small-area census data, block groups are the equivalent of enumeration districts in the built-up portions of the 145 largest SMSA's.

CENSUS COUNTY DIVISIONS (CCD's): These are statistical subdivisions of counties in 21 States where minor civil divisions were not suitable for presenting census data. In these States the MDC's are either too small, have lost nearly all meaning locally, or have frequent boundary changes. Over 7,000 CCD's have been established as relatively permanent statistical areas by the Bureau of the Census in cooperation with State and local groups. The States containing CCD's are: Alabama, Arizona, California, Colorado, Delaware, Florida, Georgia, Hawaii, Idaho, Kentucky, Montana, New Mexico, North Dakota, Oklahoma, Oregon, South Carolina, Tennessee, Texas, Utah, Washington and Wyoming.

CENSUS TRACTS: Tracts are statistical subdivisions of SMSA's and a limited number of areas outside of SMSA's. Tract boundaries are determined by a local committee and approved by the Bureau of the Census; they conform to county lines. Tracts are generally designed to be relatively homogeneous with respect to population characteristics, economic status and living conditions. The average tract has about 4,000 to 5,000 residents. For the 1970 census all SMSA's that were recognized at the time of the census were completely tracted. Over 100 counties, cities, or parts of counties outside SMSA's are also tracted.

CENTRAL BUSINESS DISTRICT (CBD): This Census Bureau term refers to an area of very high land valuation; characterized by a high concentration of retail businesses, offices, theaters, hotels, and service businesses and high traffic flow. CBD boundaries are defined in terms of existing census tract lines. CBD's are located in cities with 100,000 or more population.

COMPLETE COUNT OR 100-PERCENT DATA: This census terminology refers to items of information collected for every individual and housing unit at the time of the census.

COUNTIES: Counties are the primary political and administrative divisions of the States. The only major exceptions are in Louisiana where the divisions are called parishes, and in Alaska where 20 census divisions have been created as county equivalents. There are a number of cities which are independent of any county organization and, because they constitute primary divisions within their states, are accorded the same treatment as the county units in the preparation of census tabulations. NOTE: The District of Columbia and the independent cities within the States of Maryland, Missouri, Nevada and Virginia are all identified as county equivalents.

COUNTS: First-Sixth: The terms used by the Census Bureau to identify sets of 1970 Census data summaries on computer tape for certain kinds of geographic areas. Numbering refers to the order in which the data sets were released.

DUAL INDEPENDENT MAP ENCODING (DIME): The GBF/DIME file is a system for representing map features numerically for processing by computer. It is essentially a file of segment records where a typical segment is a portion of a street defined by intersecting streets or civil boundaries. Other segments may be defined for natural features, railroad tracks, jurisdiction boundaries and the like. The basic feature of a DIME file is that each node (intersection) and block (area bounded by segments) is uniquely identified. "Dual Independent" refers to the fact that each boundary segment in the network is described by specifying the nodes at the ends and the blocks to the right and left. With each node and block uniquely numbered, the computer can then construct two independent networks and match them to insure the existing network is completely represented and all land accounted for. Spatial information is added to the DIME file by assigning geographic coordinates to each intersection node.

ENUMERATION DISTRICTS (ED's): These are small areas defined by the Census Bureau, which have an average population of about 800. ED's are newly defined with each census for use as administrative units for the control of census operations. (In addition, some census tabulations are prepared for ED's). They never cross the boundary of a city, township, or other area (except city blocks) for which census data are tabulated.

GEOGRAPHIC BASE FILE (GBF): A description of the geographic attributes of an area (streets, intersections, geographic codes, addresses) in computer-readable form. An address coding guide and a DIME file are examples.

MINOR CIVIL DIVISIONS (MCD's): These are the primary political and administrative subdivisions of counties. The most common type of MCD is the township, but there are also towns, precincts, magisterial districts, gores, etc.

PLACES: There are two types of places recognized in the census tabulations--incorporated and unincorporated. Incorporated places are political units incorporated as cities, towns, villages, or boroughs, with the following exceptions: Towns in New England, New York, and Wisconsin and the boroughs in Alaska are not recognized as places; rather they are considered to be MCD's. Most incorporated places are subdivisions of the minor civil divisions (or census country divisions). Some incorporated places, however, constitute MCD's or cross MCD and county lines.

Unincorporated places are densely settled population centers without legally defined corporate limits. Each has a definite residential nucleus, and boundaries are drawn by the Bureau of the Census to include, insofar as possible, all the densely settled area.

REGIONS AND DIVISIONS: Census regions and divisions are large geographic areas which have been used for many decades for the purpose of providing summary figures at intermediate levels between those for the United States and those for individual States. The nine divisions are groupings of contiguous States, except for Alaska and Hawaii. The four regions are composed of groups of divisions.

STANDARD INDUSTRIAL CLASSIFICATION (SIC) CODE: Classified establishments by the type of industrial activity in which they are engaged. The categories are set up to provide a simple and workable arrangement of separating and classifying groups for statistical analysis. Comparability to census classifications is discussed in 1970 Census of Population, Classified Index of Industries and Occupations.

STANDARD METROPOLITAN STATISTICAL AREAS (SMSA's): As reported in the 1970 Census, Standard Metropolitan Statistical Areas consist of a county or group of contiguous counties (except in the New England States) containing at least one city of 50,000 inhabitants or more, or "twin cities" with a combined population of at least 50,000. In addition to the county(s), containing the central city(s), contiguous counties are included in an SMSA if they are metropolitan in character and are socially and economically integrated with the central city. In the New England States, SMSA's consist of groups of towns and cities instead of counties because these geographic areas are administratively more important than counties. The title of an SMSA identifies the central city or cities. The SMSA's are defined by the Office of Management and Budget. The 1970 census data were tabulated for a total of 247 SMSA's. NOTE: An additional 21 SMSA's were created in 1971 and 1972 as a result of post censal revision of criteria; the total number of SMSA's is presently 268.

SUMMARY TAPES: This term refers to the 1970 Census computer tapes which contain more data than are available in the published reports. However, no information specified to an individual person or housing unit is made available. The data about individuals and housing units are aggregated by areas; e.g., block groups or enumeration districts, census tracts, etc.

SUPPRESSION: The Census Bureau term which refers to the insertion of a code to replace an actual value to prevent disclosure of information concerning an individual or firm.

URBANIZED AREAS: An urbanized area consists of a central city (or "twin cities") of 50,000 inhabitants or more plus the surrounding densely settled incorporated and unincorporated areas which meet specific criteria of population size or density. Urbanized areas differ from SMSA's in that rural portions of the counties composing an SMSA are excluded, as are those places which are urban in nature but separated by rural territory from the densely populated area surrounding the central city. The urbanized area population is sometimes divided into those in the central city and those in the remainder of the area or the urban fringe.

APPENDIX

List of SMSA's With GBF/DIME Files

Alabama

Birmingham
Birmingham Regional Planning Commission

Florence
Muscle Shoals Council of Local Governments

Gadsden
City Planning Department

Huntsville
Huntsville City Planning Commission

Mobile
South Alabama Regional Planning Commission

Montgomery
Bureau of Urban Planning

Tuscaloosa
Planning Department

Alaska

Anchorage
Greater Anchorage Area Borough Planning Department

Arizona

Phoenix
Maricopa County Planning and Zoning Commission-Department

Tucson
Tucson Area Transportation Planning Agency

Arkansas

Fayetteville-Springdale

Fort Smith, Ark.-Okla.
Arkhoma Regional Planning Commission

Little Rock-North Little Rock
Planning Department

Pine Bluff
Southeast Arkansas Regional Planning Commission

California

Anaheim-Santa Ana-Garden Grove
Orange County Forecast and Analysis Center

Bakersfield
Kern County Planning Department

Fresno
Fresno County Planning Department

Los Angeles-Long Beach
Los Angeles County Regional Planning Commission

Modesto
Stanislaus County Planning Department

Oxnard-Simi Valley-Ventura
Ventura County Planning Department

Riverside County
Riverside County Planning Department

San Bernardino County
San Bernardino Planning Department

Sacramento
Sacramento Regional Area Planning Commission

Salinas-Seaside-Monterey
County of Monterey Planning Department

San Diego
Comprehensive Planning Organization

San Francisco-Oakland
Association of Bay Area Governments

San Jose
Santa Clara County Center
Santa Barbara County - Cities Area Planning Commission

Santa Cruz

Santa Rosa
Sonoma County Planning Department

Stockton
San Joaquin County Planning Department

Valejo-Fairfield-Napa
(See San Francisco)

Colorado

Colorado Springs
Pikes Peak Area Council of Governments

Denver-Boulder
Denver Regional Council of Governments

Pueblo
Pueblo Regional Planning Commission

Connecticut

Bridgeport
The Greater Bridgeport Regional Planning Agency

Bristol
(See New Britain)

Danbury
Housatonic Valley Council of Elected Officials

Hartford
Meriden
City Planning Commission
New Britain
Central Connecticut Regional Planning Agency

New Haven-West Haven
Regional Planning Agency of South Central Connecticut

New London-Norwich, Conn. - R.I.
Southeastern Connecticut Regional Planning Agency

Norwalk
Southwestern Regional Planning Agency

Stamford
Planning Board

Waterbury
Central Naugatuck Valley Regional Planning Agency

Delaware

Wilmington, Del.-N.J.-Md.
Wilmington Metropolitan Area Planning Coordinating Council

District of Columbia

Washington, D.C.-Md.-Va.
Metropolitan Washington Council of Governments

Florida

Daytona Beach
Volusia Council of Governments

Fort Lauderdale-Hollywood
Broward County Area Planning Board

Fort Myers

Gainesville
North Central Florida Regional Planning Council

Jacksonville
Jacksonville Area Planning Board

Lakeland-Winter Haven
Imperial Polk County Board of Commissioners

Melbourne-Titusville-Cocoa

Miami
Metropolitan Dade County Community Improvement Program

Orlando
East Central Florida Regional Planning Council

Pensacola
Escambia-Santa Rosa Regional Planning Council

Sarasota

Tallahassee
Tallahassee-Leon County Planning Department

Tampa-St. Petersburg
Tampa Bay Regional Planning Council

West Palm Beach-Boca Raton
Area Planning Board of Palm Beach County

Georgia

Albany
Albany Planning Commission

Atlanta
Atlanta Regional Commission

Augusta, Ga.-S.C.
Augusta-Richmond County Planning Commission

Columbus, Ga.-Ala.
Department of Community Development

Maco
Middle Georgia Area Planning Commission

Savannah
Chatham County-Savannah Metropolitan Planning Commission

Hawaii

Honolulu
State Department of Transportation

Idaho

Boise City
Ada Council of Governments

Illinois

Bloomington-Normal
McLean County Regional
Planning Commission

Champaign-Urbana-Rantoul
Champaign County Regional
Planning Commission

Chicago
Northeastern Illinois
Planning Commission

Decatur
Department of Community
Development

Peoria
Tri-County Regional Planning
Commission

Rockford
City-County Planning Commission
Rockford-Winnebago County

Springfield
Springfield-Sanagamon County
Regional Planning Commission

Indiana

Anderson
City Planning Department

Evansville, Ind.-Ky.
Southwestern Indiana & Kentucky
Regional Council of Governments

Fort Wayne
Three Rivers Coordinating Council

Gary-Hammond-East Chicago
(See Chicago)

Indianapolis
Census Use Study

Lafayette-West Lafayette
Tippecanoe Area Plan Commission

Muncie
Delaware-Muncie Metropolitan
Planning Commission

South Bend
Area Plan Commission of St. Joseph
County

Terre Haute
Area Planning Department for
Vigo County

Iowa

Cedar Rapids
Linn County Regional Planning
Agency

Davenport-Rock Island-Moline,
Iowa-Ill.
Bi-State Metropolitan Planning
Commission

Des Moines
Central Iowa Regional Association
of Local Governments

Dubuque
Dubuque County Metropolitan Area
Planning Commission

Sioux City, Iowa-Neb.
Siouxland Interstate Metropolitan
Planning Council

Waterloo-Cedar Falls
Iowa Northland Regional Council
of Governments

Kansas

Topeka
Topeka-Shawnee Metropolitan
Planning Commission

Wichita
Wichita-Sedgwick County
Metropolitan Area Planning
Department

Kentucky

Lexington
City-County Planning Commission

Louisville, Ky.-Ind.
Falls of the Ohio Metropolitan
Council of Governments

Owensboro, Ky.
Owensboro Metropolitan
Planning Commission

Louisiana

Alexandria
Rapides Area Planning
Commission

Baton Rouge
Capital Region Planning
Commission

Lafayette
Lafayette Planning Commission

Lake Charles
Calcasiue Regional Planning
Commission

Monroe
Ouachita Council of Governments

New Orleans
Regional Planning Commission

Shreveport
Caddo-Bossier Council of
Local Governments

Maine

Lewiston-Auburn
Androscoggin Valley Regional
Planning Commission

Portland
Greater Portland Council of
Governments

Maryland

Baltimore
Regional Planning Council

Massachusetts

Boston

Brockton

Fall River, Mass.-R.I.
Fall River Planning Board

Fitchburg-Leominster
Planning Board, City Hall

Lawrence-Haverhill Mass.-N.H.
Division of Planning & Development

Lowell, Mass.-N.H.
Lowell City Development Authority

New Bedford
City Planning Department

Pittsfield
Pittsfield City Planning Board

Springfield-Chicopee-Holyoke, Mass.-Conn.
Springfield Planning Department

Worcester
Worcester City Planning Department

Michigan

Ann Arbor
(See Detroit)

Battle Creek

Bay City
Bay Regional Planning Commission

Detroit
Southeast Michigan Council of Governments

Flint
Genesee County Metropolitan Planning Commission

Grand Rapids
Kent Ottawa Regional Planning Commission

Jackson
Jackson Metropolitan Area Regional Planning Commission

Kalamazoo-Portage
Kalamazoo Metropolitan County Planning Commission

Lansing-East Lansing
Tri-County Regional Planning Commission

Muskegon-Muskegon Heights
West Michigan Shoreline Regional Planning Commission

Saginaw
Office of the County Planner

Minnesota

Duluth-Superior, Minn.-Wis.
Head of the Lakes Council of Governments

Minneapolis-St. Paul, Minne.-Wis.
Metropolitan Council of the Twin Cities Area

Rochester
Planning Director, City of Rochester

St. Cloud
St. Cloud Metropolitan Area Planning Commission

Mississippi

Biloxi-Gulfport
Gulf Regional Planning Commission

Jackson
Jackson City Planning Board

Missouri

Columbia
County Administration Building

Kansas City, Mo.-Kan.
Mid-American Regional Council

St. Joseph
Metropolitan Planning Commission

St. Louis, Mo.-Ill.
East-West Gateway Coordinating Council

Springfield
Zoning and Planning Commission

Montana

Billings
Billings-Yellowstone, City-County Planning Board

Great Falls
Great Falls City-County Planning Board

Nebraska

Lincoln
Lincoln City-Lancaster County Planning Department

Omaha, Neb.-Iowa
Omaha-Council Bluffs Metropolitan Area Planning Agency

Nevada

Las Vegas
Clark County Regional Planning Council

Reno
The Regional Planning Commission of Reno, Sparks & Washoe County

New Hampshire

Manchester
Southern New Hampshire Planning Commission

Nashua
Nashua Regional Planning Commission

New Jersey

Atlantic City
Atlantic County Planning Department

Jersey City

Long Branch-Asbury Park
Monmouth County Planning Board

Newark

Morris County
Morris County Planning Board

New Brunswick-Perth Amboy-Sayreville
Middlesex County Planning Board

Paterson-Clifton-Passaic
(See New York, N.Y.)

