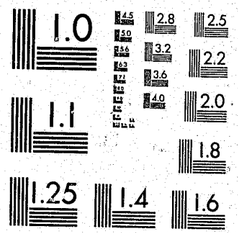


MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

U.S. Department of Justice
Bureau of Justice Statistics

Proceedings of the Second Workshop on Law and Justice Statistics

Statistics in the policy process

Statistical methodology in law and justice statistics

istical analysis in the courts

istical issues at the State level

statistician as expert witness

Bureau of Justice Statistics reports
(revised June 1984)

Single copies are available free from the National Criminal Justice Reference Service, Box 6000, Rockville, Md. 20850 (use NCJ number to order). Postage and handling are charged for multiple copies: up to 10 titles free; 11-40, $10; more than 40, $20; libraries call for special rates (301/251-5500).

Public-use tapes of BJS data sets and other criminal justice data are available from the Criminal Justice Archive and Information Network, P.O. Box 1248, Ann Arbor, Mich. 48106, (313/764-5199).

National Crime Survey

Criminal victimization in the U.S.:
- 1973-82 trends, NCJ-90541, 9/83
- 1981 (final report), NCJ-90208
- 1980 (final report), NCJ-84015, 4/83
- 1979 (final report), NCJ-76710, 12/81

BJS special reports:
- The economic cost of crime to victims, NCJ-93450, 4/84
- Family violence, NCJ-93449, 4/84

BJS bulletins:
- Criminal victimization 1983, NCJ-93869, 6/84
- Households touched by crime, 1983, NCJ-93658, 5/84
- Violent crime by strangers, NCJ-80829, 4/82
- Crime and the elderly, NCJ-79614, 1/82
- Measuring crime, NCJ-75710, 2/81

The National Crime Survey: Working papers, vol. I: Current and historical perspectives, NCJ-75374, 8/82

Crime against the elderly in 26 cities, NCJ-76706, 1/82

The Hispanic victim, NCJ-69261, 11/81

Issues in the measurement of crime, NCJ-74682, 10/81

Criminal victimization of California residents, 1974-77, NCJ-70944, 6/81

Restitution to victims of personal and household crimes, NCJ-72770, 5/81

Criminal victimization of New York State residents, 1974-77, NCJ-66481, 9/80

The cost of negligence: Losses from preventable household burglaries, NCJ-53527, 12/79

Rape victimization in 26 American cities, NCJ-55878, 8/79

Criminal victimization in urban schools, NCJ-56396, 8/79

Crime against persons in urban, suburban, and rural areas, NCJ-53551, 7/79

An introduction to the National Crime Survey, NCJ-43732, 4/78

Local victim surveys: A review of the issues, NCJ-39973, 8/77

National Prisoner Statistics

BJS bulletins:
- Capital punishment 1983, NCJ-93925, 7/84
- Prisoners in 1983, NCJ-92948, 4/84
- Prisoners 1925-81, NCJ-85861, 12/82

Prisoners in State and Federal institutions on December 31, 1981 (final report), NCJ-86485, 7/83

Capital punishment 1981 (final report), NCJ-86484, 5/83

1979 survey of inmates of State correctional facilities and 1979 census of State correctional facilities:

BJS special reports:
- The prevalence of imprisonment, NCJ-93657, 6/84
- Career patterns in crime, NCJ-88672, 6/83

BJS bulletins:
- Prisoners and drugs, NCJ-87575, 3/83
- Prisoners and alcohol, NCJ-86223, 1/83
- Prisons and prisoners, NCJ-80697, 2/82
- Veterans in prison, NCJ-79632, 11/81

Census of jails and survey of jail inmates:
- Jail inmates 1982 (BJS bulletin), NCJ-87161, 2/83
- Census of jails, 1978: Data for individual jails, vols. I-IV, Northeast, North Central, South, West, NCJ-72279-72282, 12/81
- Profile of jail inmates, 1978, NCJ-65412, 2/81
- Census of jails and survey of jail inmates, 1978, preliminary report, NCJ-55172, 5/79

Parole and probation

BJS bulletins:
- Probation and parole 1982, NCJ-89874 9/83
- Setting prison terms, NCJ-76218, 8/83

Characteristics of persons entering parole during 1978 and 1979, NCJ-87243, 5/83

Characteristics of the parole population, 1978, NCJ-66479, 4/81

Parole in the U.S., 1979, NCJ-69562, 3/81

Courts

The prosecution of felony arrests, 1979, NCJ-86482, 5/84

Habeas corpus (BJS special report), NCJ-92949, 3/84

State court caseload statistics: 1977 and 1981 (BJS special report), NCJ-87587, 2/83

State court organization 1980, NCJ-76711, 7/82

State court model statistical dictionary, NCJ-62320, 9/80

A cross-city comparison of felony case processing, NCJ-55171, 7/79

Federal criminal sentencing: Perspectives of analysis and a design for research, NCJ-33683, 10/78

Variations in Federal criminal sentences, NCJ-33684, 10/78

Federal sentencing patterns: A study of geographical variations, NCJ-33685, 10/78

Predicting sentences in Federal courts: The feasibility of a national sentencing policy, NCJ-33686, 10/78

State and local prosecution and civil attorney systems, NCJ-41334, 7/78

Expenditure and employment

Justice expenditure and employment in the U.S., 1979 (final report), NCJ-87242, 12/83

Justice expenditure and employment in the U.S., 1971-79, NCJ-92596 (forthcoming)

Privacy and security

Computer crime:
- Electronic fund transfer and crime, NCJ-92650, 2/84
- Computer security techniques, NCJ-84049, 9/82
- Electronic fund transfer systems and crime, NCJ-83736, 9/82
- Legislative resource manual, NCJ-78890, 9/81
- Expert witness manual, NCJ-77927, 9/81
- Criminal justice resource manual, NCJ-61550, 12/79

Privacy and security of criminal history information:
- A guide to research and statistical use, NCJ-69790, 5/81
- A guide to dissemination, NCJ-40000, 1/79
- Compendium of State legislation: NCJ-48981, 7/78 1981 supplement, NCJ-79652, 3/82

Criminal justice information policy:
- Research access to criminal justice data, NCJ-84154, 2/83
- Privacy and juvenile justice records, NCJ-84152, 1/83
- Survey of State laws (BJS bulletin), NCJ-80836, 6/82
- Privacy and the private employer, NCJ-79651, 11/81

General

BJS bulletins:
- Federal drug law violators, NCJ-92692 2/84
- The severity of crime, NCJ-92326, 1/84
- The American response to crime: An overview of criminal justice systems, NCJ-91936, 12/83
- Tracking offenders, NCJ-91572, 11/83
- Victim and witness assistance: New State laws and the system's response, NCJ-87934, 5/83
- Federal justice statistics, NCJ-80814, 3/82

Sourcebook of Criminal Justice Statistics, 1983, NCJ-91534, forthcoming 10/84

Report to the nation on crime and justice: The data, NCJ-87068, 10/83

BJS five-year program plan, FY 1982-86, 7/82

Violent crime in the U.S. (White House briefing book), NCJ-79741, 6/82

Dictionary of criminal justice data terminology: Terms and definitions proposed for interstate and national data collection and exchange, 2nd ed., NCJ-76939, 2/82

Technical standards for machine-readable data supplied to BJS, NCJ-75318, 6/81

Justice agencies in the U.S., 1980, NCJ-65560, 1/81

Indicators of crime and criminal justice: Quantitative studies, NCJ-62349, 1/81

A style manual for machine-readable data, NCJ-62766, 9/80

See order form on last page.

PROCEEDINGS OF THE SECOND WORKSHOP ON LAW AND JUSTICE STATISTICS 1983

Edited by Alan E. Gelfand

Papers presented in Toronto, Canada, August 13-14, 1983, at a workshop developed by the American Statistical Association Committee on Law and Justice Statistics.

Workshop sponsored by the Bureau of Justice Statistics and the American Statistical Association's Continuing Education Department with special acknowledgment to the Criminal Justice Statistics Association.

The papers and discussions in this proceedings volume are reproduced essentially as received from the authors. The papers in this volume have been reviewed by the editor but have not been subjected to a formal refereeing process. Statements or positions taken by the authors do not necessarily represent the views of either the Bureau of Justice Statistics or the American Statistical Association.

Foreword

The Bureau of Justice Statistics is pleased to have sponsored the second biannual "Workshop on Law and Justice Statistics" held in conjunction with the American Statistical Association's 1983 Annual Meeting in Toronto, Canada. The workshop, organized and conducted by the ASA Committee on Law and Justice Statistics, provided a forum through which criminal justice statisticians and researchers, in both the private and public sectors and at all levels of government in the United States and Canada, could meet to exchange ideas and discuss current work. Furthering such communication is fundamental to the BJS goal of improving the availability and dissemination of high quality statistical information on crime and the criminal justice system.

By publishing these Proceedings, we seek to share the papers with a much larger audience than the individuals who were able to attend the sessions. We believe that the potential benefits of peer interaction will be magnified many times as those who work daily with criminal justice statistics, whether in government, academia, or the courtroom, read the contributions of experts in various fields.

The Bureau's interest in supporting the work of the Committee on Law and Justice Statistics goes beyond biannual workshops. The committee represents a resource to the Bureau of extensive expertise in the field of criminal justice statistics, a resource that the Bureau frequently draws upon for a variety of purposes. Most significant, perhaps, is the accessibility to methodological and statistical talents to address thorny issues in data collection and statistical application. During 1983, such "methodological reviews" were conducted for four BJS data collection programs: the quinquennial jail census and survey of inmates and the quinquennial prison census and survey of prisoners.

The committee also recommends peer reviewers for many BJS reports and publications. In the same manner, it identifies and schedules individuals for in-house seminars for BJS staff and their counterparts in State and local governments; such seminars help the attendees maintain statistical and analytic skills and keep them current with new advances in statistical applications in criminal justice.

The relationship that has developed between the American Statistical Association's Committee on Law and Justice Statistics and the Bureau of Justice Statistics has proven to be a mutually satisfying one that allows the Federal Government to utilize the talents and experience of national experts in criminal justice statistics, allows feedback from highly sophisticated users of BJS data, and allows input from those users and others into the design of data collection and analysis and into dissemination strategies to increase the utility of the data for a wide range of analytic needs. It is an example of BJS's commitment to produce high quality data that meet the needs of all of its potential users.

Steven R. Schlesinger
Director
Bureau of Justice Statistics

This conference was conceived, designed, and arranged by the Committee on Law and Justice Statistics of the American Statistical Association. Our committee was created to provide an interface for the Association with the legal, judicial, and criminal justice communities. As such it attempts to:

(1) help to disseminate information about legal and justice statistics activities throughout the statistics community;

(2) promote the development of quality statistical activities in civil and criminal justice settings;

(3) consider and report upon relevant issues to guarantee the integrity of statistical programs maintained by the U. S. Department of Justice, and other appropriate agencies and organizations.

One of the activities our committee has elected to be involved in is to offer a biennial workshop to workers and researchers in the criminal justice community who are interested in methodology. This conference is the second such effort (the first took place in Detroit, in August, 1981).

We hope that by recording the technical papers of this conference in a Proceedings volume, the criminal justice community will be served by making its workers aware of important methodological issues with which they must be concerned. We hope that people from diverse disciplines who are interested in the methodological aspects of law and criminal justice system will be brought closer together.

We are grateful to the American Statistical Association and to the Bureau of Justice Statistics for their sponsorship and their strong emotional support.

S. James Press
Chair, Committee on Law and
Justice Statistics of the
American Statistical Association

Preface

These Proceedings document the presentations given at the Second Workshop on Law and Justice Statistics held just prior to the national American Statistical Association (ASA) meetings on August 13 and 14, 1984, in Toronto, Canada. This workshop was developed by the ASA's Committee on Law and Justice Statistics (CLJS). It was sponsored by the Bureau of Justice Statistics (BJS) and the ASA's Continuing Education Department with assistance from the Criminal Justice Statistics Association (CJSA). The first such workshop was held in Detroit, Michigan, in 1981 prior to that year's annual ASA meeting.

By all accounts, this workshop was a success. It brought together, from both Canada and the United States, more than eighty individuals, each with a strong interest in law and justice statistics. The participants included academics and attorneys, government employees and private consultants bringing training in statistics, law, criminology, sociology, economics and political science. There were sessions on statistics in the policy process, statistical methodology for law and justice statistics, statistical issues at the state level, statistical examination of the courts and the statistician as expert witness.

The ASA's CLJS began as an ad hoc committee of six in 1977. It became a standing committee in 1980 and has grown to twelve members with an agenda much increased in scope. One of its primary activities involves its contractual relationship with BJS to supply both a methodological and educational capability. As Chair of the Education subcommittee, I will comment briefly on the three areas within the educational sphere which are being pursued.

First the committee has arranged several in-house seminars, usually on statistical subjects, requested by BJS. Sue A. Lindgren, our BJS liaison person, has been very helpful in identifying subject matter and coordinating the presentations. A second area involves the development of workshops at the state level. Here the initiative and commitment of Thomas A. Henderson, Executive Director of the CJSA and CLJS member, have been invaluable.

The last area is the creation of a workshop at the national level joining together researchers and practitioners in law and justice statistics. Again, these Proceedings record the contributions of an impressive gathering of such individuals. The articles appearing herein have been reviewed and edited. It is hoped that the reader will find much that is interesting and stimulating. Planning has already begun for the Third Workshop on Law and Justice Statistics!

Thanks are due to numerous individuals who helped to either bring the second workshop together or this volume to fruition. They include the other members of the ASA Committee on Law and Justice Statistics educational subcommittee, Kathy Wallman, George Woodworth, Albert Reiss, Jr., and Thomas Henderson. They include Sue A. Lindgren and Marilyn Marbrook at BJS. Finally, they include the people at ASA who have provided unflinching support and aid: Fred Leone, Jill Stormer, Ede Denenberg, Jo Przystas and Irene Stefanski.

Alan E. Gelfand
Editor
Workshop Organizer
Vice Chair, ASA-CLJS

The papers and discussions in this Proceedings volume are reproduced essentially as received from the authors. The papers in this volume have been reviewed by the editor but have not been subjected to a formal refereeing process. Statements or positions taken by the authors do not necessarily represent the views of either the Bureau of Justice Statistics or the American Statistical Association.

Providing Statistical Support to the Policy Process

Sue A. Lindgren
U.S. Bureau of Justice Statistics

Abstract

At the American Statistical Association's 1981 annual meeting, Timothy Sprehe of the Office of Management and Budget presented a paper on "Implementing a New Federal Data Access Policy." In it, he stressed the importance of making statistical data more accessible to government policymakers by making the data readily comprehensible to nontechnical users and by providing statistical assistance in the use of those data. This paper reports what one agency, the Bureau of Justice Statistics, is doing to ensure that its data are being used in the policy process.

History of the BJS program

The Department of Justice's statistical program began in 1969 as the National Criminal Justice Information and Statistics Service of the now defunct Law Enforcement Assistance Administration. Its early efforts were directed at centralizing in one agency existing justice statistical programs throughout the federal government (programs such as the Juvenile Detention and Correctional Facility Surveys from the Department of Health, Education, and Welfare and the National Prisoner Statistics Program from the Bureau of Prisons) as well as establishing new data series such as the National Crime Survey. The early reports of NCJISS, as the agency was then called, were fat books crammed full of tables, with scant narrative explanation in technically precise, stilted language.

The reports were nearly impenetrable to many policymakers and their staffs. People who needed the data for decisionmaking were usually extremely busy and, even if they could find the data they were looking for, most likely did not have the time to decipher or summarize their content. As a result, when our numbers were used, they often were used incorrectly.

To help policymakers and many other users of our data, we began to include "executive summaries" in the front of our reports. These summaries highlighted the most important findings and tried to avoid the stilted style of operational definitions and statistical jargon.

We also tried the use of "table finding guides," modeled after those in decennial census publications, to lead users to the tables containing the numbers they desired.

These early attempts provided easier access to our data, but they fell far short of the type of data access program envisioned by Sprehe. Today we are embarked on a program to get our data into the hands of policymakers in a format that is easy to understand and use. This program takes two forms: proactive and reactive strategies to improve the use of our data in the policy process.

Our proactive strategy

In our proactive approach to improving the use of statistics in the policy process, we radically changed how we select topics for data analysis. In the past, we looked to the data series we maintained and attempted to mine those data as fully as our resources would allow. As part of our new program, we identify issues of current policy interest and try to use our data, sometimes along with non-BJS data, to address those issues. For example, we recently published, or are working on, analytic products on such topics as—

Career patterns in crime.
Profiles of the prison population.
The possibility that prisoners can be treated in less expensive environments without endangering society.
How much crime could be prevented by changes in parole laws.
The amount of burglary that could be prevented by simply ensuring that houses are adequately locked and secured when no one is at home.
Trends in habeas corpus filings.
Recidivism after release from prison.
Federal drug offenders.
The use of drugs by state and federal offenders.

We also made a major change in our data dissemination and publication program. We now emphasize shorter, less technically written reports, while continuing to publish the more familiar books of statistical tables for users who desire access to the more detailed data they provide.

Our newer publications, BJS Bulletins and BJS Special Reports, are limited to four to six pages. This limit is based on the assumption that busy policymakers and the general public will be more inclined to read a brief report than a 200-page book of tables. The topics are narrowly defined and most often are of current public debate and interest (such as drug use among prison inmates ... trends in prison populations ... households touched by crime ... and trends in court caseloads).

We strive to write these bulletins and reports in language readily understandable to nontechnical users. Readers are referred to the regular report series for technical information such as descriptions of survey design, standard errors, elaboration on definitions of terms, and copies of survey forms.

Bulletins are produced on a monthly basis, while Special Reports supplement the bulletins on an unstructured periodic basis. We have found these newer publications to be highly successful. Data from them appear in media accounts, expanding the audience exposed to the findings, and the findings are used more frequently than in past by local and State policymakers by the Department of Justice, the White House and the Congress.

Another proactive strategy to make data more understandable and accessible to policymakers and the general public is an ambitious undertaking to write a comprehensive book on crime and criminal justice using all available data sources. This publication, Report to the Nation on Crime and Justice, is the culmination of over two years of staff effort. Unlike many BJS reports, rather than taking a single data set and attempting to describe what the data say on a particular topic, the National Report began with an overview of the justice system and sought data to address the various issues. Where data are not available, the particular process or issue is described and the paucity of data noted.

The National Report adopted a "magazine" or "newspaper" format, with headlines and table titles that convey information rather than describe the variables in the table or text. For example, a traditional table title that might read:

> "Percent distribution of justice spending by level of government, 1979"

appears in the National Report as:

> "State and local governments pay 87% of all government costs for criminal and civil justice."

The National Report also relies heavily on graphics because they are useful for presenting quantitative information to nontechnical audiences. It is printed in four colors to further improve accessibility. The underlying notion is to produce a book that is inviting to the reader and that would, once the reader began, hold his interest.

The National Report was published in December 1983 and received extensive media attention. Officials in Congress, the Department of Justice, and the White House have praised it and requests for copies have far outpaced any report BJS has published in the past. Within the first three weeks, more than 20,000 copies were ordered, compared with a usual print order of considerably less than 10,000 for other BJS reports. The reception the National Report has received clearly indicates the need for data producers to provide statistical information in a nontechnical format.

The Sourcebook of Criminal Justice Statistics takes a different type approach to making information readily available to policymakers and others in need of criminal justice data. This book of close to one thousand pages is virtually limited to tabular presentation of data; there is no text describing the data, only very brief text describing the availability of data on various topics. The Sourcebook, in essence, is a Statistical Abstract for criminal justice. It makes available, in a single volume, reference data from close to a hundred different sources, many of which are not well-known or easily obtainable. While the basic sources contain much more detail than is presented in the Sourcebook, the Sourcebook contains the data that are most frequently needed and used.