Trenton
(See Philadelphia, Pa.)

Vineland-Millville-Bridgeton
Cumberland County Planning Board

New Mexico

Albuquerque
Middle Rio Grande Council of Governments of New Mexico

New York

Albany-Schenectady-Troy
Capital District Regional Planning Commission

Binghamton, NY-Pa.
Southern Tier East Regional Planning Board

Buffalo
Erie and Niagara Counties Regional Planning Board

Elmira
Chemung County Planning Board

Nassau-Suffolk
Nassau-Suffolk Regional Planning Board and Nassau-Suffolk Regional Marine Resources Council

New York-NY-N.J.
Tri-State Regional Planning Commission

Poughkeepsie
Dutchess County Planning Board

Rochester
Genesee Finger Lakes Regional Planning Board

Syracuse
Syracuse-Onondaga Planning Agency

Utica-Rome
Herkimer-Oneida Counties Comprehensive Planning Program

North Carolina

Asheville
Metropolitan Planning Board

Burlington

Charlotte-Gastonia
Charlotte-Mecklenburg Planning Commission

Fayetteville
Cumberland County Joint Planning Board

Greensboro-High Point
Guilford County Planning Department

Winston-Salem
City-County Planning Board

Raleigh-Durham
Triangle J Council of Governments

Wilmington-New Hanover
Planning Department

North Dakota

Fargo-Moorhead-Minn.
Moorhead Planning Commission

Ohio

Akron
(See Cleveland)

Canton
Stark County Area Transportation Study

Cincinnati, Ohio-Ky.-Ind.
Ohio-Kentucky-Indiana Regional Planning Authority

Cleveland
Northeast Ohio Areawide Coordinating Agency

Columbus
Mid-Ohio Regional Planning Commission

Dayton
Miami Valley Regional Planning Commission

Hamilton-Middletown
(See Cincinnati)

Lima
Lima-Allen County Regional Planning Commission

Lorain-Elyria
(See Cleveland)

Mansfield
Richland County Regional Planning Commission

Springfield
Clark County-Springfield Regional Planning Commission

Steubenville-Weirton, Ohio-W.Va.
Brooke-Hancock-Jefferson Metropolitan Planning Commission

Toledo, Ohio-Mich.
Toledo-Lucas County Planning Commissions

Warren
Trumbull County Planning Commission

Youngstown
County of Mahoning Planning Commission

Oklahoma

Lawton
Lawton Metropolitan Area Planning Commission

Oklahoma City
Association of Central Oklahoma Governments

Tulsa
Tulsa Metropolitan Area Planning Commission

Oregon

Eugene-Springfield
Lane Council of Governments

Portland, Oreg.-Wash.
Columbia Region Associaton of Governments

Salem
Mid-Willamette Valley

Pennsylvania

Allentown-Bethlehem-Easton, Pa-NY
Joint Planning Commission
Lehigh-Northampton Counties

Altoona
Blair County Planning Commission

Erie
Erie County Metropolitan Planning Commission

Harrisburg
Tri-County Regional Planning Commission

Johnstown
Cambria County Planning Commission

Lancaster
Lancaster County Planning Commission

Scranton
Lackawanna County Regional Planning Commission

Wilkes-Barre-Hazleton
Luzerne County Planning Commission

Philadelphia, Pa.-N.J.
Delaware Valley Regional Planning Commission

Pittsburgh
Southwestern Pennsylvania Regional Planning Commission

Reading
Berks County Planning Commission

Williamsport
Lycoming County Planning Commission

York
York County Planning Commission

Puerto Rico

Caguas
(See San Juan)

Mayaguez
(See San Juan)

Ponce
(See San Juan)

San Juan
Puerto Rico Planning Board

Rhode Island

Providence-Warwick-Pawtucket, R.I.-Mass.
Statewide Planning Program

South Carolina

Charleston
Berkeley-Charleston-Dorchester Regional Planning Council

Columbia
Central Midlands Regional Planning Council

Greenville-Spartanburg
Greenville County Health Dept.

South Dakota

Sioux Falls
City Planning Department

Tennessee

Chattanooga, Tenn.-Ga.
Chattanooga-Hamilton County Regional Planning Commission

Kingsport-Bristol, Tenn.-Va.

Knoxville
Knoxville/Knox County Metropolitan Planning Commission

Memphis, Tenn.-Ark.-Miss.
Memphis and Shelby County Planning Commission

Nashville-Davidson
Metropolitan Planning Commission

Texas

Abilene
West Central Texas Council of Governments

Amarillo
Panhandle Regional Planning Commission

Austin
Austin City Planning Department

Beaumont-Port Arthur-Orange
South East Texas Regional Planning Commission

Brownsville-Harlingen-San Benito
Planning and Zoning Commission

Bryan-College Station
Director of City Planning

Corpus Christi
Coastal Bend Council of Governments

Dallas-Fort Worth
North Central Texas Council of Governments

El Paso
West Texas Council of Governments

Galveston-Texas City
(See Houston)

Houston
Houston-Galveston Area Council

Killeen-Temple

Laredo
City Planning and Zoning Commission

Lubbock
Planning and Zoning Board

McAllen-Pharr-Edinburg
Planning and Zoning Commission

Midland
Planning Department

Odessa
Odessa Planning Department

San Angelo
Concho Valley Council of Governemnts

San Antonio
Alamo Area Council of Governments

Sherman-Denison
Texoma Regional Planning Commission

Texarkana, Texas-Texarkana, Ark.
Ark-Tex Council of Governments

Tyler
City Planning Department

Waco
Waco City Planning Department

Wichita Falls
Nortex Regional Planning Commission

El Paso
Department of Planning and Research

Utah

Provo-Orem
Utah County Council of Governments

Ogden
Weber Area Council of Governments

Salt Lake City
Salt Lake County

Virginia

Lynchburg
Central Virginia District Planning Commission

Newport News-Hampton
Peninsula Planning District Commission

Norfolk-Virginia Beach-Portsmouth Va.-N.C.
Southeastern Virginia Planning District Commission

Petersburg-Colonial Heights-Hopewell
Crater Planning District Commission

Richmond
Richmond Regional Planning Commission

Roanoke
Fifth Planning District Commission

Washington

Richland-Kennewick
Benton-Franklin Governmental Conference

Seattle-Everett
Puget Sound Governmental Conference

Spokane
Spokane Regional Planning Conference City Planning Department

Tacoma
(See Seattle)

Yakima
Yakima County Planning Commission

West Virginia

Charleston
Kanawha County Planning Commission

Huntington-Ashland, W.Va.-Ky.-Ohio
KYOVA Interstate Planning Commission

Parkersburg-Marietta, W.Va.-Ohio
Wood County Planning Commission

Wheeling, W.Va.-Ohio
Bel-O-Mar Interstate Planning Commission

Weirton
Brooke-Hancock-Jefferson Metropolitan Planning Commission

Wisconsin

Appleton-Oshkosh
East Central Wisconsin Regional Planning Commission

Green Bay
Green Bay-Brown County Planning Commission

Kenosha
(See Milwaukee)

La Crosse
Mississippi River Regional Planning Commission

Madison
Dane County Regional Planning Commission

Milwaukee
Southeastern Wisconsin Regional Planning Commission

Racine
(See Milwaukee)

CHAPTER 15

SMALL AREA POPULATION ESTIMATION

Leo Schuerman

This discussion on small area population estimation techniques was adapted from a chapter of Community Health Indicator System: A Model for Program Planning and Evaluation - Los Angeles Metropolitan Area (to be published). This report was written by Leo Schuerman of the University of Southern California while a staff member of the Census Use Study, Bureau of the Census.

INTRODUCTION

This Chapter was written for a person with a prior background in advanced statistical estimation techniques. For those who have not had such a background, the section will still be useful in providing an orientation to these techniques and tools. In most instances consulting assistance may be required to implement them.

The population estimation techniques discussed in this section were developed specifically to enable the estimation of denominators for rates. The denominator in a rate is the "population-at-risk." For example, if the event of interest is the occurrence of arrests of male juveniles ages 10-17, the numerator would be the reported number of male juvenile arrests and the denominator would be the total number of males, ages 10-17. As such, the denominator represents the upper limit of the phenomenon.

A number of sources of information were investigated for possible use as denominator values. These included Federal and local censuses and special surveys. None was discovered that could serve as a source for denominator values. The national census of population and housing occurs only at ten year intervals. Therefore, using decennial census information for denominators becomes increasingly less reliable during the decade between censuses. Special censuses conducted by either states or the Census Bureau occur only infrequently and at a considerable expense. Furthermore, special censuses rarely cover large areas of metropolitan regions. They are usually collected for a city when that city feels that its official population figure is lower than the true value and hence an inequity exists in the allocation of funds. Since the primary focus of special censuses is the updating of total population counts, such censuses usually do not contain the types of information needed for more specific subpopulations of interest. Therefore, a need exists to develop methods of arriving at yearly estimates for various population components.

ESTIMATION REQUIREMENTS

Given the need for denominators, and the inability of existing data sources to fulfill the requirements for this information, various types of estimation procedures were investigated. This process was guided by the general requisites that any estimation procedure should comply with the following parameters.

1. The estimation methodology should consist of a set of generalized procedural steps.

2. The procedure should utilize existing administrative data sets as input.

3. The procedure should not be "data bound," that is, tied to specific types or formats of administrative data.

4. The procedure should yield estimates that are rich in population composition detail. (e.g., age, sex, and ethnicity).

5. The procedure should have the flexibility to estimate areal units smaller than county or city totals.

6. Estimates should be obtainable on at least a yearly basis.

7. The benchmark year information should be gathered from locally derived files or from obtainable Federal Decennial Census files.

Among demographers and urban data specialists, there is a long tradition of interest in methods of population estimation. From this tradition comes the notion that to be truly useful, an estimation procedure should be chosen on the co-basis of the needs for the estimates and the existing data environment of the area for which estimates are required. In general, the needs are that estimates be annual, timely, of fine area resolution and rich in population composition detail. Further, estimates are specifically required at the level of census tracts rather than some larger areal units. This requires the use of administrative data which can be aggregated to this level.

Using the available administrative data files for Los Angeles County, seven administrative files were discovered that could efficiently be processed and employed as input to a system of population estimation. These files were already being collected and processed in machine readable form by various agencies in the county. All files were geocodable, that is, it was possible to locate each event in space in terms of a census tracting system. These files pertained to the following events:

1. Births
2. Deaths
3. Fetal Deaths
4. Reportable Disease
5. Mental Health Clinic Records
6. Adult Probation Cases
7. Juvenile Probation Cases

Given the specific requirements for estimates and the existing data environment in Los Angeles County, a method of poupulation estimation was decided upon that combined aspects of two standard estimating methodologies. These were the Ratio Correlation Method (also called the Multiple Regression method) and the Bogue-Duncan Composite method. Both methods employ data that is indirect in nature. Indirect data is information that is designed to measure something other than the population parameters for which estimates are needed. Indirect data serve as surrogates for the population parameters of interest. The distinction between direct and indirect data arises from the use of the files. The same data set may serve both functions for different applications. For example, the Juvenile Probation file may be considered direct information when used to investigate the characteristics of juvenile probationers and indirect when used to estimate the number of juveniles in the population ages 10 to 17. All data sources for the estimation procedure chosen here, may be considered indirect except for the use of the 1970 Federal Census of Population and Housing which was used for the bench mark dependent variables.

BOGUE-DUNCAN COMPOSITE METHOD [1]

The Bogue-Duncan Composite method[1] involves creating separate estimates for various subgroups of the population and then summing these components to arrive at totals. The method uses data sources specific to certain subgroups of the population as predicted by logical or demographically sound data relationships. Events such as the records of births are used as symptoms of population composition. Throughout this discussion, such events when they are recorded and tabulated are referred to as both symptoms (their logical status) and independent variables (their statistical utility). Specifically, the Bogue-Duncan method makes separate estimates of the population over 45 years of age from recorded deaths for both males and females over 45 years, and the sex and age specific death rates for persons over 45. The population ages 15 to 44 is estimated by inspection of birth records. The estimation procedure first estimates the number of females from the number of births and birth rates, and then estimates the number of males from the estimated number of females and the male-female ratios. Children aged five to fourteen are estimated from school enrollment data and school enrollment rates. Children less than five years of age are inferred from ratios of children under five and the female population. After these separate component estimates are completed, estimates for age, sex and total population may be constructed by simple addition. While it is not usually done, it would be possible to introduce a classification variable, such as ethnicity, if the information about births, deaths, and school enrollment is available by ethnic categories. The reason that this method is not recommended for small areas is that the rates of births, deaths and school enrollment vary widely over a population of census tracts in a culturally and socially diverse metropolitan area.