Another proactive strategy to increase the policy-relevance of BJS products is our external analysis program. The BJS staff is very limited in size; thorough examination of particular topics is simply beyond our resources. By using researchers in academia and the private research community, we have been able to tap expertise external to BJS and to bring fresh insight to issues of current policy debate.

In developing this external analysis program, we solicited ideas for analytic projects from leading criminologists throughout the country. We requested that the proposals use BJS data, if at all possible, and place emphasis on such topics as—Career criminals ... recidivism ... the deterrent and incapacitative effects of incarceration ... plea bargaining ... police patrol strategies ... the deterrent effect of police arrests ... bail reform ... and the insanity defense. Nine modestly priced projects were funded in 1983, and we expect that a number of them will produce results that can be published in our Special Reports series.

A larger, more ambitious component of our external data analysis program involves soliciting proposals for more extensive projects. More than forty proposals were received, and we expect to fund ten in 1984.

Our reactive strategy

In addition to our proactive program to provide data to inform the policy process, we have continued our reactive role of providing statistical data and support to policymakers when requested. Our most ambitious effort in this area was our participation in the National Indicators System. The function of the system is to brief the President, the Vice President, and the White House staff of social, demographic, and economic trends in the United States.

Our briefing for the White House was held in September 1981 and used a briefing book prepared by BJS according to specifications developed by Dr. Richard Beal, then Director of the White House Office of Planning and Evaluation. The format of the briefing book is almost entirely graphical. Each graphic is followed by several short statements that interpret the graphic or present additional information on the topic.

Several aspects of the National Indicator System as a tool for policy formation should be appreciated:
- The stated intent is to provide statistical information relevant to the policy process, but prior to policy development.
- The data are presented independent of advocacy on policy positions, budget, and legislative issues.
- Because of time constraints imposed by a limit of thirty minutes for the briefing, the issues presented were selected to highlight a number of crucial messages on the subject of criminal justice.

After the briefing, we published the briefing book as Violent Crime in the United States.

Beyond participating in the National Indicator System, BJS has been providing statistical data and support for a number of developing programs in the Department of Justice. Our work in these areas has been much more extensive over the past two years for several reasons.

- Before BJS was established as an organizational entity, the statistical program was buried in the Law Enforcement Assistance Administration with little visibility—people in and outside the federal government simply did not know that the statistical program existed or were unaware of the extent of the program.

- Before the passage of our authorizing legislation, calling for a Director to be appointed by the President and confirmed by the Senate, our program was headed by career civil servants who did not have ready access to the politically appointed policymakers in the Department of Justice. Those in need of justice statistics now know whom to contact, and they often attend high-level policy-oriented meetings in the Department with the BJS Director.

- The independence of BJS establishes it as a credible, objective agency whose data are perceived as being free from advocacy bias. Data are useful in the policy process only to the extent they have been produced objectively and are perceived as being objective. In the past, statistical products were produced by the statistical arm of LEAA and were credited to LEAA. Even though the information was objectively produced, LEAA was a political grant-making agency, which does not give the same impression of objectivity as a statistical agency.

- The justice statistical program is an infant compared to other federal statistical agencies. In the early years, it did not have much data available to support the policy process. After thirteen years of development, it now has extensive data bases that can be tapped for information to inform policymakers.

Providing statistical information to the Department of Justice in response to a request most often takes the form of a few numbers for a speech or for congressional testimony or to advise the Department in the use of BJS numbers or numbers from other sources, usually operational data bases maintained by the various components of the Department. Two major efforts are worth noting, however.

- The first of these was the development and approval of what would become the Organized Crime and Drug Enforcement Regional Task Force program. That program, put in place in 1983, identifies and targets for investigation and prosecution the people who are directing the drug traffic, people who in the past have eluded apprehension by insulating themselves within their organizations. BJS involvement began with designing the statistical graphics used to brief the Attorney General and the President on the proposed program. Once the program was approved, BJS helped to develop descriptive information about each of the regional areas to be used by Departmental officials visiting the areas. BJS advised the Department on the development of evaluation criteria and the implementation of operational data-collection systems to obtain the data needed for monitoring and evaluation.

- A second important policy initiative that made extensive use of BJS data and expertise was development and announcement of the Department's policy on prison overcrowding. Quarterly BJS reports on prison populations documented the crisis nature of this problem. This, coupled with policy changes that are increasing prison populations and the expense and time factor involved in building new prisons, led the Department to consider and adopt other strategies to reduce prison overcrowding. The Department's policy statement on this issue, delivered by the Attorney General at Vanderbilt University on March 3, 1983, made extensive use of BJS statistics.

Issues in data accessibility

Our experience in providing data and statistical support for the policy process over the years has revealed a number of issues that statistical agencies must face in attempting to increase the value of their data for policy development.

Technical vs. lay language. Sprehe's observation that "...to policymakers data access may mean receiving statistical information in a form that is readily comprehensible to the lay-person and adapted to policy needs" is well founded. In many areas of government, key policymakers must be considered to be a part of the nontechnical lay audience. This is especially true in the legal agencies, where the policymakers are most likely to be lawyers with extensive legal backgrounds but with little experience in interpreting or using statistical information. In arenas such as this, it is particularly important that statistical findings be made available in nontechnical language and in a policy-relevant format.

But this presents a dilemma for many statistical agencies. It is difficult for many data analysts to write in language that is technically correct, yet understandable to a lay audience. By training, analysts have been taught to write in very precise language—complete with all the necessary caveats and the identification of applicable statistical measures. To the uninitiated, a report presenting statistical findings can be unfathomable and therefore will communicate no useful information. If the data are to be used at all, the analyst must determine the most appropriate conclusions and present them in plain language free of statistical jargon.

Not only is it difficult for a statistician or researcher by training to adopt such a style, many analysts resist doing so because they believe that the statistical jargon and writing style establishes their credentials in the eyes of their peers.

Timing of data input—before or after policy determination. In an ideal world, policymakers would request and use statistical information before making policy determinations. In reality, most policymakers come to policy-setting positions with preconceived ideas as to what policy decisions they will make. Their policy positions are based on personal experience, political values, and previously acquired information. This underscores the need for statistical agencies with information relevant to policy decisions to adopt a proactive analysis and data dissemination strategy that informs individuals before they assume policymaking positions.

The search for confirming data. Many times, policymakers will make policy decisions and then search for data to support those decisions. This presents a problem for agencies that have information that does not support the policy that has been set. In those situations, the statistical agency has no alternative but to inform the policymakers of the data that are available and to point out that the policy decision cannot be defended using available statistical information. In many statistical agencies, this must be handled very diplomatically because those in policymaking positions are frequently the superiors of those from whom they are requesting statistical information.

Data manipulation to produce misleading statements. I know of no instance in which a data producer has been asked to manipulate data to support a policy decision when the correct interpretation of the data would be contrary to the decision that had been reached, or to withhold from public dissemination such data. That is not to say that such requests have not been made; nor does it imply that statisticians need not concern themselves with such a possibility. Professional ethics, of course, dictate the appropriate behavior in such situations.

Conclusion

To make data more useful for policymaking requires that the data producers engage in a concerted effort to present their statistical information in a readily understandable, nontechnical, policy-relevant form and to be available to assist the policymakers in interpreting and using the data. But such a participatory role for statistical agencies raises a number of issues that go to the core of maintaining an objective, unbiased statistical program. In the past, statistical agencies have insulated themselves from the policy process and the potential it poses for politicizing their programs by maintaining the stance of "fact finders," leaving it to others to figure out how to use the data and to decide what the data mean. Modern federal statistical policy expects data producers to assist policymakers and other data users.

Those of us who have passed through the revolution feel a certain nostalgia for the days when our professional life was simpler; but there is no going back. The past few decades have taught us that there is a tremendous demand for statistical information for use in the policy process and that it is we, the data producers, who are best qualified to interpret the data to ensure that the information is used correctly.

Issues in Using Statistics in the Policy Process

Marion R. Metcalf, *Criminal Division*
U.S. Department of Justice

Abstract

This paper discusses problems in selecting measures of the criminal justice system's performance. The primary illustration is the felony attrition rate, which is based on the percentage of arrests that do not result in convictions. Because this rate gives equal weight to all case outcomes other than conviction, it is of limited help to policy makers whose concern is whether significant cases go unprosecuted for avoidable reasons. The paper recommends that criminal justice researchers collect and present descriptive data when examining case processing.

When I first became interested in the policy analysis profession, I thought it was very important to study high-powered statistics. My concern was that researchers and statisticians could wield a lot of power over the policymaker who didn't understand their methods. I wanted to be able to analyze those "mystical" statistical reports independently and reach my own conclusions. So I studied regression, cross-sectional time series, and analysis of variance; and sometimes now I even use that knowledge in my work as a policy analyst in the Criminal Division.

But my experience at Justice has taught me that some of the biggest policy debates arise from research that presents nothing more elaborate than percentages. Part of the problem is that even percentages can be confusing to attorneys who are not accustomed to thinking about them. I remember once trying to demonstrate that there was something fishy about a table in a report. The table gave descriptive statistics on 100 cases. But all the percentages were carried out to the first decimal place. I showed the chart to a couple of lawyers, assuming the problem was self-evident: how could you have 19.6 percent of 100 cases? But the response I usually got was an uncomfortable, blank look and the reply, "I don't know anything about statistics."

But the biggest problem is that researchers often don't think carefully enough about which percentages to calculate and present. Even attorneys who are not math whizzes have taught me a great deal about what distinctions need to be made in analyzing various problems, and those lessons have great implications for the appropriate base to use in computing percentages.

To illustrate some of those lessons, I would like to talk about analyses of case processing, using examples from a variety of sources. Since I have started out on a critical note, I should make it clear that I have drawn some of my examples from a study I think is very good: What Happens After Arrest? by Inslaw (1), which examined cases processed in Washington, D.C. in 1974. My problem is not with that report but with some of the policy implications others have proposed either by using that report as a basis or without refining its analytical framework.

Analyses of Crime and Punishment

Let me begin by looking at the big picture. There is a general perception, which is supported by the numbers, that we have an awful lot of crime in the U.S., but, as all that crime gets processed by the justice system, we wind up punishing a disproportionately small number of offenders. A recent illustration of this view was depicted in a graphic picture of a leaky funnel in U.S. News and World Report, (2) which started out with 500 serious reported crimes and wound up with only 20 adults and 5 juveniles being sent to prison or jail.

(There are also critics who think we punish too many people. My own criticism of the "leaky funnel" view is that, if the 25 convicted persons are high-rate offenders, the ratio of one in twenty may be perfectly appropriate. But I will leave those arguments aside.)

But let me return to all that serious crime. Most of the time, we talk about the crime we can measure: crime that occurs as a discrete, observable, reportable event. Statisticians have devoted a lot of energy in recent years to measuring unreported crime, so now we have both victimization surveys and reported crime statistics on which to base estimates of the crime rate.

Those crime rates, however, miss what I think of as the "unmeasurable" crimes. For example, crimes such as public corruption, organized crime, and major narcotics trafficking are hard to detect, let alone measure. That point has important implications for the kinds of cases that make it into both the crime statistics and the case flow process. That is one reason it can

be dangerous to try to tell how "well" the criminal justice system is doing just by looking at percentages drawn from its caseload. Having identified the perception that there is "all that crime" and that the law enforcement system does an inadequate job of catching and punishing the perpetrators, analysts try to figure out why. They often focus on various segments of case processing, compute some percentages, and draw some conclusions about how we might improve the enforcement system. Too often, however, their choice of percentages does not really illustrate the problem they care about. In selecting percentages, we need to think about the following questions:

1. What are we trying to explain about the justice system?
2. What factors determine which cases make it into the base for analysis?
3. What resource factors do we need to consider?
4. Which actors may be held accountable for various outcomes?
5. Should all cases in the base be weighted equally?

An Illustration: GAO Declination Study

My first example is from a GAO report released in 1978. GAO examined cases that had been referred to prosecutors and found that 62 percent had been declined. GAO said:

> Many of these complaints were declined because of legal deficiencies, such as lack of evidence, or inability to prove intent. However, many of the declined complaints could have been prosecuted but were declined because the U.S. Attorneys believed that . . . the cases did not warrant the cost of prosecution and/or staff was not available to handle heavy workloads. (3)

GAO analyzed a sample of the declined cases and concluded that 22 percent had been "prosecutable." The report went on to make a number of points about declination guidelines, constrained U.S. Attorney resources, and the handling of minor violations. Rather than debate those points, I want to focus on the choice of percentages.

The problem GAO wanted to focus on was the possibility that prosecutable cases may be going unprosecuted.

Unfortunately, what stuck in people's minds were (1) GAO's concerns about declination policies and (2) the 62 percent declination figure, which happens to be the least interesting and useful number in the report. If GAO's concern was about the declination of prosecutable cases, its analysts should have focused on that group of declinations; instead, they cluttered up the issue by devoting so much attention to the 62 percent declination rate. In doing so, GAO portrayed all declinations as failures of the justice system. Instead, GAO might have said that, of all cases referred to federal prosecutors,

- 38 percent were accepted for prosecution;
- 48 percent were declined because they were not prosecutable cases; and
- 14 percent were declined even though they were prosecutable cases.

Or, Of "prosecutable" cases referred to federal prosecutors,

- 73 percent were accepted; and
- 27 percent were declined.

In response to the critique I have just made, researchers defend themselves by saying, "but we mentioned that all the referrals are not prosecutable." That strikes me as a lame excuse when the confusion could have been avoided in the first place. GAO's choice of percentages did not reinforce the distinction between "prosecutable" and "unprosecutable" cases; in fact, the choice blurred that distinction.

The failure of researchers to make those distinctions explicitly is a very sore point with many prosecutors. Prosecutors see it as part of their job to screen out cases that should not be prosecuted. They may disagree on the question of whether GAO should have said that "prosecutable" declinations were 14 percent of all referrals or 23 percent of all "prosecutable" referrals; but they resent being criticized for a 62 percent declination rate that lumps together legally sufficient and legally insufficient cases.

Research on Felony Attrition

Which brings me to the "felony attrition" concept. I use the term "felony attrition" to refer to research that focuses on the fraction of felony arrests (or referrals to a prosecutor) that do not result in felony convictions. Although, again, these studies generally mention that not all attrition is "avoidable," those qualifications are almost never explicitly incorporated into the analytical methods employed. The implication is that any case outcome other than a felony conviction represents a failure of the justice system. Depending on the writer's perspective, that failure may be blamed primarily on poor case preparation by police, prosecutors' guidelines, or inadequate communication between investigators and prosecutors.

Felony attrition studies can be useful as explanations of how the justice system

processes cases. But care must be taken in drawing board policy implications from them. For instance, some observers have proposed arrest-based conviction rates as appropriate evaluative measures for police case preparation, prosecutors' offices, or attempts to improve law enforcement coordination. I believe that felony attrition rates are misleading measures of all three.

The key question is not how many cases (or what fraction) do not result in conviction, but whether significant cases are going unprosecuted for avoidable reasons. The felony attrition literature has been a useful first step in understanding how cases are handled. The next step should be the development of measures that focus precisely on case significance and on aspects of case attrition that are avoidable, that should be avoided, and that are amenable to correction through improved police and/or prosecutor performance and coordination.

For the rest of this paper, I will focus on five problems with the arrest- or referral-based conviction rate:

1. It diverts attention from system output issues;
2. It ignores what happens prior to arrest or referral;
3. It overlooks resource issues;
4. It assigns equal responsibility to all actors in the system; and
5. It assigns equal weight to all cases and to all reasons for which cases may be dropped.

1. It diverts attention from system output issues. It is probably true that it would be desirable to produce more convictions. But it does not follow that the best way to measure progress toward that goal is to calculate the ratio of convictions to arrests. One way to increase the arrest-based conviction rate might be to make fewer arrests. However, it is possible that in order to increase the number of persons convicted and punished, it would be necessary to make more arrests and tolerate a lower conviction rate.

2. It ignores what happens prior to arrest or referral. Some analysts have inferred from felony attrition studies that we should encourage police to make only "convictable" arrests or investigators to refer only "prosecutable" cases. There are three problems with this argument: 1) limitations on police discretion in making arrests, 2) variations in screening practices, and 3) differences among cases.

- Limitations on discretion in making arrests. There is a notion that police have a lot of discretion in making arrests and that they can do more to produce convictable cases. I do not completely disagree with that. But in Inslaw's study of arrests in Washington, D.C., almost half of the arrests were made either during or immediately following the commission of a crime. (4) A lot of those arrests seem to have been what we might call public order maintenance arrests (e.g., street drug dealing, public drinking, or prostitution). To a lesser but important extent, the arrests were for "street crimes in progress." We certainly would not want police to assess the likelihood of conviction before making arrests during a hold-up in progress. In devising statistics to assess police performance, those arrests may need to be treated differently from arrests made after the officer has had time to build a case prior to making the arrest.

- Variations in screening practices. There appears to be a great deal of variation in the extent to which agencies attempt to screen out "unprosecutable" cases before referring them to the prosecutor. These variations are especially important in making comparisons across agencies or jurisdictions. For instance, the federal investigative agencies have different screening policies: the FBI's policy is that investigations may ordinarily be closed only by the U.S. Attorney (U.S.A.), unless the U.S.A. has developed a written declination guideline. Without a guideline, even if the agent is pretty sure the case is not convictable, he is supposed to ship it over to the U.S. Attorney to make the closure official. Other agencies, such as the Postal Inspection Service, give agents more discretion to close cases on their own. Therefore, comparisons of referral-based conviction rates across agencies can be very misleading.

- Variations in Cases. The third factor to consider in what happens before arrest is the type of case. Let me illustrate the point by comparing a commercial burglary and a street-level drug case. In the burglary, a store owner comes into the store on a Monday morning, sees some of his stock missing, and calls the police. He may have a suspect in mind. So the case is on the books as a "known suspect burglary." If an arrest is made, the prosecutor has to link the perpetrator to an event that, remember, no one actually saw take place. In a street-level drug case, the detectives often know what it takes to "make" the case. For example, they may have to witness three drug buys. That means that if they make an arrest, it is more of a "sure thing." The drug case would not go on the books until the arrest is made; but the prosecutor's job may be easier for the drug case than for the burglary, and conviction may be more likely.

So, to summarize, there are important differences in

- the strength of cases at the point of arrest,
- the extent to which cases are screened before referral to a prosecutor, and
- the degree of discretion in whether to make an arrest.

Those differences can be obscured by arrest-based conviction rates.

3. It overlooks resource issues. Some observers believe that where prosecutorial or court resources are constrained, those constraints, rather than the number of convictions the system will produce. Similarly, there is a growing body of evidence that limited punishment resources are the most critical constraint facing the criminal justice system at all levels of government.

4. It assigns equal responsibility to all actors in the system. The felony attrition approach glosses over the issue of which actors may appropriately be held accountable for unfavorable case outcomes. Police, for instance, are held equally responsible for cases that are dropped because of inadequate police investigation and cases lost because the prosecutor erred at trial. It is important to note here that the police should not be expected to assess the sufficiency of the evidence in all cases; it is generally best for prosecutive judgments to be made by a prosecutor. However, the felony attrition rate penalizes the prosecutor for declining a case based on a correct judgment that the evidence is insufficient. In addition, it does not distinguish an appropriate declination from a case lost at trial because the prosecutor failed to decline it on evidentiary grounds.

This problem may not be crucial in studies that explain how cases are processed. But if conviction rates are to be used as the basis for policy decisions or performance measures, distinctions have to be made. For example, in the sample analyzed in the Inslaw study, more cases "dropped out" because the defendant had completed a diversion program than were declined for evidentiary reasons. (5)

The proper treatment of diversion programs in analyzing case flow is problematic. "Diverted" cases do not represent a conviction, but, assuming the program is run responsibly, they are not a failure of the justice system either. In the studies I have seen so far, treatment of "diversion" as a "non-conviction" has not created a serious bias. (6) If use of the program becomes more widespread, however, its proper treatment in computing performance measures could become a serious methodological issue.