Thus, the utility of the Bogue-Duncan Composite method is the notion of creating separate estimates for various types of population subgroups, each based on the information source which should be most powerful for the estimation of that group. The procedure for summing the estimates created for subpopulations to totals remains a viable way of increasing the validity of total estimates.

RATIO CORRELATION METHOD

While the notion of subpopulation specific estimates was borrowed from the Bogue-Duncan scheme, the actual estimation method comes from another standard technique, known as the Ratio Correlation or Multiple Regression method. This set of procedures also utilizes symptomatic information as proxy measures for population parameters which are not themselves directly measurable. In its simplest form, a ratio correlation estimate involves two points in time and three known values with which to estimate one unknown value. An example might be drawn for Hypothetical City, involving telephones in service as symptoms for the estimation of the total population. Let us assume that in 1960, the total population is known from the Federal Census of Population to be 2000 persons. If in this year, telephone company records show 500 telephones in service in the city, there is a ratio of one telephone for each four persons. This ratio can alternately be expressed as .25 telephones per capita. Suppose that in 1965 telephone company records show that there are 600 telephones in service. In an attempt to estimate the number of persons at that time, an analyst might apply the ratio from the 1960 finding to the current number of telephones and conclude that the 1965 population of Hypothetical City could be estimated to be 2400 persons. The validity of such a conclusion depends on the assumption that the telephone per capita ratio remains constant or nearly constant during the five years.

A more common application of this method involves the charting of change variables over at least two points of time. This type of ratio correlation method involves two bench mark years for which both direct and indirect information can be evaluated. This application often employs the investigation of rate change between the symptom and population parameter over a period of time (perhaps between Decennial Censuses) and then uses the rate of change in the relationship to further refine the estimates. Other similar uses involve the use of proration of growth, where an estimate for a large spatial unit such as a state is allocated into estimates for that state's counties on the basis of the inferred change in the county's share of the total population.

In a still more practical application of the ratio correlation technique, an estimate does not depend on any single symptom type but rather a number of different symptoms combined in some fashion. Methods of combining multiple symptoms include: the use of a linear least squares equation (Multiple Regression), the use of symptom transformation and averaging, and the use of weights based on a priori assumptions. The most common of these alternatives is the use of Multiple Regression. Using many well chosen symptoms is believed to have the effect of partially canceling out the effects of an unforseen drastic change in the relationship between any single symptom and the dependent variable. An example would be a drastic reduction in the birth rate.

The primary shortcoming of this technique is that for many types of demographic phenomenon, especially those related to highly volatile social conditions, the temporal interval on which such change is measured (ten

years) is simply too long. Many phenomena change in ways that could not be accurately reflected in an analysis of rate of change during the period of time considered. To assume that the rate of change may be extended in a linear fashion is not sound for many investigations. A good example would be the change in the birth rate during the period of time between the 1960 and 1970 censuses. Obviously the use of a linear projection of the observed 1960 to 1970 change for the years subsequent to the 1970 census could be highly inappropriate.

RECOMMENDED ESTIMATION METHOD

The method of estimation recommended combines aspects of both the Bogue-Duncan Composite method and the statistical formulation of the Ratio Correlation or Regression method. Borrowed from the composite method is the use of separate estimates for subpopulations, summing for totals, and the concern for richness in population composition. From Ratio Correlation methods comes the regression formulation for combining various symptomatic files in an attempt to arrive at optimum predictive equations.

One additional ingredient was chosen for inclusion in the estimation model. It is the notion of controlling for type of urban area. Controlling has widespread utility in statistical investigations dealing with factors which impinge on some other relationship. In the case of small area population estimation, an attempt was made to control for differences in the relationships between symptoms and population parameters across a highly differentiated metropolitan region. This application for estimation has been suggested by Chu, Goldberg, Morrison,[2] and others.

NECESSARY ASSUMPTIONS

Shyrock and Siegal[3] have noted that the extent to which indirect data may be useful in models of population estimation is directly related to the effect that factors other than population composition have on the values of the symptomatic data.

An example which illustrated this point is the consideration of the relationship between the occurrence of a juvenile male on the juvenile probation file, and the number of males 10 to 17. The assumption which underlies the use of a regression framework for annual estimation is that over time, the relationship between the symptomatic variable and the dependent variable remains constant. This assumption is necessary in light of the fact that direct data is not available for periods of time between censuses. Obviously, the assumption becomes more tenuous with an increased interval between the estimated year and the benchmark on which the estimation equations have been constructed.

A second assumption is that the relationships between symptomatic and dependent variables hold over space. This assumption suggests that the proportion of males ages 10 to 17 which show up on a juvenile probation file in a developing suburb is the same as the proportion found in a central-city ghetto area. This assumption obviously ignores the heterogeneity of metropolitan areas. Anyone familiar with American cities would

expect that differentiation does exist and different relationships among symptoms and population parameters could be expected. This suggests the possibility of controlling for relevant characteristics of urban subareas. This was accomplished with a newly constructed variable, "The Urban Maturation Index".

DEVELOPMENT OF ESTIMATION EQUATIONS

The development of equations for the estimation of population parameters begins with assumed relationships. For example, it is assumed that the number of male juveniles which show up on the juvenile probation files would be positively related to the number of males aged 10 to 17. Similarly, we would expect the number of women ages 15 to 44 to be related to the number of reported births. Such intuitive notions about the data relationships guide the creation of estimation equations.

This paper discusses changes between the years 1968 and 1971, where it was necessary to create unique estimates of population for use as denominators, i.e. populations at risk, for these four years. Since census data is available for the year 1970 from the Decennial Census of Population and Housing, 1970 served as the benchmark for developing the estimation equations. Therefore, estimation for the years 1968 and 1969 involved reverse estimation. This was done for a number of reasons. First, the years in question were temporally closer to 1970 than to 1960, the date of the prior census. This helps minimize the changes over time on the factors which would alter the relationships between population parameters, the dependent variables, and the symptomatic variables. Secondly, in order to have a benchmark from which to build equations, it is necessary that the same data be available for the benchmark and the years to be estimated. In 1960, much of the locally collected data was not prepared in machine readable form. This limitation precluded using a forward and reverse technique for the estimation of the years between censuses or for the systematic alteration of weights based on past trends for the estimation of years subsequent to 1970.

Definition of Areas

Prior to the development of estimation equations for Los Angeles County, two spatial problems needed resolution. The first was the definition of areal subunits for which estimates are to be reported. The second problem was to define the universe of areal coverage for these estimates. For subareas, the logical choice of 1970 census tracts could not be employed since there had been a great many changes in tract boundaries between the 1960 and 1970 censuses. This problem arose as a result of the fact that during the years 1966 to 1970, the local area data files prepared by the various agencies still employed the 1960 census tracting system. In fact, until 1973, some of the agencies in Los Angeles County were still using the 1960 census tract grid for geocoding their records for individuals. This problem, coupled with the fact that the criterion variables for the benchmark year 1970 were to be selected from the 1970 census, required the use of a special set of comparable areas developed from both the 1960 and 1970 census tracting schemes.

The use of this set of comparable areas enabled the merging of data which contained either of the two geocoding schemes. The result was that data items for areas with constant boundaries could be matched. The comparability grid for all of Los Angeles County consisted of 1157 areas called "correspondence tracts". This compares to 1254 census tracts for 1960 and 1584 census tracts for the 1970 censuses.

It should be noted that for some areas of the metropolitan region, the correspondence tracts were the same as either the 1960 or the 1970 systems, since boundaries had not changed between the two censuses. For other areas, where changes in the tracting systems had occurred, it was necessary to aggregate on both of the systems until common boundaries were achieved.

The second spatial question involved the total area of coverage for this study, the universe of territory, for which estimations would be created. The determination of this area involved the elimination of correspondence tracts on the basis of three criteria taken from the 1970 census. These criteria were: total population, total number of housing units, and the percent of the total population which resided in group quarters. Correspondence tracts were eliminated if there were either less than 100 persons total population, fewer than 10 individual housing units, or greater than 50 percent of the population in group quarters. In addition to these criteria, the tracts which encompassed the channel islands and tracts which consisted of vessels in port were eliminated. The result of this elimination process provided 1142 correspondence tracts as the universe of territory.

Development of Data Files

Once the areal definition was completed, the next step was to develop the necessary data files for building the estimating equations at the benchmark year, 1970. Files used as symptoms included data which contained individual case records pertaining to births, deaths, reportable diseases, adult and juvenile probation records, and reported mental health cases.

The agencies responsible for the creation and maintenance of these files had included within each individual record, the census tract of residence for the individual to which the information pertained. This eliminated the step of adding the census tract codes either by hand or with the use of the Census Bureau's unimatch computer system.

Each of these files also contained information about the individual to which each record pertained, including age, sex, race and other relevant characteristics depending upon the agency's interest in the individual. It should be noted that the file on births contained information for both parents, as well as for the child. All other files pertained to only one individual.

One general problem with the information as collected by the various agencies should be noted. The Spanish ethnicity classification schemes used by the agencies were not all consistent with each other. For this reason, it was deemed advisable that each of the files should be processed for spanish surname codes. With one exception, this was done with the

Spanish surname syntactic analysis program. The exception was the mental health file, which contained only the first five letters in the individual's name due to considerations of confidentiality. For this file, the codes supplied by the Department of Mental Health were used. The identification of persons with Spanish surname was of concern for this project since, in Los Angeles County, individuals of Spanish heritage make up a significant proportion of the population.

At this point in the preparation of the data files, all names, addresses, and other information not relevant for the purposes of population estimation were eliminated from each of the six local area data files. The next step in file preparation was the aggregation of the individual case records to correspondence tract areas so that information comparable with the 1970 census could be achieved. The information for all individuals in each of the correspondence tracts was summed to yield area totals only. The files were then in the form of census tract summaries. For example, for tract 2032, there were 55 births in the year 1968. At this point, the characteristics of individuals were lost and only area statistics remained. This assured that all information obtained from the local agencies would remain strictly confidential.

After aggregation, separate master files for each of the years 1968 to 1971 were prepared. In addition to variables for each year from the symptomatic data files, dependent variables for the year 1970 were concatenated to the symptom file for that year to serve as the benchmark for the development of estimation equations. At this time, the variable, "Urban Maturation Index", was added to each of the yearly local area data files.

At this point, one file (1970) had been created upon which to develop estimation equations and three other files (1968, 1969, 1971) to which these equations would be applied to generate estimates for the various population parameters.

Checks For Spurious Relationships

The next procedure attempts to eliminate the possibility of using spurious variance for the estimation of population parameters. Two methods were employed. First, each of the variables was investigated with regard to minimum, maximum and mean values. Those symptomatic variables which had a mean number of persons per tract less than ten were eliminated so that estimates would not rely on symptoms which were rarely found within certain areas of the county. Secondly, in an attempt to control for for errors which arose from invalid or unreliable data preparation procedures, each variable was investigated within each tract over all four years.

The object of interest here was the occurrence of unusual changes in the values of variables between years. The method used compared the change between values of individual variables for a tract with the changes observed in the values of all other variables for that tract. For example, if the reported number of deaths to men in a tract changed radically between the years 1968 and 1969 but the number of reported deaths to women was relatively stable, as were most other symptomatic variables, the tract

was flagged for further investigation. If in fact, the reported number of deaths to men was the only variable to exhibit dramatic increase, it could be assumed that there was some irregularity in the coding or data collection procedures and that variable would be eliminated from the list of candidate variables for inclusion in the process of developing estimation equations. This was done only when the irregularities showed up over a number of tracts. The decision was based on whether or not the variable was suspect in greater than 20 percent of the tracts in the universe of territory. Few variables were actually eliminated for this second reason. Those which were eliminated came principally from the mental health file.

Selection Of Final Symptomatic Variables

From the variables which remained, the most probable symptomatic variables for each estimate were selected on the basis of a theoretical relationship between that variable and the dependent variable to be estimated. In the desire to estimate the number of women aged 15 to 44, it seemed that the number of reported births in a tract could be assumed to be related to this parameter. Similarily, the number of reported deaths to persons over the age of 64 was assumed to be negatively related to the number of children under the age of ten. Some of these a priori assumptions were rather loose since the aim here was to develop a pool of candidate variables which was manageable from the standpoint of the statistical procedures to be used. A multiple stepwise regression was used to select the most powerful symptomatic variables from this rather large pool of candidate variables. During this process, an attempt was made to select sets of variables composed of symptomatic variables from all of the six source files (births, deaths, disease, mental health, and juvenile and adult probation). This was done so that no single estimate would be heavily dependent on any single source data file. For example, in this way, it was anticipated that the estimates would be relatively independent of any single change in the relationship between a symptom and the population parameter. The aim here was to minimize the possibility that a real change in a proportion, such as a change in the birth rate, could alone significantly bias the estimates.

The symptom selection process was aided by the inspection of tables of simple product moment correlations (Pearson's R). It is important to note that variables which were felt to have important relationships on the basis of sociological assumptions were not eliminated on the basis of minimal correlations. This was done for the following reasons. First, it is possible that when a third variable is controlled, as in a multiple regression scheme, variables with little apparent first level correlations will add to the overall ability to predict the dependent variables. Secondly, the population of tracts are widely divergent in characteristics as the result of patterns of urban differentiation and specialization. For this reason, one variable could be very important for inclusion in the candidate list for certain types of areas, while its importance might be minimal for the whole county. This was found to be the case for a number of variables and will be discussed later in the description of controlling for levels of urban maturation.

The reduced set of candidate variables was then processed using the SPSS Statistical Package with the stepwise multiple regression program. On the basis of an iterative process, the procedure selects those variables which, in linear combination, do the best job of predicting the dependent variables. The choice of variables includes consideration of the optimum explained variance and the significance of the slopes. One stepwise multiple regression was run for each of 28 estimates using as dependent variables the known or criterion variables from the 1970 census and the symptomatic variables from the local agencies for the year 1970. From the variables selected by the statistical procedure, a subset of fewer than ten which were most powerful were noted. These variables were then used to form the estimation equations for the total county.

CONTROLLING FOR DIFFERENTIATION WITHIN METROPOLITAN AREAS

There are volumes of literature based on description and classification of the differentiation within urban metropolitan areas. In the thirties, Park and Burgess[4] and their students described the differentiation within the Chicago metropolitan area using a framework which became popular within the field of Human Ecology, known as the concentric zone theory. Others have shown that, while the concentric zone pattern of urban differentiation may have fit expanding Chicago quite well, the same pattern was not applicable to all cities. However, the search for a purely spatial urban typology did sensitize researchers to the fact that activities and cultural differentiation within various subareas of cities result in different patterns of behavior.

Coming from economic studies is the notion of stages of development in the use of land within metropolitan areas. Hoover and Vernon[5] postulated an urban land use cycle which proceeds from stages of settlement and building up, to intensive residential use, to use patterns other than residential -- principly commerical and industrial, and finally to stages of urban renewal. This perspective on urban development suggests that urban morphology should be viewed in terms of the land use stage of sections of a metropolitan area.

From this perspective, a number of studies have employed the notion of "date built up." (B. Duncan et al, Van Arsdol and Schuerman, 1971).[6] This concept has been used as a tool for the investigation of changes in neighborhood cohorts with respect to population aggregation and expansion, and the aggregation and expansion of ethnic populations. The use of "date built up" attempts to get at the processes of development of metropolitan subareas. For this estimation project, a variable similar to "date built up" was constructed.

For the purposes of population estimation, such a notion may control for differentiation within metropolitan regions. Given the assumption that the relationship between various symptomatic or indirect data and the population component values is to be controlled within a situation of urban differentiation, distinguishing subareas on the basis of status within an urban typology of stages of growth would seem appropriate. Toward this end, a new index of urban subarea position within a cyclical developmental process has been devised. The original variable, "date

built up" was based on the date at which a census tract achieved a single dwelling unit density of at least two units per acre. This measure does not however, reflect more advanced stages of urban development. There is the possibility that a subarea may exhibit persistence in terms of the type of land use. A compound index was constructed which takes into account the possibility that an area may or may not change from a "built up state" to one of intensive residential use. This measure is called the Urban Maturation Index.

Urban Maturation Index

Specifically, the Urban Maturation Index includes the notion of the use of land for both single and multiple dwelling units. As such, the Urban Maturation Index reflects both the first stage of settlement-subdivision and the transition to intensive residential land use. Figure 1 demonstrates that a cross-tabulation of these two measures of urban maturity was employed. The variable on the Y axis is the original variable "date built up", as used by Duncan. The X axis contains an analogous measure constructed from the date at which a census tract achieved a multiple dwelling unit density of at least two units per acre. When these two variables are cross-classified, a matrix of nine cells results. The cells were then ranked on the basis of the concept of "urbanness." These ranks are displayed as the uppermost number within each cell (See Figure 1).

It should be remembered that although there are nine discrete classifications to this newly developed scheme, they are descriptions of position within what is essentially a continuous process of urban development. In the same light, it must be stressed that although these positions in an urban typology are ranked for convenience during statistical procedures, these are essentially nominal classes rather than ordinal ones.

Area one is the classification which represents the most urban condition for this scheme. This value is ascribed to areas that met the criteria of two multiple dwelling units prior to 1940, but have not yet met the criteria with regard to single dwelling units. This is, at first, misleading since it seems to imply that the criteria for single dwelling units will be met at some date subsequent to the 1970 census. This is not the case, for if a census tract met the criteria for multiple dwelling units before 1940, but did not in 1940 meet the criteria for single dwelling units, one must assume that the intensity of residential use for that area existed in a stage which is beyond, in a developmental sense, that of an area which meets both criteria. An area that in 1940 met both criteria must be assumed to be in an earlier stage of development since the land is not utilized to its potential for high density residential activities. If it were, the prevalence of multiple dwelling units would cause the single dwelling unit criterion not to be met. Construction of an index in this way fits the Hoover and Vernon notion of stages of development in the land use cycle in a way that accounts for the cyclical nature of the pattern of land use change. A further extension of this notion would be one that incorporated historical measures on commerical and industrial land usage.

Date of Maturation
Multiple Dwelling Units

Date of Maturation Single Dwelling Units	Pre 1940	Pre 1960	Pre 1970	
Pre 1940	2 65 (5.7)	4 86 (7.5)	7 60 (5.2)	18.4
Pre 1960	3 41 (3.6)	5 139 (12.2)	8 410 (35.9)	51.7
Post 1970	1 58 (5.1)	6 39 (3.4)	9 244 (21.4)	29.9
	14.4	23.1	62.5	1142 100.0

Figure 1. URBAN INDEX MATRIX

Inspection of the ranks of the urbanization index in Figures 1 and 2 demonstrate that the measure is logically consistent with the notion of stages of development. Area nine is the least urban of the possible values. In this type of area, neither the criterion for multiple or single dwelling units is met.

The use of this newly constructed variable for development of a regression type estimation model proceeds from the need to adjust the prediction of the dependent variable. The basis of adjustment is the differences in the relationships among areas which are differentiated in terms of the land use cycle. The notion of land use is itself not critical for the estimation of population component values, since it is not land use, but rather population composition that is of interest. The notion of land use instead implies a range of demonstrated differences in the characteristics of the classified populations. There are important social conditions which correlate with the developmental and historical characteristics of neighborhoods. These correlates imply different levels of fertility, morbidity, incidence of mental health disorders, mortality, age structure, and crime. That the relationship between the proportion of occurrence of such events should vary systematically by urban maturation level is sufficient reason to control for differentiated areas when attempting to estimate population characteristics from the occurrence of these events.

Statistical Analysis Steps

First, a statistical test was performed to discover whether or not there were significant differences in the slopes of the regression hyperplanes among the nine urban maturation areas. If there was shown to be no significant differences among slopes, this new variable could be added to the equations as additional information for the correction of estimate values using a multiple eta procedure. If there was found to be significant differences among slopes, then a multiplicative relationship rather than an additive one could be assumed. In this second case, separate estimating equations would have to be developed for use in each of the nine urban maturation areas. The test used for this investigation was the multiple analysis of covariance.

The result of this test was that for all dependent variables, interaction was shown to occur across all nine urban maturation areas (i.e., there were significant differences among slopes). In one sense, this result is a testament to the utility of the urban maturation area scheme for the description of subareas within a metropolitan region. It demonstrated that the variable. "Urban Maturation Index", is indeed, useful for the statistical differentiation of areas on the basis of the relationships among population parameters.

1. Multiple dwelling unit density achieved before 1940, single dwelling unit density not yet achieved in 1970.
2. Multiple dwelling unit density achieved before 1940, single dwelling unit density achieved before 1940.
3. Multiple dwelling unit density achieved before 1940, single dwelling unit density achieved before 1960.
4. Multiple dwelling unit density achieved before 1960, single dwelling unit density achieved before 1940.
5. Multiple dwelling unit density achieved before 1960, single dwelling unit density achieved before 1960.
6. Multiple dwelling unit density achieved before 1960, single dwelling unit density not yet achieved in 1970.
7. Multiple dwelling unit density not yet achieved in 1970, single dwelling unit density achieved before 1940.
8. Multiple dwelling unit density not yet achieved in 1970, single dwelling unit density achieved before 1970.
9. Multiple dwelling unit density not yet achieved in 1970, single dwelling unit density not yet achieved in 1970.

Figure 2. URBAN MATURATION INDEX AREA DEFINITIONS

While the power of the Urban Maturation Index was suggested by the above finding, the unfortunate practical result of the multiple analysis of covariance test was that separate estimation equations had to be constructed for each of the nine areas. Therefore, the original pool of candidate variables was again searched for symptomatic variables specific to each of the urban maturation areas, since using symptoms based on relationships developed for the entire county had been demonstrated to be less powerful for certain of the Urban Maturation Index areas.

Again, each of the potential variables for use in the list of symptoms was investigated for number of cases within tracts contained in each of the Urban Maturation Index areas. Variables which had too few cases were again eliminated from the lists of potential candidates. Following this selection procedure, the nine sets of candidate variables were then processed with multiple stepwise regression in an attempt to select sets of 5 to 10 variables to be used in the actual estimation equations for the tracts contained within each of the nine areas. There is another way of demonstrating that creating separate estimating equations for each of the areas significantly increased the ability to predict or estimate accurately. Reference to Figure 3 will demonstrate the differences between the R squared values for the estimating equations for the entire county and the values of the R squares for the equations developed for each of the nine Urban Maturation Index areas.

When all 252 equations were finalized, they were stored for input to the ESTIMATOR program. ESTIMATOR was run for each of the study years for which estimates were needed. The files of yearly estimates were then sorted and included in the master data file for the construction of the desired rates as described in the Introduction to this Chapter.

NOTES

[1]Donald J. Bogue and Beverly Duncan, "A Composite Method for Estimating Postcensal Population of Small Areas by Age, Sex, and Color," Vital Statistics--Special Reports, Selected Studies [an HEW, National Center for Health Statistics serial publication], Vol. 47 (August), 1959.

[2]Shih-Fan Chu, "On the Multiple Regression Method of Population Estimation," Nevada Population Studies: Technical Report (University of Nevada, n.d.); David Goldberg, et al., "Estimates of Population Change in Michigan, 1950-1960," Michigan Population Studies, No. 1 (Ann Arbor: University of Michigan, Program for Research in Population and Human Ecology, 1960); David Goldberg, et al., "A Test of the Accuracy of Ratio Correlation Population Estimates," Land Economics, Vol. 40, 1964, pp. 100-102; Peter A. Morrison, Demographic Information for Cities: A Manual for Estimating and Projecting Local Population Characteristics, R-1373-NSF (Santa Monica: RAND Corporation, 1971).

[3]Henry S. Shryock, Jr. and Jacob S. Seigal and Associates, The Methods and Materials of Demography (Washington, D.C.: U.S. Bureau of the Census, 1973).

DEPENDENT VARIABLE		AREAS									
NUMBER	LABELS	1	2	3	4	5	6	7	8	9	TOT	AVE
3	TOTAL MALES	.819	.822	.945	.886	.940	.885	.821	.925	.949	.879	.912
5	POP < 10	.846	.840	.923	.906	.915	.895	.897	.939	.962	.900	.924
7	MALES 10-17	.748	.475	.820	.731	.802	.804	.734	.809	.917	.634	.800
8	FEM 10-17	.666	.513	.768	.721	.802	.870	.792	.814	.913	.594	.801
12	MALES > 64	.862	.880	.924	.929	.969	.942	.837	.960	.915	.897	.931
13	FEM > 64	.873	.923	.932	.932	.975	.939	.836	.950	.961	.920	.942
15	MALES > 17	.818	.854	.932	.899	.854	.883	.818	.920	.937	.852	.902
16	FEM > 17	.741	.844	.883	.868	.958	.841	.804	.939	.962	.895	.913
17	FEM 15-44	.751	.783	.892	.849	.927	.898	.828	.908	.946	.846	.894
231	POP 18-64	.765	.796	.890	.869	.944	.826	.772	.919	.942	.885	.896
22	BLK MALES	.908	.976	.981	.958	.970	.998	.984	.949	.957	.938	.958
23	SPAN MALES	.916	.953	.981	.952	.940	.971	.988	.959	.966	.935	.958
25	BLK FEMALES	.942	.968	.974	.957	.978	.998	.970	.954	.965	.948	.963
26	SPAN FEMALES	.917	.953	.969	.948	.932	.959	.982	.966	.968	.928	.958
28	BLK MALES 10-17	.907	.966	.900	.942	.969	.990	.960	.927	.962	.875	.945
29	SPAN MALES 10-17	.856	.942	.977	.919	.809	.954	.962	.914	.941	.803	.912
31	BLKS < 18	.956	.980	.975	.971	.982	.991	.978	.961	.973	.941	.970
32	SPAN < 18	.905	.962	.976	.958	.916	.969	.974	.954	.962	.907	.952
34	BLK FEM 15-44	.922	.965	.976	.959	.976	.999	.962	.937	.959	.940	.954
35	SPAN FEM 15-44	.906	.954	.952	.951	.933	.930	.989	.952	.958	.919	.950

Figure 3. R SQUARED COMPARISONS

[4]Robert E. Park, Ernest W. Burgess, and Robert D. McKenzie, eds., The City (Chicago: University of Chicago Press, 1967).

[5]Edgar M. Hoover and Raymond Vernon, Anatomy of a Metropolis (Garden City, N.Y.: Doubleday, 1962).

[6]Beverly Duncan, Georges Sabagh, and Maurice D. Van Arsdol, Jr., "Patterns of City Growth," American Journal of Sociology, Vol. 67 (January), 1962, pp. 419-421; Maurice D. Van Arsdol and Leo Schuerman, "Redistribution and Assimilation of Ethnic Populations: The Los Angeles Case," Demography, Vol. 8, 1971, pp. 459-480.

CHAPTER 16

TOOLS FOR FORECASTING

Frank Lockfeld

INTRODUCTION

When the locations or coordinates of criminal events are located on a map along with geographic identifiers the resulting display can be a useful tool in detecting and diagnosing crime patterns. The statistical data resulting from the map encoding of events can, with proper caution, be used in forecasting events.

COMPUTER MAPPING

The ability to identify the locations of criminal events can significantly enhance crime analysis. Knowing where different crimes occur assists in the diagnosis of problems, facilitates the development of more accurate baseline measures, provides a basis for focused programs, enables program evaluation, and is valuable for projecting levels of activities.