5. It assigns equal weight to all reasons cases may be dropped and to all cases. Despite all the things I have said so far, I do think it is important to keep working to measure and improve the way police and prosecutors do their jobs. But the fact that some cases do not get prosecuted or convicted does not tell us much about how well our system works. As a starting point, I would like to see researchers start to be explicit in distinguishing among these "non-convictions":

1. Cases the prosecutor did not wish to pursue as felonies (e.g., because the case was trivial or a first-time offender had completed a diversion program);
2. Cases the prosecutor would have pursued but could not for evidentiary reasons or witness-related problems (this group then deserves a second look to try to distinguish cases in which the problem might have been fixed from those in which nothing more could be done);
3. Cases the prosecutor would have pursued but couldn't because of resource constraints; and
4. Acquittals.

But even improving our understanding of what happens after cases get into the system does not necessarily tell us whether the "right" cases are making it into the system. In the GAO "62 percent declination" study, a lot of those cases got counted because someone could see that the crime had happened: there was an auto theft or a theft from an interstate shipment.

Unfortunately, not all of the important crime is that easy to count. For example, at the time of GAO's study, the governor of my home state of Tennessee was selling pardons to convicts. He finally got caught, but his crime was much tougher to detect than a typical auto theft. His crime was not reportable in the usual sense of the word. And even though I thought the guy was a crook, I could not have recorded my perception if I had been surveyed for a victimization study.

One of my biggest problems with felony attrition studies is that they give an auto theft case exactly the same weight as one involving a corrupt governor. Yet for me, distinguishing cases by their significance is central to the policy question of whether the justice system is doing what it is supposed to. I wish I had detailed and concrete suggestions for how researchers could help us make these judgments, but I don't. My office has looked at the idea of case-weighting systems, which try to make case-quality criteria explicit. But those systems are controversial even within my office, never mind with the prosecutors we work with.

I do have one suggestion, though: collect and present as much descriptive information as you can on the cases you analyze -- data on such things as: Who are the victims? Who are the perpetrators? How much money was involved? If it was a violent crime, was a weapon involved? If it was a program fraud, was the offender a provider or a recipient? The more of those details we have, the more we can start to at least tell how the justice system handles major and minor cases, and maybe start to figure out if we are doing the right things.

I said at the outset that I had learned that percentages are often more important to the policy debate than fancier statistics. I think that politicians, the press, and people in general gravitate toward percentages because they think they understand them. Unfortunately, the numbers do not always mean what they seem to. I think we have just as much responsibility to exercise care in citing percentages as we do in using our more "mystical" statistical methods.

(1) Brian Forst, et. al., What Happens After Arrest?, Publication no. 4, PROMIS Research Project, Washington, D.C., Institute for Law and Social Research, 1977.
(2) U.S. News & World Report, Nov. 1, 1982.
(3) General Accounting Office, U.S. Attorneys Do Not Prosecute Many Suspected Violators of Federal Laws, page i, GGD-77-86.
(4) Forst, et al., op. cit., p. 33.
(5) An estimated 1430 cases in the sample were dropped after initial acceptance because the defendant completed a diversion program; approximately 1290 cases were declined either initially or subsequently for evidentiary reasons. Calculated from pages 67 and 69, Forst, et al., op. cit.
(6) Diverted cases constituted 8 percent of the arrests in What Happens After Arrest?.

Structure and Process of Collecting and Analyzing Justice Statistics in Canada

Dr. Gaylen A. Duncan, *Canadian Centre for Justice Statistics*

The following quote from the report that recommended the creation of the Canadian Centre for Justice Statistics gives a true picture of the frustration that was facing managers and researchers in Canada:

"We can trace a hog from farm to market, but we cannot trace an individual through the justice system."

I could add: "We can track the family lineage, and all relevant, and sometimes irrelevant, information on the life and times of a race horse at no cost to the state. As a manager in the public service, the task I face is the creation of basic information on people in the system and on the operation of the system."

The real challenge facing the Centre is to create similar systems for the justice system without imposing unreasonable costs on governments.

Introduction. The theme of this paper can be summarized as follows: Explain the Centre in terms of its history, its rationale for existence, its recent experiences and its likely future. To do this, first, the paper describes some fundamental differences and similarities in Canadian and U.S. constitutional structures, and then places Canadian experience into context based on what was done before 1981, what is being done now and what will be done in the future. The paper also adds a few comments on the Canadian Uniform Crime Reporting Survey.

In many ways this paper is a series of good news/bad new stories. Using Canadian and U.S. examples, the good news/bad news theme is not being expressed in terms of Canadian/U.S. development but in terms of good news and bad new about justice information in Canada.

I. Canadian Constitutional Reality. Treatises have been written that compare the Canadian and American constitutions. For the purposes of this paper, it is only necessary to understand and appreciate the fact that the split between American federal and state jurisdictions regarding the administration of the justice system has not been followed in Canada. The split between the federal and provincial jurisdictions differs in most, if not all, respects from the U.S.

For example, in Canada the federal government appoints all judges for the Supreme Court of Canada, Court of Appeal for each province, superior trial courts, etc.... On the other hand, the province has the sole authority to employ the judge's secretary, to purchase office furniture, to build court rooms, and to hire additional support staff. Resources to pay for travel by support staff are controlled by the provinces. Resources for travel by federally appointed judges are controlled by the federal government.

For example, Canada has a single, unified criminal code. Thus it is not a difficult task to develop a standardized classification system for the Uniform Crime Reporting system or for criminal caseload statistics. On the other hand, the administration of the criminal justice system falls under provincial jurisdiction. Because of this, it is extremely difficult to develop comparative workload measures for courts and court administration. Similar difficulties exist for comparative prosecutorial workload measures.

For example, jurisdiction over correctional insitutions is split between the federal level and the provincial level based solely upon the length of the original sentence. If the sentence is for two or more years, the offender goes to a federal institution. This particular split could be altered by amending the statute which sets the two year limit.

For example, Canada has a unified court structure. Judges may be appointed by the federal government (superior courts) or by the provincial government (lower courts), but all appeals go through the same structure, eventually landing before the Supreme Court of Canada. It should be noted that the one major exception to this is the Federal Court of Canada which has limited jurisdiction to hear certain cases (e.g. actions against the federal government or appeals from federal tribunals).

In summary, the key points to remember are:

Federal Jurisdiction

- 1 Criminal Code
- federal police for federal statutes
- appoint and pay superior judges
- limited prosecutorial role (federal acts)
- 2 year plus in correctional institutions
- cost share few areas
- legal aid
- new Young Offenders Act
- develop and articulate national policy

Provincial Jurisdiction

- administration of the justice system
- contract federal police to perform local police functions or control provincial police forces or local police forces
- appoint and pay lower court judges, but run all courts
- 2 years and less in correctional area
- implement and pay for national policy

This dichotomy in jurisdictional authority is also reflected in the structure and operations of the Canadian Centre for Justice Statistics. The Centre is part of Statistics Canada, a federal agency, but it is also controlled by a Justice Information Council, a form of board of directors, with representatives from both levels of government. This is not surprising since the provincial jurisdictions are the source of most of the necessary data.

II. Creation of the Centre. President J.F. Kennedy once said - "Washington is a city of southern efficiency and northern charm."

The Centre is an agency designed by a committee of lawyers, premised on the split Canadian jurisdiction in justice, and directed by consensus decisions of all participants.

The Evaluator of the Centre, in a recent report, said - "Seen objectively, this is not the most efficient way to collect national justice statistics."

This may not be the most efficient way to collect national justice statistics, but it is proving to be the most effective way. Over the past twenty-five years, all previous attempts have failed. After two years of considering alternative approaches, a conscious decision was taken by all parties to try this route.

The key point to remember is that the Centre was created in June, 1981 to run for 3 years, to be evaluated by an external evaluator and then to have a committee of 22 key people, the board of directors, consider its future. If nothing is done, the Centre will fold in April, 1984.

III. What Was it Like Before 1981. The following information was being produced in the years preceding 1981 (the creation of the Centre):

- police administration statistics showing the number of police and civilian personnel;
- uniform crime statistics;
- homicide information;
- nothing on legal aid programs;
- civil court statistics - 1 study was coordinated in 1973;
- criminal court statistics covering part of two provinces;
- correctional statistics covering institutions only and covering only 4 of the smaller provinces.

The reports were plagued with problems of timeliness, accuracy, quality, coverage and gaps.

The system was based on heavy clerical resources and centralized processing.

IV. What We Have Today. The following summarizes the information now available and the experimental projects now underway:

- police administration statistics as before but twice as timely;
- uniform crime statistics as before, but more timely and with more graphical presentation;
- homicide information as before but more timely;
- legal aid programs were described in two recent reports and a new national survey has been proposed;
- Civil Courts information has been produced on administrative tribunals at the federal level and for five provinces, descriptive material for all jurisdictions on family law has been produced and a survey is being considered, descriptive material on civil court processes is being prepared and an experiment has been started using a legal information retrieval system to see if it can form a basis for civil court information;
- Criminal Courts information on manpower, resources and costs, a national directory of courts has been produced for the first time, all but one jurisdiction has started or agreed to start reporting basic caseload statistics for the current year, a proposal for detailed caseload and offender reporting has been circulated to jurisdictional information system planners, a court workload study has been started in three jurisdictions and a major test is underway to see if court sentencing and offender information can be culled from the existing fingerprint system;
- prosecutorial information on manpower resources and costs was recently released;
- correctional information covering both institutions and correctional services for all jurisdictions is now produced annually;
- juvenile court data for all jurisdictions using standardized reporting is now underway.

Based upon the achievements outlined above, achievements which have been attained after only 24 months of operation, it is fair to reiterate both the conclusion of the Centre's

Evaluator and my own reaction. Seen objectively, this is not the most efficient way to collect national justice statistics. But it appears to be the most effective way.