Products developed by the Bureau of the Census have improved the availability and cost-effectiveness of location identification through computer techniques. A computerized map, called a Geographic Base File (GBF/DIME) was developed for most metropolitan areas as part of the 1970 decennial Census. These maps have been corrected, updated and extended in many Standard Statistical Metropolitan Areas (SMSA's) throughout the country. The GBF/DIME identifies the address ranges along each street segment together with geographic identifiers relating to each side of the segment, such as census tract and block number, corporate jurisdiction, zip code, etc. Additional geographic identification, such as beat or precinct number, may be added to the GBF/DIME. The GBF/DIME identifies x-y coordinates at each end (node) of the segment. These coordinates relate to latitude-longitude and to state plane coordinate systems for states where such systems exist.

"Geocoding" (Criminal Justice System Standard 4.8) is the process of matching the specified location of an event to the GBF/DIME and attaching the appropriate geographic identifiers to the other information recorded about that event. The Census program "UNIMATCH" enables events referenced to a street address, street intersection or place name to be matched against the GBF/DIME. Another Census computer program, "GRIDS," enables computer line printer maps to be produced from the x-y coordinate values given these events (see Figures 1 and 2). Other computer programs, such as SYMAP, developed by the Harvard University Laboratory for Computer Graphics and Spatial Analysis, are also available for the production of maps from geocoded data.

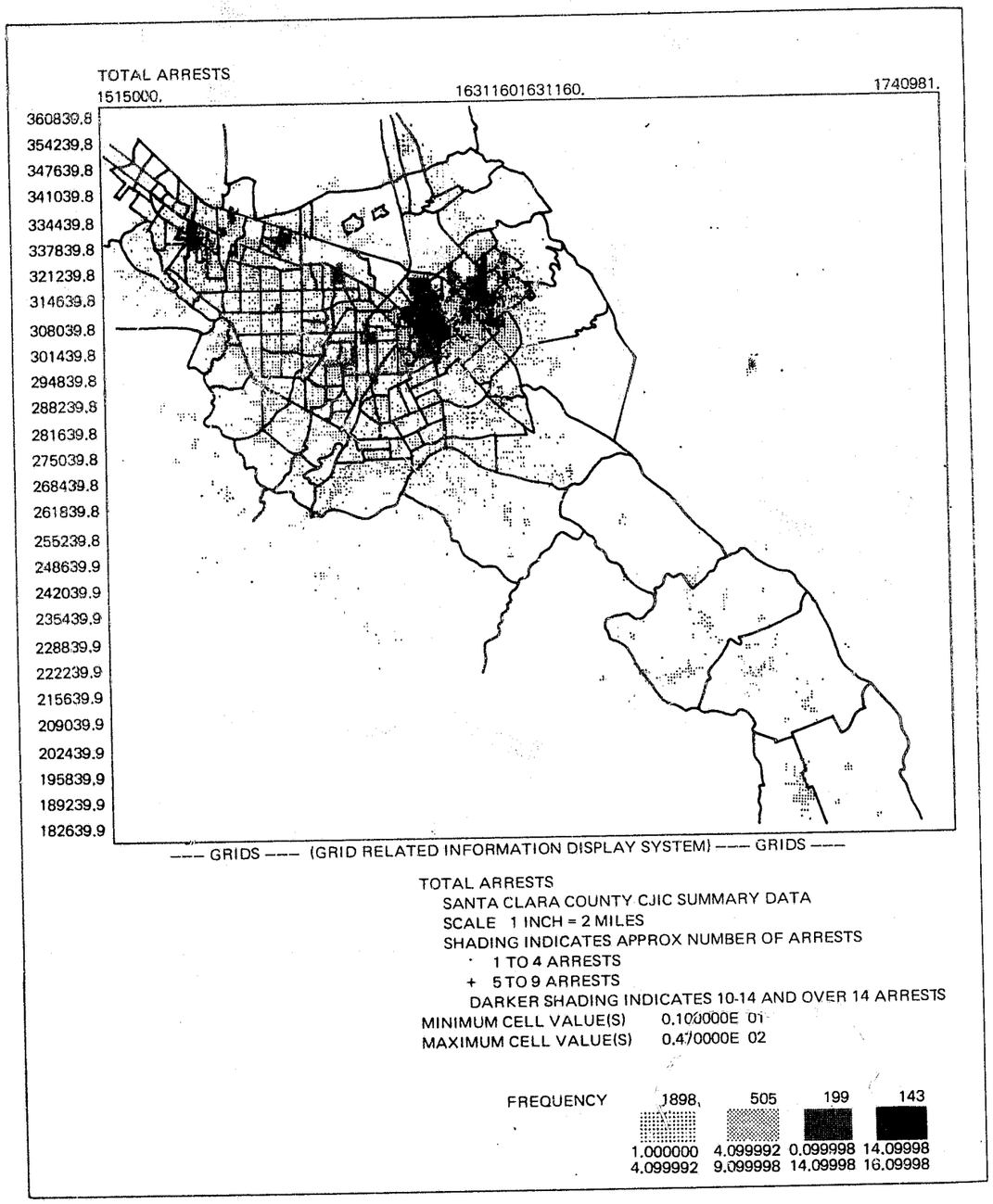


Figure 1. EXAMPLE OF "GRIDS" SHADED MAP.

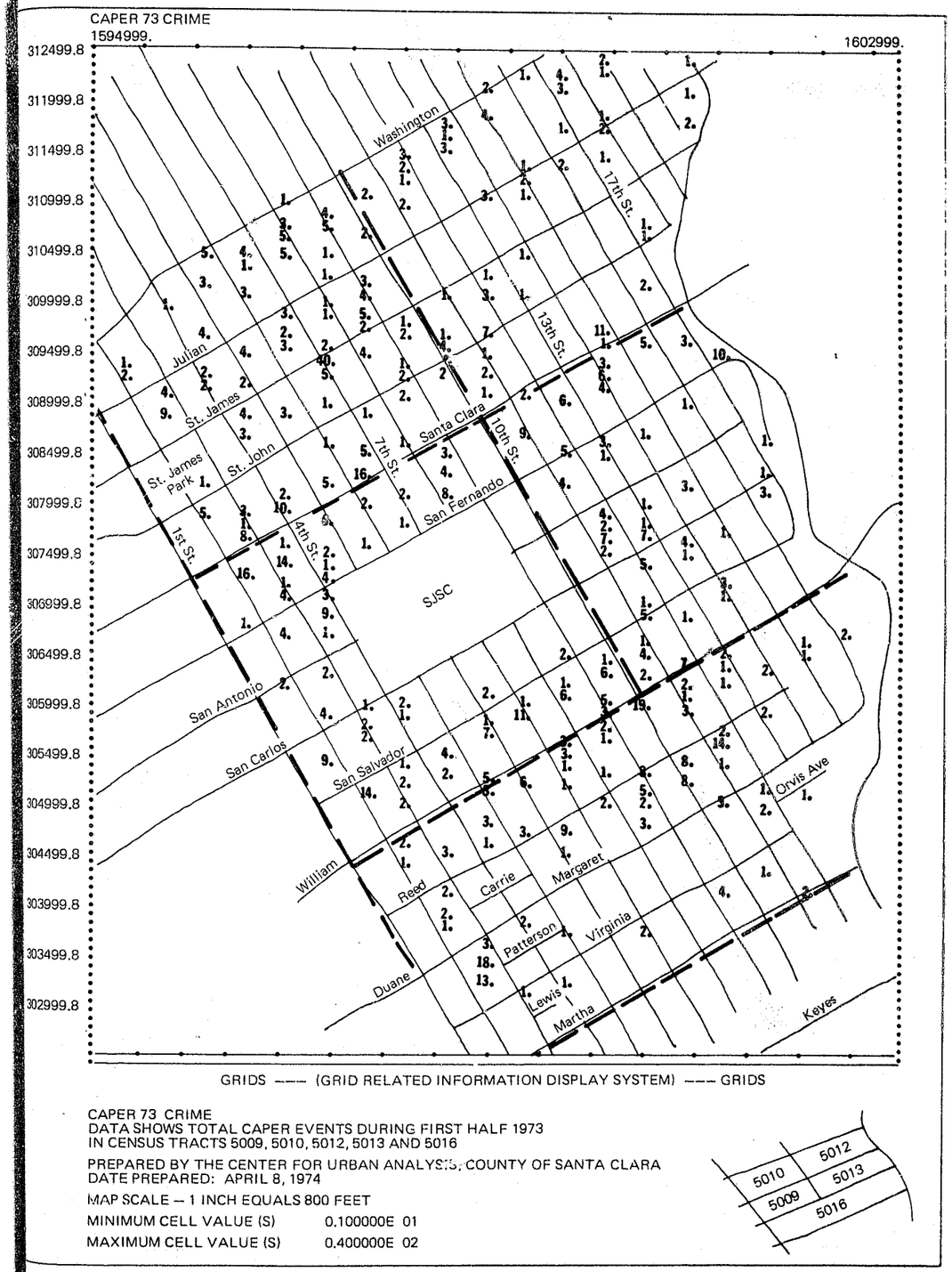


Figure 2. EXAMPLE OF "GRIDS" VALUE MAP.

While map displays can facilitate crime analysis, the power of geocoded crime data is most effectively realized when organized as a data base in an interactive computer system capable of selectively extracting, summarizing, manipulating and displaying the data directly on a map image.[1] (See Figures 3A, 3B and 4A, 4B.)

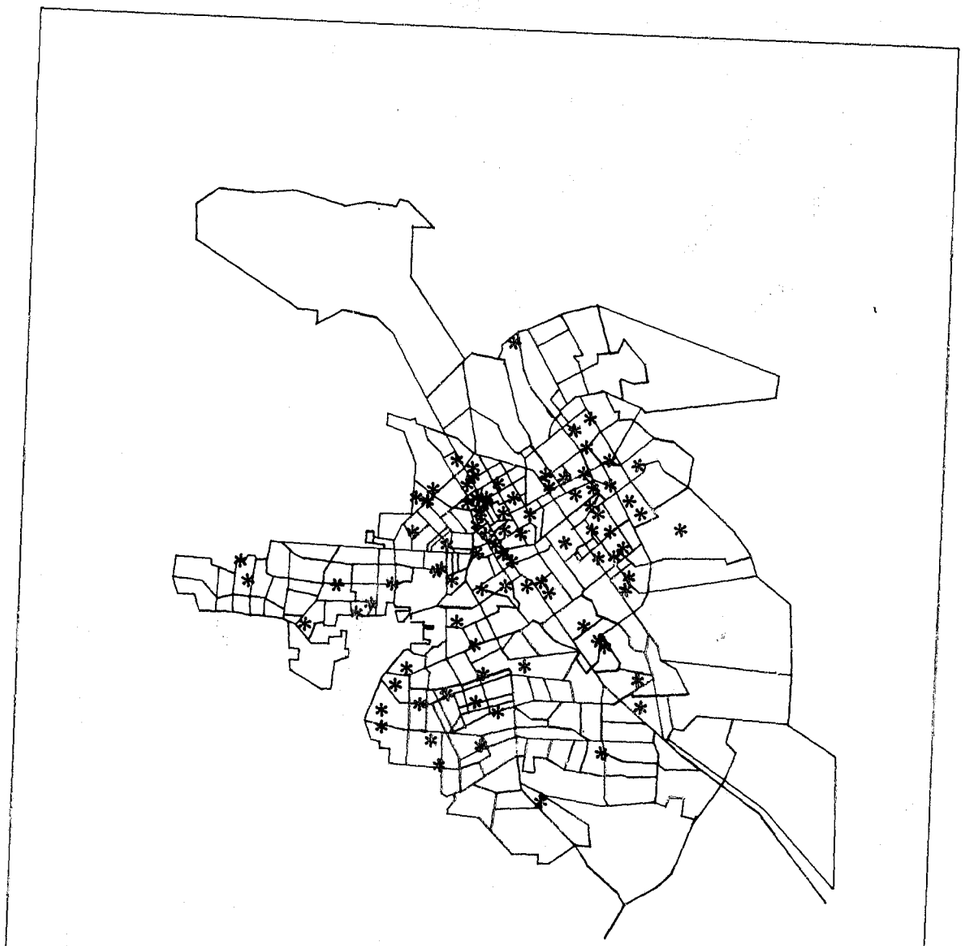
One of the main advantages of geocoded crime data is that the events may be aggregated to relate to geographic units for which other data is available. Census data and property data are important complementary data items for crime analysis. A property data source, for example, that distinguishes the number of commercial establishments within particular areas enhances the analysis of crimes against such activities. Census data provides detailed characteristics of population and housing for small areas (blocks, block groups and census tracts) that enables the association of such characteristics with crime data. The use of geocoded calls-for-service data against property and census data facilitates beat or patrol design that equalize workloads,[2] permits organization for strategic response to anticipated demands for police services,[3] and provides documentation for estimating workforce requirements associated with changing characteristics of the serviced area.

FORECASTING WITH GEOCODED DATA

Most crime-specific analysis techniques are based on cross-sectional data, that is, data that is referenced to a single point in time or independent of time. A major advantage of location-coded data is that a greater variety of analytic techniques may be employed, especially statistical methods of spatial analysis where each area serves as an independent observation. A classic study analyzing the spatial distribution of crimes within a city and their relationships to significant economic, demographic and social determinants is C. F. Schmid "Urban Crime Areas."[4] This study used offense and arrestee data during 1949-1951 together with 1950 Census tract data to generalize "gradients" and "isopleths" of crime and to relate crime to urban typologies drawn from sociology. The ecological pattern of crimes were also analyzed with factor analysis of the offense, arrestee and census data. In a similar study, Howard[5] used cluster analysis scores for census tracts, based on 1970 Census data, as well as linear multiple regression, and 1972 arrestee data to estimate relationships between underlying social dimensions and arrestee rates.

Techniques such as factor analysis and regression analysis may yield statistically meaningful associations of characteristics, but do not necessarily demonstrate causality. Equations derived from regression that statistically "explain" the variation among the observations for a given place at a given time, or replicate observed values within certain confidence limits, are not necessarily transferrable to other places or other times. This limitation is of particular concern in the contest of crime analysis and the forecasting or projection of crime.

The genesis of crime is recognized to be in the community. but no theoretical construct--sociological, psychological or economic--has been accepted as the explanation of the etiology of crime.[6,7] While regression is a powerful tool of analysis, its application to data other than that in



"EXTRACTING" the data base for in-progress calls-for-service, day 25 of the sample (San Jose Police Department data), and displaying the locations on a map composed of 272 reporting and analysis zones.

Figure 3A. INTERACTIVE COMPUTER GRAPHIC SYSTEM (GADS - Geo-Data Analysis and Display System) RELATING A DATA BASE TO A MAP.

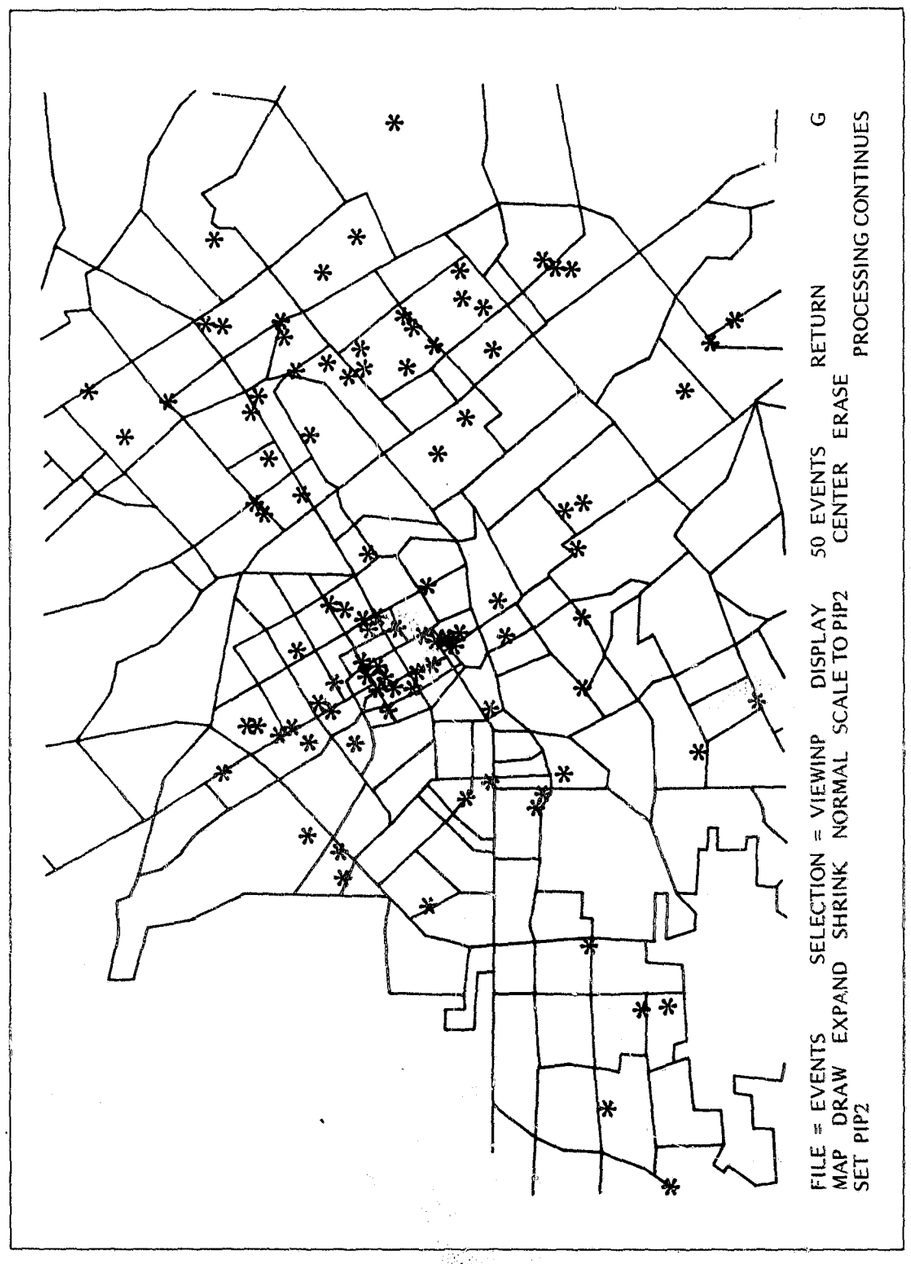


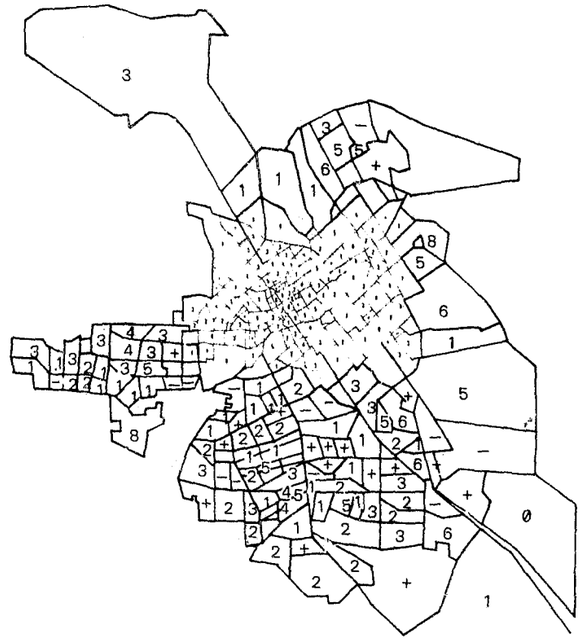
Figure 3B. GADS – ENLARGING THE MAP IMAGE OF THE "EXTRACTION" DISPLAY OF FIGURE 3A TO FOCUS ON EVENTS IN THE CENTRAL AREA.

Value ranges for symbols are shown at bottom of illustration. The item represented, TOTCTIM is the total consumed officer time responding to calls-for-service in the 1974 sample. For example, a "9" is displayed when the total consumed time for a zone exceeds 9000 minutes, an "8" is displayed when the consumed time totals between 8000 and 8999 minutes, etc.

CURRENT SUPERZONES

GADS BASIC ZONE MAP OF SAN JOSE WITH DISPLAY SYMBOLS

ZONES
0



SAVE EXTEND CLEAR & REDRAW
PRINT NORMAL SCALE OVERLAY WITH MAP 1 FIND ZONE
GET CREATE ENLARGE ZONE SYMB 1 = ON DOTS
RETURN SHRINK SSMODE NUMBER

CURRENT STATEMENTS

TOTAL CONSUMED TIME ALL DAYS 1974

DISPLAY SYMBOLS FOR EACH ZONE DEPENDING ON TOTAL CONSUMED TIME

```
9 TOTCTIM(91)  > 9000,  8:  TOTCTIM(91) >  8000,  7:  TOTCTIM(91) >  7000,
6 TOTCTIM(91)  > 6000,  5:  TOTCTIM(91) >  5000,  4:  TOTCTIM(91) >  4000,
3 TOTCTIM(91)  > 3000,  2:  TOTCTIM(91) >  2000,  1:  TOTCTIM(91) >  1000,
+ TOTCTIM(91)  > 500,   -:  TOTCTIM(91) >  0,     0:  TOTCTIM(91) =  0,
  SAVESYMB(1),
```

Figure 4A. GADS–DISPLAYING SYMBOLS FOR EACH ZONE ON THE MAP IMAGE.

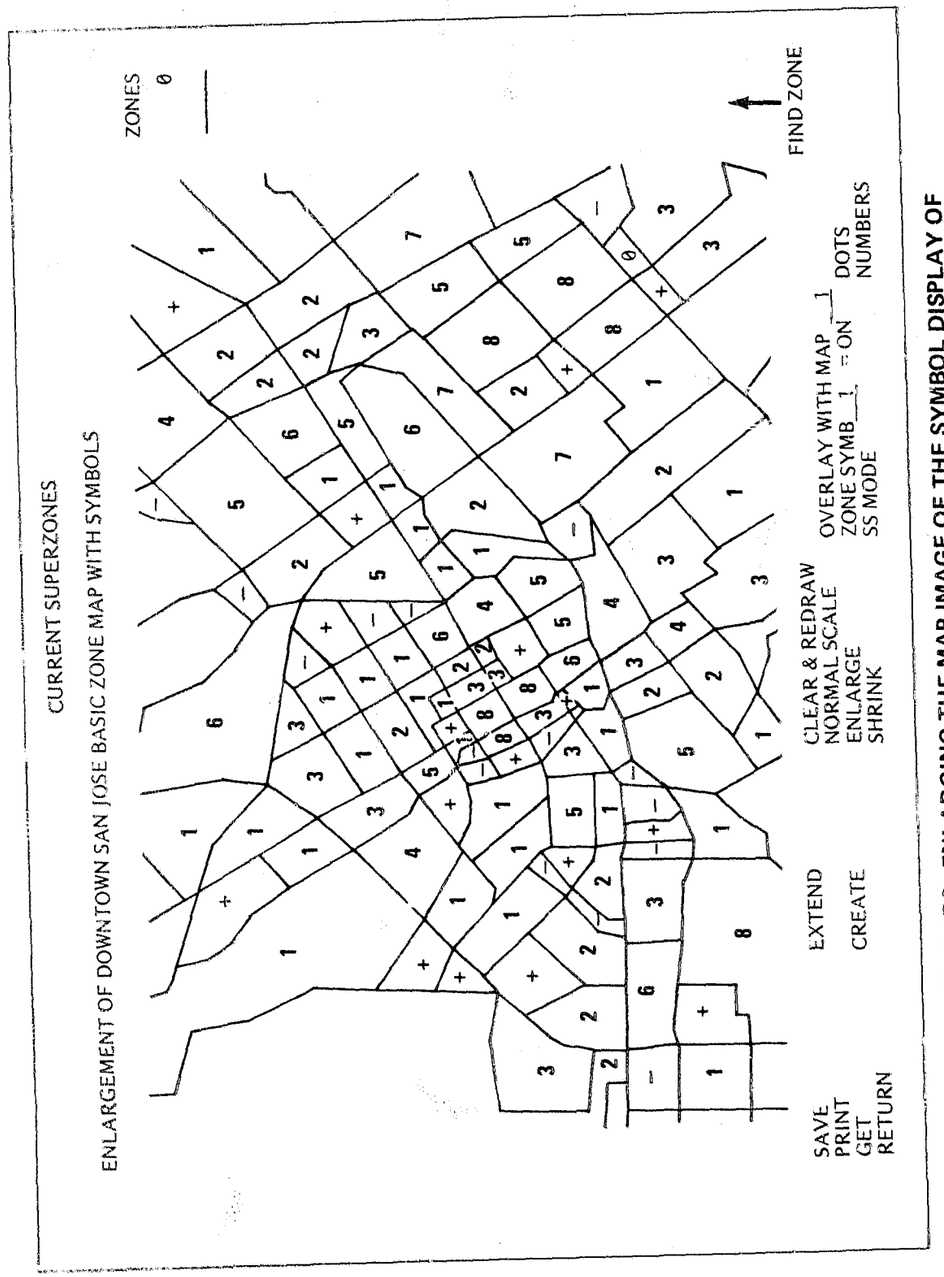


Figure 4B. GADS—ENLARGING THE MAP IMAGE OF THE SYMBOL DISPLAY OF FIGURE 4A TO FOCUS ON ZONE TOTALS IN THE CENTRAL AREA.

which the equations were estimated suffers from the assumption of ceteris paribus--all things remain in the same relationship. This assumption is particularly difficult for social phenomena such as crime. Technology, social and cultural mores, and other attributes difficult to measure are recognized to be associated with crime, but are rarely explicitly included in such analysis. Even if such characteristics were included in regression however, values of these characteristics at future times would be required to use the estimating equations in a forecasting or projecting mode. Still, these statistical methods should be considered. If there is any doubt about their application in a given instance, the advice of a statistician should be sought.

Two techniques are useful for projecting crime over a short-term period (e.g., up to 2-3 years), though neither resolves the fundamental problems above. The first technique is time series analysis. Exponential smoothing is a particular method of time series analysis that gives greatest weight to recent information and least weight to older data.[8] Spectral analysis is a technique useful for events of a cyclical pattern, such as crimes with weekly or seasonal variations. A second approach is the analysis of first differences, a method used in modeling other urban behavior. Multiple regression is a primary tool for such analysis, but the variables are represented as changes over a unit of time rather than cross-sectional. The projection period established by such analysis is the unit of time originally represented. Like other forms of projections through equations, estimates of future values of the independent variables are required. This approach can use either independent estimates of such changes or functions of previously observed states. An expression that uses both variations might be as follows:

$$dASSAULTS_i = a + b_1\ MALES,\ 17\text{-}25_i + b_2\ dUNEMPLOYMENT,MALES + b_3 dCOMMCL_i + E_i$$

where:

dASSAULTS = change in the number of assaults in area i from time t to t+1

$MALES,\ 17\text{-}25_i$ = number of males aged 17-25 resident in area i at t+1

dUNEMPLOYMENT, MALES = change in the areawide rate of male unemployment from t to t+1

$dCOMMCL_i$ = change in the number of commercial establishments in area i from t to t+1

a, b_1, b_2, b_3 = constants estimated by regression

E_i = error term for area i

The usual cautions and criteria for validity of multiple regression equations apply to first difference estimates.

NOTES

[1]P. E. Mantey, and E. D. Carlson, "Integrated Data Bases for Municipal Decision-Making," National Computer Conference, 1975 (Montvale, N.J.: AFIPS Press, 1975), pp. 487-493.

[2]R. V. Bradshaw, "Man-Machine Interaction for Police Field Manpower Deployment," Urban and Regional Information Systems Association Conference Proceedings, 1974.

[3]R. C. Larson, "A Hypercube Queuing Model for Facility Location and Redistricting in Urban Emergency Services," The Rand Corporation, R-1238-HUD, March 1973.

[4]C. F. Schmid, "Urban Crime Areas," American Sociological Review. XXV, no. 4 and no. 5 (August and October 1960).

[5]M. K. Howard, "Predicting Arrestee Residence Rates," Santa Clara County Criminal Justice Pilot Program, April 1973.