V. The Structure of the Centre. It is difficult to summarize the geographic and demographic facts about Canada. Given our size, our location, and our population, we are fundamentally a nation that should not exist. Our population is 50% more than the population of the State of Texas - a state that prides itself on its wide open spaces - yet our land mass is 15 times that of Texas. Canada is divided into a federal government, 10 provincial governments and two territorial governments in the north.

Within each jurisdiction, there may be as few as 1 ministry responsible for the administration of justice or as many as four (Corrections, Attorney General, Solicitor General, Justice Policy Secretariat, all in Ontario) with yet another ministry dealing with juveniles.

Canada's statistical system is heavily premised on the centralized functions of Statistics Canada, of which the Centre forms a part.

The Centre is controlled by a board of directors called the Justice Information Council (JIC), consisting of 22 deputy ministers responsible for the administration of justice in all the jurisdictions plus the Chief Statistician of Canada. This board approves all projects, approves annual work plans, sets new directions and priorities, and receives the evaluation of the Centre. The board also has an executive committee.

Each representative on the JIC has appointed one delegate to the Liaison Officers Committee which acts, in effect, like a management committee. The LOC receives all detailed plans for projects, identifies policy issues for submission to the JIC and ensures that all processes function smoothly.

There is a plethora of other committees that assist, monitor, or direct the Centre. They are collectively referred to as the support structure for the entire initiative.

Each program at the Centre is advised and supported by a Program Development Committee (PDC) which is made up of experts in a given field. These committees are not based on jurisdictional representation but rather they consist of selected experts in a given field. Recommendations go from their PDC's to the LOC and directions for further work go from LOC to PDC's.

The Centre itself is headed by an Executive Director, and has two main program groups (the Technical Assistance Directorate (TAD) and the Statistics and Information Directorate (SID)and one staff group (Policy, Planning and Evaluation)). The SID group has programs in Law Enforcement, Courts (Criminal, Non-Criminal) Legal Aid, Corrections, Juvenile Justice and Integration and Analysis. The TAD group performs a similar function to what was originally conceived for L.E.A.A. in the U.S. This group provides funds and technical resources to support the development of operational information systems to feed data to collection efforts that support national priorities.

VI. Centre Priorities: 1981 and Today. The documents that led to the creation of the Centre identified a mandate that was premised first on "informing the public" and secondly on "supporting national justice policy development". In 1981, the main policy issues facing the Centre were matters of "micro data projects" versus "aggregate data" and the degree of analysis that should be done by the Centre versus the policy groups in the various jurisdictions.

After two years of effort (the major results of which I have already outlined), the JIC has authorized the Centre to look into the future and to prepare a plan of action for the 3 to 5 years that will follow our original 3-year lifespan. For the past few months, we have finally had an opportunity to ask "what is needed and what should be done" rather than to focus on "how do we do what was originally asked of us". As a result, we are beginning to put our current projects into a new perspective. Our ideas are now being reviewed by the jurisdictions so everything that I say from now on should be viewed as my best bet on what might be approved before April 1984.

The priority for 1984 and onwards will not be "to inform the public" but will be to assist in the development and use of management information, to deal with on-going issues as a result of budget restraints, cost-sharing discussions and current initiatives to increase the role of evaluation projects and long-term planning and budgetting in the justice system.

Simply put, our focus will be to serve the JIC. To do this we will take each sector, describe it, cost it, measure caseflow through it and develop measures to forecast demand on that sector. This last item is key, though we do not fully understand the program and project implications of it.

All of this work will be premised on the need to facilitate the production of activity indicators.

The second prime thrust will be the provision of support to develop systems and provide training on the use of information from such systems. The third and last thrust will be the provision of information as a service to the JIC.

Before I move on to discuss Canada's U.C.R. system, let me tie together one key point that relates to the structure of the Centre and to our likely future directions. The major potential benefit of the Centre vis-a-vis all previous initiatives is that we have clearly identified our client. It is the JIC . Further, the JIC represents the vast majority of our data providers and the JIC is also our major information user. This user/producer committee is also our board of directors. This is a simplisitic view of reality, but from this simplistic model we derive our greatest likelihood of success. And if we fail, it is this very group that must pass judgement on us - even if the final conclusion is that we became irrelevant because the JIC failed to provide adequate direction.

VII. Canada's U.C.R. System. The similarity between the U.S. and Canadian U.C.R. systems begins with the fact that they are both descendants of the efforts of the International Association of Chiefs of Police. Thereafter, in almost every way, the similarities end.

Remember that Canada has 1 federal criminal code and that this code applies throughout the country. This quirk of our constitutional structure has allowed us to develop a single system, covering over 99% of the relevant forces and over 99% of our population. We collect data on over 100 crimes with about 10 variables for each. This wealth of information is the good news. The other side of the coin is that we too face major criticisms from policy makers, managers and academics on quality, scope, coverage, timeliness, detail, etc. We too have started a major review of the role and need for national crime data and other police information.

I expect that we will evolve to a point where we have 1 survey to monitor basic crime trends, and a second sytem, based on existing and future automated operational systems in the major police forces, to provide access to detailed data for in-depth studies. I say this even though we do not yet know the final results or reactions to Canada's first full blown victimization survey which is due for release shortly.

VIII. Summary. The Centre is a three-year experiment that is already two years old. It has a staff of about 80 that is expected to do the work originally set out at our creation while simultaneously planning for the future and maintaining a heavy commitment to the consultation process. We are working on short-term projects (collect and analyse information relevant to today's problems) and long-term solutions (build local systems and develop local expertize). To do this, we are learning from past mistakes in Canada and benefitting from current efforts in the U.S.

IX. Revisiting Some of the Messages. Having already summarized my main points, I should finish with a major vote of thanks to various U.S. organizations, many of which are represented here today. Recently some of the Centre staff, plus representatives of two provinces, visited the Bureau of Justice Statistics, the National Center for State Courts, the Federal Administrator of Courts, the National Institute for Justice and the Federal Judicial Centre.

They were royally treated and we received reams of documentation and many useful ideas. I trust that this is just the beginning of a series of such meetings.

Now, having concluded the main part of the paper, I want to take a few minutes to make a few personal comments on where I think the Centre is going. I hope these points will help to draw together the various themes I have already presented.

First, the Centre will focus on each sector of the justice system by first describing it, then costing it, then quantifying caseloads and finally identifying measures of demand. In my opinion, the idea of measures of demand is a critical component of our future, even though we still have much to learn about this concept.

We are not trying to build a model of the justice system. This might be the long-term future of our efforts, but, in my opinion it is the wrong way to start. We have realized that the output of one sector of the system is the input to the rest - no great breakthrough. Having realized this, we now recognize that the pressure for good police statistics does not just come from the police, but also from prosecutors who need this information to predict prosecutorial demand. Similarly, court data is needed to plan correctional services. We are beginning to realize that the need for demand data in one sector can be used to drive efforts in the previous sectors. I predict that this framework approach will turn out to be as important in the future as our structure of a JIC (representing users, producers, controllers and evaluators) is critical to our success today.

Secondly, success in getting detailed data on special issues and conducting offender tracking studies can only be achieved by putting operational systems into jurisdictions.

Third, based upon these systems, the Centre must be flexible enough to respond to JIC requests for information that are based on current and evolving needs.

These three initiatives must be successfully achieved in an environment where users and producers totally control us, where we are

expected to take statistical information from operational systems and not to impose new, unreasonable costs on jurisdictions, where we are required to minimize the number of new surveys and at the same time where we are to dramatically increase the quantity and quality of information on the justice system, and finally that we are to live under a permanent sunset clause.

That is the future I see for the Canadian Centre for Justice Statistics, and in my opinion it is a bright future.

The Organization of the Ministry of the Solicitor General of Canada and the Role of the Statistics Policy Advisor

Penny Reedie, *Ministry of the Solicitor General, Canada*

The Ministry of the Solicitor General was created in 1966, bringing the major operational elements of the federal government concerned with the administration of the criminal justice system under the direction and supervision of the Solicitor General of Canada. These major elements are the three operational agencies, the Royal Canadian Mounted Police (RCMP), the Correctional Service of Canada (CSC), and the National Parole Board (NPB), and the Ministry Secretariat.

The RCMP is the federal law-enforcement agency, which also provides policing services under contract to eight of the ten provinces, 190 municipalities, and two territories. CSC is the federal penitentiary agency which is responsible for administering sentences of two years or more. The NPB is responsible for the granting of conditional release. The Ministry Secretariat supports the Deputy Solicitor General in his role as chief policy advisor to the Minister. Thus the Secretariat provides policy advice, co-ordinates Ministry programs, and provides certain centralized common services in areas such as research and statistics. Organizationally, the Secretariat has three branches: Policy, Police and Security, and Programs.

The Statistics Division is one of five divisions in the Programs Branch, which advise the Minister in the development of long-range policies on criminal justice matters. The Statistics Division provides professional and technical services to the Minister, the Secretariat executive, and the Ministry agencies, and promotes the development of better information and statistics in the Canadian criminal justice system. Divisional activities include statistical policy, statistical methodology, statistical studies, computer systems technology, and short-term requests for data and advice.

Because the Ministry provides direct services, much of the work of the Statistics Division is in generating data bases and producing statistical information which anticipate future policy requests and which help to contextualize policy objectives and specify policy questions. Some recent projects are Conditional Release, Mandatory Supervision, Recidivism, Dangerous Offenders, Natives, and the Effects of Long Term Incarceration. To promote the development of better information and statistics, the Division is involved in operations research and the development of computer-systems technology. A major responsibility of the Division is identifying the limitations of data and interpreting statistics produced by the ministry agencies and by external organizations such as the Canadian Centre for Justice Statistics.

The Canadian Urban Victimization Study is a large, high profile project which exemplifies the dual role of the Statistics Division. The Victimization Survey addressed five areas:

1. the extent of reported and unreported crime during 1981,

2. the risk of criminal victimization,

3. the impact of crime,

4. public perceptions of crime and the criminal justice system and victims' perceptions of their experiences, and

5. the needs of victims.