[6]M. D. Maltz, "Measures of Effectiveness for Crime Reduction Programs," Operations Research. XXIII, no. 3 (May-June 1975), pp. 452-474.

[7]J. B. Kau and P. H. Rubin, "New Estimates of the Determinants of Urban Crime," The Annals of Regional Science. IX, no. 1 (March 1975), pp. 68-76.

[8]S. I. Gass, "Models in Law Enforcement and Criminal Justice," A Guide to Models in Governmental Planning and Operations, Prepared for Office of Research and Development, Environmental Protection Agency, Washington, D.C., Contract No. 68-01-0788, August 1974, pp. 231-275; see especially pp. 244-261.

Part V.

DEVELOPING AN ANALYTIC UNIT AT THE LOCAL LEVEL

CHAPTER 17

CRIME ANALYSIS: ORGANIZATION OF A TEAM TO IMPACT THE LOCAL CRIMINAL JUSTICE SYSTEM

Charles D. Weller

INTRODUCTION

A professional crime analysis team is an important element in planning and implementing crime reduction activities. The organization and development of a professional crime analysis capability requires advance thought and preparation in order to systematically plan and implement crime reduction activities at the city or county level. The decision to organize this capability suggests that elected officials have reached agreement on the purpose and objectives of the team and that certain assumptions are shared by those who will be affected by the products of the team. This chapter discusses how the crime analysis team can be staffed and made to function with other elements of government.

PRELIMINARY PLANNING FOR A CRIME ANALYSIS TEAM

Far too often preliminary preparations for criminal justice planning are overlooked. The result is that local officials may not approach a local criminal justice planning program with a full understanding of its purpose, functions, and benefits. Traditionally, mayors, city managers, county commissioners, and even governors have passively supported the need for some criminal justice planning, originally stimulated by the need to qualify for grants-in-aid from the Law Enforcement Assistance Administration. Experience has shown that these government officials often fail to fully appreciate or completely exploit the potential of good planning until such a unit has been functioning for a year or two, without their active interest and input. Frequently, attention is drawn to criminal justice planning needs as the result of a critical budget decision in which federal funding emerges as the only solution to a lack of local revenue. Therefore, an important part of the preliminary planning is to familiarize local officials with the need for and benefits of an organized, ongoing crime analysis capability and to illustrate how it can be made to support major procedural and policy decisions in the criminal justice system.

There is a heightened interest on the part of mayors and city managers in acquiring a criminal justice policy analysis capability. Future national funding opportunities may further stimulate this interest. For the purpose of this paper, attention will be focused on the development of a crime analysis capability at the city and/or county level, fully acknowledging its critical value at other levels of government.

Regardless of its organizational location, there are certain critical considerations that should be reviewed and addressed.

1. Does the city and county jurisdiction really want the capability to measure the effectiveness of its local criminal justice system, recognizing the risk that hard data may emerge within two or three years which declares that the existing system may be ineffective and perhaps mismanaged?

2. Is the local political environment and administrative philosophy secure enough to tolerate research, planning, and burdensome data collection without unnecessary interference?

3. Do city and agency administrators have the management sophistication to implement change if research and planning justify certain changes in criminal justice procedures, operations, or resource allocation?

4. Does the local government or its agencies possess some data processing skills and experience with which to technically support the crime analysis team? (This is not to suggest that each jurisdiction must own a computer, but rather should possess minimum and general data collection and processing knowledge and skills.)

If the answer to these four questions is an unqualified "yes," then it is appropriate to proceed to consider locating a crime analysis unit near the elected decision-makers of the jurisdiction. For planning and crime analysis to have impact, the unit must be at a staff position with direct access to the Chief Executive(s).

STAFFING AND RECRUITING A CRIME ANALYSIS TEAM

The principal thrust of this chapter suggests that a finite number and quality of personnel skills are required for a crime analysis team, without respect to the minimum size of the population served. Larger jurisdictions will desire more specially skilled and experienced personnel, but there appears to be a minimum staffing level below which such a unit should not go if any real benefit were to be realized from it.

Jurisdictions with sophisticated personnel departments or civil service units may wish to create job specifications, salary ranges, and qualifications unique to the function of this new unit. Others may wish to appoint the staff without tenure and have the crime analysis staff serve at the pleasure of the elected Chief Executive(s). A third alternative is to do both--certify the staff in the personnel classification system, but require the Director of the unit to serve at the pleasure of the Chief Executive(s). There are some advantages and disadvantages to all three alternatives. Most important, the Director of the unit must build and maintain the confidence of his superiors, because policy analysis usually results in "tough" decision-making, decisions that emerge from force fields representing every facet of community life. It is probably a foregone conclusion that the crime analysis unit will need six months to a year to develop its credibility, but this testing period can be shortened dramatically if the professional qualifications of the staff are exemplary. This means that the recruitment, screening, and hiring of the staff must attain the highest priority and interest of the Chief Executive's office.

If local officials are willing to give this degree of support to the creation of a crime analysis team, there will be reduced pressure on the staff during the first six months of operation when their principal activities are devoted to the collection and analysis of baseline crime and operational event data. Their early activities will take them into the dark records rooms of every local criminal justice agency. Such efforts must also involve the state agencies which serve residents and offenders, etc., within the local jurisdiction.

Staffing the unit with highly qualified professional persons who are capable of interacting with all levels of criminal justice personnel is a prerequisite to building and maintaining a successful crime analysis effort. How they are recruited and linked to local government can be flexible as long as the Chief Executive of the jurisdiction has easy and routine access to their skills and products.

The personnel recruited for the crime analysis team must be reasonably experienced and familiar with operations of the criminal justice system. More important, they should have educational and occupational backgrounds that demonstrate the possession of an analytical mind.

There should be within the team a demonstrated capability to understand and conduct applied research as opposed to theoretical research. The staff should have a "real-world" perspective of the community and the criminal justice system.

The ideal crime analysis team member has had occupational level experience in more than one criminal justice agency, complemented by graduate level education in data collection, data analysis, and elementary research methods and design. Obviously, it is possible to possess all of these characteristics without having been involved in any post-secondary education, but the numbers of these individuals are so few that job qualifications and concomitant salary ranges should be structured to attract the more formally educated and professionally experienced candidates. This should not prove difficult, because there are more and more individuals possessing the educational and experience characteristics needed to pursue careers in this new profession.

No crime analysis "staff per population" ratio guideline has yet been set forth for developers of such a crime analysis unit. The degree of sophistication for planning and analysis tasks currently available within the operating agencies of a jurisdiction may be such that the central city or county team serve more as leaders, coordinators and stimulators. More typical jurisdictions lacking agency-based special analysis units require team personnel with a full range of the minimum skills described below. It is probably a fact that most of the urban areas of the United States are at the point of building upon an existing modest capability.

Without addressing the amount of resources required to support a crime analysis team, it appears from past experience that one-half of a crime analysis team, or a team without all the proper skills, really is not much of an advantage over no team at all. Quite frankly, each jurisdiction must decide to what extent its crime analysis unit should get involved in the broad functions of criminal justice policy analysis as opposed to, or in addition to, federal grants administration.

If the Chief Executive's interest is simply to comply with LEAA grant participation requirements, local and state authorities can provide reasonable estimates of the number of staff necessary to attract and monitor the expenditures of the annual allocation of LEAA dollars to that jurisdiction. However, if an effective "crime analysis and planning unit" is desired that builds its foundation on crime analysis for the purpose of providing decision-makers with hard data about policy, procedures, and effectiveness of a criminal justice agency or system, then the unit must possess a multiplicity of experience and skills no matter how small the population be be served. Smaller jurisdictions should be encouraged to give priority to crime analysis and "pay the freight," so to speak, in the interest of moving their criminal justice agencies toward a better system of accountability for the local tax dollars expended. If this is not an acceptable expenditure for tax payers' dollars by the local unit of government, then decision-makers should not launch half-hearted attempts under the guise of fullblown professional criminal justice planning efforts.

Five full-time professional persons supported by at least one clerical person is a minimum requirement for engaging in full-fledged criminal justice planning. This obviously means that smaller cities and counties must determine that "increasing the productivity of their criminal justice agencies" is nearly the top priority of the community. The author believes it should be, but recognizes that other urban problems in some jurisdictions must take precedence over crime fighting. Be that as it may, if the local units of government "cannot stand the heat, they should stay out of the kitchen," with respect to professional crime analysis and criminal justice planning at the local level. For a weary number of years, too many jurisdictions have desired membership in the profession without fully understanding that municipal and county level policy analysis of any kind will be an expensive proposition if it is to be an effective tool for decision-making by elected officials.

WHAT SHOULD THE TEAM LOOK LIKE?

The Director

Nowhere in the staff of the team is the diversity of experience and skills more important than in the team leader. A key requirement of the Director is the ability to establish, as quickly as possible, professional credibility with criminal justice agency staff and administrators. Further, being able to talk "cop talk" is essential to the courtship that must take place with the Sheriff or Chief of Police. Similarly, knowing what an arraignment is, and the difference between probation and parole, is crucial to the unit's credibility with other elements of the system. Within 60 to 90 days, agency heads will put this person to the test and make the judgements about his/her knowledge and skills that will determine the unit's relationship with that agency far into the future.

The Director should be a generalist with criminal justice operational experience, thoroughly familiar with the operations of the whole criminal justice system. He must have sufficient management skills so that he avoids being a "do it your-selfer." He must believe in what he is doing. He must be able to identify problems from data and conversation and set objectives and assign tasks to staff who will be held accountable for product completion.

This means that the Director must be generally familiar with quantitative analysis, systems analysis, and financial accountability. While he need not be an expert in each of these fields, he will use his general knowledge to make himself meaningfully conversant with the "tools" of the trade and the experts in special areas that may provide special advice.

The crime analysis unit Director must be a communicator, striving constantly to integrate information and form bridges between disassociated criminal justice agencies. He is a "justice of the peace" bringing together administrators who traditionally have preferred to remain isolated and protective of their own areas in the overall system. He must be useful; demonstrating, not asking for professional recognition of his competence. And finally, he must be able to conceptualize and sell action, and serve as an expeditor in getting proposed changes implemented.

The Director and his staff, as a team, must show a service commitment to assisting the operating agencies. For example, the unit should plan with, not for the police department, so that real services and assistance tasks can be identified and delivered. The Director continually searches for collateral activities to conduct that make the criminal justice agency administrator's job easier. This is a creative task of matching staff competence with opportunities to assist. Clinically, crime analysis teams should be considered a "helping profession."

If the Director possesses these qualities, he will stand the early credibility tests, and the unit will have access to agency files, records, procedures, and even confidential problems which really are the cornerstones of a local crime analysis unit. Without hard data, the unit is nothing. Over time, the Director's elected supervisors should trust him as much as their campaign managers; yet a sensitive line should separate the crime analysis unit from elected officials and their subordinate agency administrators. The Chief Executive must set forth the ground rules with agency heads to prevent them from making end-runs around the crime analysis unit. The team's special status needs to be openly protected. In effect, all the jurisdictions' agencies must administratively recognize and respect the crime analysis unit for its expertise.

The skills and attributes of the remaining four members of the team do not necessarily suggest any particular organizational hierarchy. That can be organized according to the local decision-makers' assessment of the unit's functions and the abilities of its personnel. With only five or six persons involved, the hierarchy should be flat, all reporting to the Director. When the unit reaches ten to twenty staff members (as in larger jurisdictions) with more specialized personnel, an additional management person may be justified.

The Data Expert

One member of the team must be trained and educated to bring sophisticated quantitative skills to the unit. He or she should possess system analysis skills, statistical analysis skills, data processing skills, and the ability to report on the results of data analysis activities. If the data expert also knows the criminal justice system well, he or she is the ideal candidate. If a choice must be made between the person with technical

skills and no criminal justice experience, as opposed to the person with the desired experience and weak technical skills, the Director should choose the technical expert. Manipulating criminal justice data over time will eventually cause the technical expert to achieve vicarious occupational experience in the criminal justice system.

The Police Specialist

There are some functional speciality skills required on the team. One staff member must be extensively familiar with the field of law enforcement. Today many college-educated patrol officers are leaving the security of the police to enter criminal justice planning, so an excellent manpower pool is available. However, merely having some patrol experience and a college degree does not qualify one for entry to the unit. Among patrol officers with these qualifications, the Director should seek candidates who demonstrate a propensity for analytical problem solving, individuals who appear to enjoy working with data, and those who write with some command of the English language. These individuals are likely to possess skills which can be broadened to accomplish many tasks not integral to police work but nevertheless important to a police agency.

Courts/Corrections Staff

The basic skills required for crime analysis at the local level do not indicate the need for a full-time attorney or even a full-time correctional specialist. What is needed is an individual who knows and understands the judicial process and dispositional alternatives for convicted criminal offenders. It is readily possible to find staff who have education and/or experience in adjudication and correctional programs who, similar to the patrol officer, possess analytical interests. One staff person can adequately struggle with local courts and corrections problems as his principal functional specialty, if a basic team unit is all that can be afforded. As the fourth person on the team, he does double-function duty, but it is expected that the team will have a well-rounded technical expert to help with adjudication and correction problems, and identification and analysis duties.

The Budget Analyst

Finally, the fifth person is a budget analyst with an accounting background, education, and/or experience. And, if the financial analyst has worked within one or more criminal justice agencies, he is the ideal candidate. The unit does not need a "budget butcher." Rather, the skills of an analytical accountant and financial manager are desired to assist in "budget building." Here is the one person on the team who should possess technical expertise desired by every operating agency. Furthermore, the availability of this team member to provide technical assistance to the operating agencies should be a publicly known fact. He must strive to change attitudes about accountability, work unit measurement, budgeting, forecasting needs, assessing past-expenditure experience, and most of all, he must sincerely believe and portray his image as a member of the total team, not just an accountant for the program oriented staff.

THE CRIME ANALYSIS TEAM

Ideally, all members of the crime analysis team must be able to write and speak effectively. They should be accustomed to meeting firm deadlines. Team members, as well as the Director, should be able to recognize opportunities to be useful to the operating agencies and seize upon these opportunities to act as a catalyst for effecting other changes already deemed desirable.

A special aptitude for marrying the practical and academic worlds will help the team build personal relationships with agency personnel that are based upon competence, rather than authority. Finally, the crime analysis team itself must tolerate isolation, both personal and academic. The unit functions best when it is only one-half a friend to everyone. That makes the job lonely, but one, nevertheless, that offers each member of the team a great deal of personal pride, for each undoubtedly will contribute to the achievement of the impossible--successfully changing the criminal justice system in America.

Part VI.

CRIME-ORIENTED PLANNING AND PROGRAM DEVELOPMENT

CHAPTER 18

CRIME-ORIENTED PLANNING AND PROGRAM DEVELOPMENT

R. Allen Payne

INTRODUCTION

In January, 1972, the Law Enforcement Assistance Administration launched the High Impact Anti-Crime Program, a major demonstration which provided up to $20 million to eight large American cities to reduce the incidence of stranger-to-stranger crime and burglary by 5 percent in two years and 20 percent in five years. Underlying this program was the concept of "crime-oriented planning": that crime (and particularly, specific classes of crime) was the starting point for criminal justice program development, not the symptoms of crime, such as overworked or understaffed police, courts, or corrections agencies.

DATA -- THE BASIS FOR PROGRAM DEVELOPMENT

The basis for planning and program development in the crime-oriented planning model is the collection and analysis of data: crime data, which tells us how, when, where, and with what frequency individual crimes such as robbery, rape, or burglary are committed; offender data, such as age, sex, race, and personal characteristics of those who commit crimes; and victim data which indicates the characteristics and responses of those who have been victimized.

The analysis of these data need not necessarily involve extensive computer operations and programs. Basic information about a crime incident frequently is available from police reports. Narrative incident reports are the source of additional information. However, as LEAA discovered in the Impact Program, data availability, quality, and utility can present significant initial difficulties, even in larger cities. On the other hand, transforming existing data into useful, retrievable data is not an insurmountable task.

Atlanta used a small early Impact grant to redesign its police field reporting system. Denver, in a modified manual replication of the CAPER System functioning in Santa Clara County, California, added to the offense and arrest information available through the Police Department by hiring part-time undergraduate coders to cull additional desired information from police reports and case files. Eventually student coders were doubled in number and assigned similar tasks in every criminal justice agency in Denver, thus providing for an expanded base both of crime and systems response data.

Although data sources such as Uniform Crime Reports, Offender Based Transaction Statistics (OBTS) and victimization surveys are discussed in detail elsewhere in this manual, a further word about data is in order.

Data needs for crime-oriented planning are hierarchical, as is the planning model itself. That is, the state planner addresses statewide burglary rates and geographic concentrations of burglary. The city planner focuses on burglary occurrence by frequency, type (commercial vs. residential), time, and location (census tract or police district). The project level planner concentrates more on methods of operation, short term trends in frequency, and specific time and place of occurrence.

PROBLEM IDENTIFICATION

Analysis of data, therefore, gives us a profile of individual crimes by characteristics of the crime event, setting or environment, offender, and victim. It also tells us in quantitative terms how the criminal justice system responds to the commission of certain offenses. From this analysis comes the identification of problems: crime problems (e.g., the 15 percent rise in the statewide robbery rate or the 20 percent rise in citywide residential burglaries); offender problems (e.g., juveniles comprising 30 percent of burglary arrestees); victim-related problems (e.g., preponderance of elderly robbery victims within a certain census tract); systems problems related to the response to specific crimes (e.g., city court burglary cases are backlogged with a 240 day average from arrest to trial). Problems then must be prioritized, which leads to the next step, the establishment of goals.

CRIME-ORIENTED GOAL* SETTING

Once a specific crime problem has been identified, there are two basic responses: reduce the cause of the problem and improve control of the problem. Cause is reduced either by attacking underlying social, economic, or other basic conditions that promote the commission of crime. Obviously, much program activity in this regard is outside the authority and responsibility of the criminal justice system. Nonetheless, social, demographic, and other economic data related to crime analysis can point to obvious areas of interagency or intergovernmental cooperation of benefit to the community in areas beyond that simply of crime control. On another level, "cause" can be reduced by intervening in criminal careers through programs involving identified offender groups or through what frequently are called "prevention" programs, such as youth service bureaus or school attendance projects in high juvenile crime areas.

Control is improved either by reducing opportunity for the commission of criminal acts or by increasing the risks of offending. Examples of the

*The term "goal" is sometimes used interchangeably with the term "objective." In either case, the attainment of a goal or an objective must be judged on a measurable interim or final achievement.

former include target-hardening activities such as physical security improvement or property identification. Examples of the latter include concentrated police patrol and the speedy adjudication of target crime cases.

Crime-oriented planning, therefore, leads to the establishment of comprehensive program goals that involve the whole system of criminal justice. It is not simply a police suppression model.

Goals should be hierarchical and quantitative. There may be initial resistance by operating agency personnel to the establishment of quantitative goals. However, quantitative goals provide for clarity of purpose and the measurement of program/project results. They are an integral component of crime-oriented planning.

In hierarchical terms, at the state level, a major program goal might be the reduction of burglary by 5 percent within one year in the five largest metropolitan areas. At the city level, this goal might translate to a reduction of residential burglaries by 10 percent and a reduction of commercial burglaries by 5 percent.

PROJECT DEVELOPMENT

The selection and development of projects continues this process of quantification. Tactical alternatives to meet state or city goals must be weighed and chosen. What type or types of program activities respond most directly and effectively to identified problems and established goals? The utility of data collection and problem analysis is severely hindered if the full examination of alternatives is shortchanged, or if data is collected and molded solely to justify the selection of a predetermined project activity.

The process of alternative selection takes place through a screen of real-world constraints that can be political, legal, economic, social, moral or ethical in nature. Highly competent planners and a sophisticated rational supervisory board might agree that a particular alternative offers the best possibility for reducing the target crime, yet they may not select that alternative for implementation because it is economically unfeasible, patently illegal, or otherwise contrary to prevailing community standards. Constraints are natural and they must be understood by planners; the difficulty arises in not letting the possibility of constraints that may be imposed unduly hinder the planner in the full development and examination of alternate tactics and programs.

The selection from among alternatives yields a plan of related program activities that can span the criminal justice system and include some activities even outside the system. Selected activities can include programs related to the prevention of the target crime; the detection, deterrence, and apprehension of offenders; the diversion of offenders from the criminal justice system; the adjudication of offenders; and the rehabilitation of offenders, through both institutional and non-institutional programs. Data collection and analysis also may point out the need for legislative change, additional data and information systems, or additional research, all of

which frequently are overlooked in the search for alternative courses of action.

Based upon choices from among alternatives, we can imagine the following project activities and goals developed in accordance with the reduce cause - improve control approach to a city-level burglary reduction goal:

1. A property identification project to enroll X persons and mark Y amount of property within the census tracts with highest burglary rates.

2. A concentrated crime patrol aimed at providing X amount of patrol and Y response time during peak hours in the high burglary census tracts, increasing by Z the amount of burglary arrests.

3. A reduction in court backlog for burglary cases by X, and an increase in burglary convictions by Y, through the addition of judicial and prosecutorial personnel.

4. Offender-based programs for burglary offenders such as drug diversion (where offender drug abuse is substantial) or juvenile diversion projects (where juveniles are found to constitute a substantial percentage of burglary offenders).

Several comments can be made about these project examples. First, many of the particular activities undertaken as a response to crime-oriented problem analysis are not necessarily different or unique from those that are undertaken where reliance on data is less extensive. The difference is that these activities are undertaken in response to specifically identified and bounded problems and their goals and operation, location, and scope of services or products are carefully related to the magnitude and nature of the problem. Second, products or results (effectiveness objectives) and levels of service (efficiency objectives) are clearly specified, providing for clarity of purpose and the evaluation of results and performance.

PERFORMANCE MEASUREMENT

In criminal justice, evaluation is a concept whose time has come. It remains an integral link in the crime-oriented planning chain of data-based problem analysis and goal formulation. Good evaluation depends upon a clear statement of project purpose, and indeed, good evaluation begins with the establishment of quantified, time-phased project goals, objectives, and milestones at the project design stage.

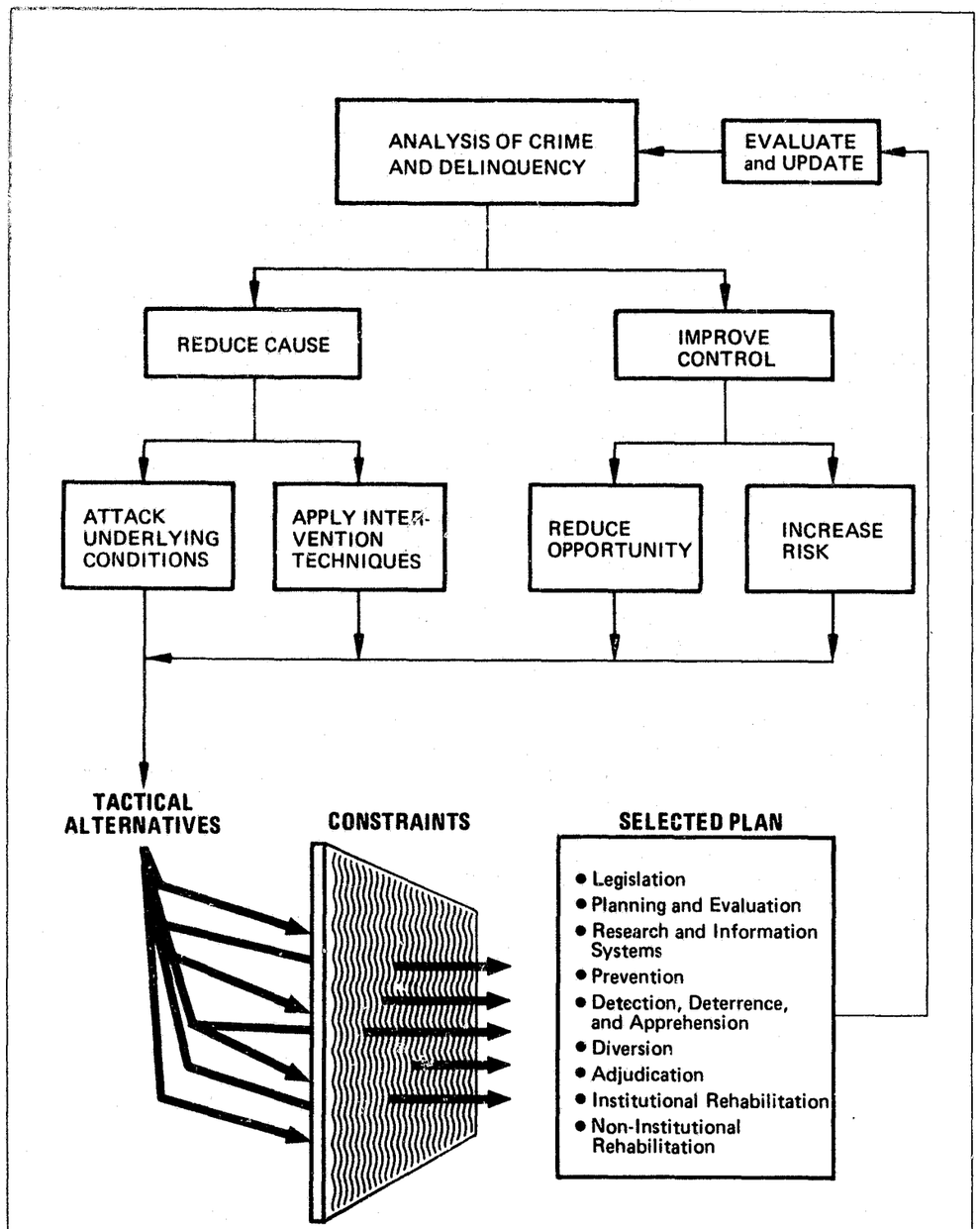
A direct link to specific crime reduction goals cannot be established for every project, nor can elaborate control or experimental groups be utilized in every case. Nonetheless, project level self-assessment can be accomplished by any grantee where initial attention has been paid to clearly delineating the goals of the project and where sufficient attention has been paid to defining performance objectives and assuring for the availability, collection, and analysis of data.

Linkages are as important to the evaluation process as they are to the earlier stages of problem analysis and goal formulation. After a project strategy to achieve a program goal has been selected, effectiveness objectives and efficiency objectives must be established. Examples of the former can include specific crime reduction objectives, or objectives addressing recidivism reduction or decreased court processing time. Activity objectives "operationalize" or explain quantitatively what levels of services or activities will be performed or provided. For both outcome and activity objectives, specific data measures are required to assess project progress. After data elements are identified, a plan must be prepared for the collection, analysis, and interpretation of the data.

Several general comments about evaluation of crime-oriented programs are in order. "Hard" evaluation is possible even where objectives appear "soft". As the evaluation design of the St. Louis Home Detention Impact project shows, initial objectives such as "Keeping [assigned] youths trouble free and available to the court" are amenable to the establishment of specific effectiveness measures. Second, evaluation is not the end of the crime-oriented planning cycle but the beginning of the recycle. Projects such as Denver Impact's Special Crime Attack Team (SCAT), which continuously uses evaluation data to adjust its concentration on target areas and target crimes, bears out the effectiveness and utility of evaluation results upon project operations.

CONCLUSIONS

Crime-oriented planning techniques, in summary, reflect first a classical, accepted framework for problem solving, and second, a reliance on data and data analysis for the identification of problems and priorities, the establishment of goals, and the evaluation of results. Stripped bare of planner's jargon, this approach is an effective aid in telling us what our crime problems are, what responses appear warranted, what kinds of performance and activities we can and should expect, and whether in fact desired results were achieved. (See Figure 1 for a graphic presentation of crime-oriented program development.)



Source: IMPACT (LEAA brochure), 1973.

Figure 1. CRIME-ORIENTED PROGRAM DEVELOPMENT

INDEX

*U.S. GOVERNMENT PRINTING OFFICE:1975 211-063/535 1-3

END