The methodology and questionnnaire were designed and pre-tested by "technicians" in the Division in consultation with Statistics Canada. "Translators" ensured that pertinent issues identified in discussions with policymakers were included in the design (e.g. crime prevention, victims, police resources) and later will assure that the findings are presented to policymakers in a way that is useful and used.

The survey was conducted in early 1982 in seven major urban centres, Greater Vancouver, Edmonton, Winnipeg, Toronto, Montreal, Halifax-Dartmouth, and St. John's. From the results of the survey we have begun to produce

1. Bulletins on topics of general interest -- elderly victims, victims of violence by intimates and violence by strangers, break and entry offences, motor vehicle theft, vandalism, etc.;

2. User Reports to respond to specific information requests;

3. Reports that deal more extensively with such topics as the measurement of crime and victimization, methodology, etc.; and

4. Special reports directed to pertinent issues identified by policymakers.

The Victimization Survey provides not only a wealth of statistical information but also a rich data base from which we can integrate and answer a range of policy and research questions. The data base also allows for a quick turnaround time in answering questions and specific information requests.

The Victimization Survey is illustrative of the role of the Statistics Division as distinct from the Canadian Centre for Justice Statistics, which has both federal and provincial responsibilities. The work of the Centre cannot be and should not be oriented specifically to the policy concerns of the federal Ministry of the Solicitor General, whereas the work of the Statistics Division is and must be. It is this area of policy and the activities of statistics policy advisors, i.e., translators, that I would like to concentrate on here.

The Statistics Policy Advisor is a unique merging of two areas: statistics and policy. The Statistics Policy Advisors act as liaison persons between the Ministry and other government departments, among users of criminal justice statistics within the Ministry, and between the Ministry and the Canadian Centre for Justice Statistics. It has been our experience that a statistics policy advisor/-translator/liaison person can provide useful services in every policy-user area. The Statistics Policy Advisors also are the link between, or the translators for, the policymakers and the statisticians. The Statistics Policy Advisors must know the limitations and the variety of interpretations that can be drawn from different data produced internally or obtained from external sources. The Advisors inform the policymakers on the interpretation and limitations of particular criminal justice statistics and communicate to the statisticians and data collectors new initiatives being considered by policymakers.

The Statistics Advisors frequently act as mediators, communicating to statisticians the concerns of policymakers regarding such things as timeliness, quality, and utility of data, and to policymakers the concerns of statisticians about such things as quality control. Most important, perhaps, the Advisors help policymakers define their information needs within the context of what is and can be available.

The Statistics Policy Advisors must maintain dialogue with the statisticians in order to keep informed of limitations of particular data sets and consequently the interpretations of such data and to keep policymakers advised of the meaning of criminal justice statistics. As a result of a continuing dialogue and of the recognition of the concerns of both statisticians and policymakers, the victimization survey emerged. Although such surveys have their limitations, the victimization survey nevertheless provided a very large volume of information about the extent of reported and unreported crime during 1981, the risk of criminal victimization, the impact of crime, public perceptions of crime and the criminal justice system, and victims' perceptions of their experiences.

It has been our experience that the two greatest dangers facing technicians are

1. the tendency toward reification of the data they are collecting, and

2. sacrificing timeliness for methodological and technical considerations.

The two biggest problems that policymakers have are

1. a lack of appreciation of the value and uses of statistics, and

2. an inability to define their statistical needs.

These sets of problems are not unrelated. Indeed, if statistics were better and more timely, then policymakers would better see their worth, and if policymakers were better able to define their needs, then the statistics would be better.

The Statistics Division has organized itself around this dilemma so that the workload is roughly divided between technicians and statistics policy advisors. The technicians work in the area of operations and information systems and serve the immediate needs of the Ministry. They answer short-term requests for information and advice from Ministry officials and external parties and are also involved in long-term research and statistical studies. More important is the unique area of statistics

policy where statistics policy advisors serve as translators of statistics produced internally by the different agencies and by external bodies.

An examination of the Uniform Crime Report (UCR) system illustrates the dilemma faced by criminal justice policymakers confronted with statistics. The UCR database, the single most important source of Canadian criminal justice statistics, is generated by federal, provincial and municipal police agencies. Usually, UCR data is interpreted by the media to indicate a rising rate of crime, but many alternative analyses are possible, depending on the ends to be served. For example:

1. look at how crime has risen;

2. look at how rare violent crime is;

3. UCR shows no cause for fear and panic;

4. UCR is false -- it minimizes the amount of crime, omitting the hidden or "dark figure".

One role of the translator therefore is to caution the policymaker about facile conclusions. Another is to caution the technician against reifying his data base.

There is presently underway in Canada a large, slow and complex program to develop comprehensive and accurate national and local statistics at every level and for every element of the criminal justice system. This program involves the federal, provincial and municipal governments and all criminal justice agencies. Within the next decade, data will be available which will allow reliable statements to be made about all aspects of the Canadian criminal justice system. The Statistics Policy Advisor will then be even more important than now, in reducing the mass of data to its policy implications.

Analysis of a Y-Stratified Sample: The Georgia Charging and Sentencing Study

George G. Woodworth, *The University of Iowa*

This is a report on methods which I have used to analyze complex survey data on general purpose statistical software. The survey in question concerns charging and sentencing in homicide cases in Georgia between March 1973 and January 1980. During that period there were about 2,500 convicted offenders, 100 under death sentences, 876 with murder convictions and a life sentence, and 1,480 with voluntary manslaughter convictions.

In this discussion I shall focus on one outcome variable (whether or not an offender received a death sentence) and upon 11 independent variables listed in Table 1.

Mnemonic	Description
Independent Variables	
BLACKD	Defendant was black.
WHVICRC	Victim was white.
BLVICMOD	Family, lover, liquor or barroom quarrel.
STRANGER	Victim was a stranger.
TWOVIC	Defendant killed two or more victims.
FELMUR	Defendant was involved in a felony at the time of killing.
VPCARBR	Defendant had one or more prior convictions for violent personal crimes, burglary, or arson.
TORTURE	Victim was physically tortured.
MENTORT	Victim was mentally tortured.
NOKILL	Defendant was not the killer.
LDFB7D	Crime involved rape, armed robbery, kidnap, killing to silence a witness, execution style killing, or a victim pleaded for his/her life.
Dependent Variable	
DSENTALL	Death sentence given conviction: 1 = death 0 = life.
Case Weighting Variable	
WEIGHT	Reciprocal of probability of selecting the case.

Table 1. Selected variables from the charging and sentencing study.

The Stratified Sampling Plan

Sentencing in capital cases in Georgia is bifurcated into a guilt trial followed by a penalty trial at which additional evidence may be presented relating to the defendant's prior record and to other factors which could not be presented in the guilt trial. For various reasons, chiefly the prosecutor's not seeking a death penalty, not all defendants convicted of murder have a penalty trial.

The sampling plan used in this study was stratified by four case types (death, life sentence after penalty trial, life with no penalty trial, and voluntary manslaughter) in each of the state judicial circuits. One-hundred-percent samples were taken in the first two categories and varying sampling fractions were used in each judicial circuit in the other two categories. In analyzing the data I treated the sampling plan as if an independent, biased coin toss had been made for each offender to determine whether he or she entered the sample, the probability of heads being equal to the sampling rate for the subpopulation containing that offender.

Descriptive Statistics

Descriptive statistics, such as estimated population totals, need to be adjusted in some way to remove the bias introduced by stratified sampling. The most straightforward adjustment is of course the use of case weights equal to the reciprocal of the sampling fraction. In other words, in any accumulation, whether a count of cases or a sum of products, the contribution of each case is multiplied by the case weight. This produces consistent, if not unibased estimates.

Inferential Statistics

Statistical inference generally involves the estimation not only of a population parameter but also of the standard error of the parameter estimate. Complex sampling generally invalidates the usual standard error formulas applicable to simple random sampling. This does not present an insuperable problem for simple descriptive statistics such as proportions, totals, and means, but it does create difficulties in the use of more complex multivariate methods such as linear and logistic regression and log-linear models.

Since stratification on independent variables does not bias their coefficients in linear or logistic regression, there is no need to apply case weights or any other adjustment to obtain unbiased, fully efficient estimates of regression coefficients unless, as in the Charging and Sentencing Study, the stratification scheme does involve the dependent variable. Although regression coefficients may be biased, it is not a particularly difficult theoretical problem to derive consistent, efficient estimates of regression coefficients under Y-stratification; however, statistical software for this purpose is not widely available nor sufficiently general to handle all cases. To give one example, the correct method of adjusting for sampling bias in log-linear models fit by the iterative proportional scaling algorithm is to set the initial fitted value in each cell to the sampling fraction rather than to 1; one program which has this capability is BMDP4F.

Efficient estimation of linear or logistic regression models is not quite so easily achieved on existing software (Goldberger, 1981; Woodworth, 1982; Manski and Lerman, 1977). One fairly simple, consistent, but inefficient method is Manski and Lerman's WESML, which, briefly described, consists of weighting the logarithm of the likelihood of the ith case by the reciprocal of the sampling fraction. For any exponential type likelihood this amounts to computing the sufficient statistics using case weights, then proceeding as in the unweighted case. In other words, use weighted linear regression or use weighted cell counts in logistic regression (BMDPLR has this capability via the COUNT option.

The problem with WESML is that the "nominal" standard errors of regression coefficients as printed by the computer program will be unreliable; however, I have made use of two simple techniques for getting consistent estimates of these standard errors: for weighted linear regression, I adapted Cressie's "safe" method for estimating the standard error of a misweighted mean (Cressie, 1982), which requires no more than a weighted linear regression program capable of placing residuals on file. For logistic regression, I applied the ideas of Grizzle, Starmer, and Koch (1969) and Bishop, Fienberg, and Holland (1975, p. 148) to derive what I provisionally call the Modified Mantel-Haenszel Procedure.

A Safe Standard Error

In the weighted regression of, say, death sentence (DSENTALL) on race of victim (WHVICRC) and the 10 other independent variables in Table 1, the regression coefficient for race of victim (X) can be expressed as:

$$\text{COEFF} = \frac{\text{Sum of } (\text{XRES}*Y*W)}{\text{Sum of } (\text{XRES}^2*W)}$$

where XRES is the residual from the weighted regression of race of victim on the 10 other independent variables and W is the case weight. In other words the regression coefficient is a weighted sum of values of the dependent variable (Y), and consequently the standard deviation of the regression coefficient is

$$\text{STDEV OF COEFF} = \frac{\sqrt{\text{Sum of } (\text{YVAR}*(\text{XRES}*W)^2)}}{\text{Sum of } (\text{XRES}^2*W)}$$

where YVAR is the residual variance of the dependent variable for the ith case. In practice, a conservative estimate of YVAR is the squared residual of Y divided by $(1-\text{HAT})^2$, where HAT is the diagonal entry in the "hat" matrix. The resulting "safe" standard error estimate for the regression coefficient is obtained by replacing YVAR in the above equation by this quantity.

Exhibit 1 shows how to compute this using MINITAB.

```
MTB > INFO

COLUMN       NAME        COUNT

  C1         DSENTALL     1066
  C2         BLACKD       1066
  C3         WHVICRC      1066
  C4         BLVICMOD     1066
  C5         STRANGER     1066
  C6         TWOVIC       1066
  C7         FELMUR       1066
  C8         VPCARBR      1066
  C9         TORTURE      1066
  C10        MENTORT      1066
  C11        NOKILL       1066
  C12        LDFB7D       1066
  C13        WEIGHT       1066

 MTB > REGRESS C1 11 C2-C12;
SUBC > WEIGHTS C13;
SUBC > HI C14;          (save HAT in col. 14)
SUBC > RESIDUALS C15. (save Y-residuals in col.15)

                                ST. DEV.      T-RATIO =
COLUMN       COEFFICIENT        OF COEF.      COEF/S.D.

             -0.04473           0.02259       -1.98
BLACKD        0.03930           0.02007        1.96
WHVICRC       0.08020           0.02040        3.93
BLVICMOD     -0.01747           0.01313       -1.33
STRANGER      0.04343           0.01803        2.41
TWOVIC        0.17429           0.03774        4.62
FELMUR        0.05038           0.02193        2.30
VPCARBR       0.03104           0.01287        2.41
TORTURE       0.42199           0.05684        7.42
MENTORT       0.22550           0.05173        4.36
NOKILL       -0.09444           0.02528       -3.74
LDFB7D        0.13924           0.02682        5.19

 MTB > NAME C14 'HAT' C15 'Y RESID'
 MTB > REGRESS C3 10 C2 C4-C12;
SUBC > WEIGHTS C13;
SUBC > RESIDUALS C16. (save X-residuals in col. 16)

 MTB > NAME C16 'X RESID'
 MTB > LET C17 = C13*(C16**2)
 MTB > SUM C17 K2
     SUM = 207.06
 MTB > LET C18 = ((C13*C15*C16)/(1-C14))**2
 MTB > SUM C18 K4
     SUM = 22.767
 MTB > LET K4 = SQRT(K4)/K2
 MTB > PRINT K4
 K4     0.0230443 (safe standard error for WHVICRC)
```

Exhibit 1. Computation of safe standard error estimate.

The Modified Mantel-Haenszel Procedure

To study the relationship between death sentencing and race of victim, controlling for race of defendant and the other 9 independent variables, it is necessary to develop a scale summarizing the influence of these variables, since it is impractical to cross-tabulate on all of them. The method which I used was to run the 11 variable weighted regression (linear or logistic) and then compute an 'aggravation' scale by summing the 9

nonracial variables times their regression coefficients. This step is illustrated using MINITAB in Exhibit 2.

```
MTB > INFO

COLUMN       NAME       COUNT
  C1         DSENTALL    1066
  C2         BLACKD      1066
  C3         WHVICRC     1066
  C4         BLVICMOD    1066
  C5         STRANGER    1066
  C6         TWOVIC      1066
  C7         FLEMUR      1066
  C8         VPCARBR     1066
  C9         TORTURE     1066
  C10        MENTORT     1066
  C11        NOKILL      1066
  C12        LDFB7D      1066
  C13        WEIGHT      1066

 MTB > REGRESS C1 11 C2-C12 C21 C22;
SUBC > WEIGHTS C13.

COLUMN       COEFFICIENT   ST. DEV. OF COEF.   T-RATIO = COEF/S.D.
                -0.04785        0.02239            -2.14
BLACKD           0.04308        0.01996             2.16
WHVICRC          0.08422        0.02027             4.15
BLVICMOD        -0.01775        0.01315            -1.35
STRANGER         0.04391        0.01799             2.44
TWOVIC           0.17220        0.03766             4.57
FELMUR           0.05076        0.02202             2.31
VPCARBR          0.02978        0.01287             2.31
TORTURE          0.42050        0.05677             7.41
MENTORT          0.22503        0.05174             4.35
NOKILL          -0.09701        0.02522            -3.85
LDFB7D           0.14012        0.02689             5.21

MTB > LET C32 = C22-.04308*C2-.08422*C3
MTB > MIN C32 K1
      MINIMUM = -0.16261
MTB > MAX C32 K2
      MAXIMUM =  0.83247
MTB > LET C33 = (C32-K1)/(K2-K1)
MTB > LET C34 = ROUND(C33*10+.5)
MTB > RECODE 5 11 C34 5 C34
MTB > NAME C34 'AGGLEVEL'
```

Exhibit 2. Computation of aggravation scale.

The aggravation scale is then discretized into 5 levels of aggravation (1 = lowest, 5 = highest). I used 10 equal length intervals and collapsed the upper 6 into one since they contained very few cases. The data were then aggregated into 10 two-by-two tabulations as shown in Table 2.

The column labelled "LogCPR" contains the logarithm of the cross product ratio for each of the two-by-two tables (Race of Victim by Sentence). It is a measure of the disparity between the death sentencing rates for white- and black-victim cases defined as the natural logarithm of the ratio of the odds on death for white-victim cases divided by the odds on death for the black-victim cases. In the table for aggravation level

Level of agg	Race of def	Race of vic	Life	Death	LogCPR	STD ERR
1	WHITE	BLACK	19.7	0	UNDEFINED	
		WHITE	77.0	0		
	BLACK	BLACK	550.7	0	UNDEFINED	
		WHITE	32.3	0		
2	WHITE	BLACK	36.1	0	UNDEFINED	
		WHITE	315.8	13.4		
	BLACK	BLACK	760.2	10.2	1.90	.60
		WHITE	55.7	5.0		
3	WHITE	BLACK	1.0	1.0	-1.90	1.48
		WHITE	55.1	8.2		
	BLACK	BLACK	42.3	1.0	2.35	1.11
		WHITE	48.2	12.0		
4	WHITE	BLACK	5.4	1.0	.69	1.22
		WHITE	49.1	18.0		
	BLACK	BLACK	51.8	3.0	2.23	.71
		WHITE	39.6	21.2		
5	WHITE	BLACK	0	0	UNDEFINED	
		WHITE	5.0	17.2		
	BLACK	BLACK	14.5	4.0	3.45	1.24
		WHITE	1.4	12.0		

Table 2. Weighted total number of cases by race of victim and sentence controlling for level of aggravation and race of defendant.

4 for black defendants, for example, the odds on death for black-victim cases are 3.0 to 51.8 or .059 and are 21.2 to 39.6 or .535 for white-victim cases, so the cross product ratio is .535/.059 or 9.07 and the logarithm is 2.23 as indicated in Table 2. Note that a negative logCPR indicates that white-victim cases had a lower risk of death while a positive logCPR indicates that white-victim cases had a higher risk of death.

In order to judge the overall significance of the pattern of logCPR's it is necessary to estimate their standard errors. Assuming a Poisson model for the occurrence of a case with a particular set of characteristics, it can be shown that a consistent estimate of the variance of a weighted count is the sum of squared case weights for the cases contributing to the count. A standard "propagation of error" argument (Wilks, 1962, theorem 9.3.1) shows that the variance of the log of a weighted count is approximated by the sum of squared case weights divided by the square of the sum of the case weights. Exhibit 2 illustrates how to compute the sums and sums of squares of case weights using MINITAB.

Since the logCPR is a linear combination of four log counts, its variance is the sum of variances of its four component log counts. The hand calculation of logCPR and its standard error is displayed in Exhibit 4.


```
MTB > INFO

COLUMN       NAME       COUNT
  C1         DSENTALL    1066
  C2         BLACKD      1066
  C3         WHVICRC     1066
  C13        WEIGHT      1066
  C14        WT2         1066 (squared case weights)
  C34        AGGLEVEL    1066

 MTB > LET C51 = 10*'WHVICRC' + 'DSENTALL'
 MTB > LET C51 = 10*'AGGLEVEL' + 'BLACKD'
 MTB > NAME C51 'WV X DS' C52 'AG X BD'
 MTB > TABLE C52 C51;
SUBC > SUMS C13.

ROWS: AG X BD      COLUMNS: WV X DS

          Black Victim                White Victim
          Life      Death             Life      Death
           00        01                10        11
10       19.7320     ---             262.5539   1.0000
11      550.7267     ---              32.3052    ---
20       36.0966     ---             315.7794  13.2000
21      760.2344  10.2000             55.6543   5.0000
30        1.0000   1.0000             55.1354   8.2000
31       42.2908   1.0000             48.2030  12.0000
40        5.4200   1.0000             49.0809  18.0000
41       51.8375   3.0000             39.5577  21.2000
50         ---      ---               5.0000  17.2000
51       14.5310   4.0000             1.3800  12.0000

 CELL CONTENTS --
        WEIGHT:SUM

 MTB > TABLE C52 51;
SUBC > SUMS C14.

ROWS: AG X BD      COLUMNS: WV X DS

          Black Victim                White Victim
          Life      Death             Life      Death
           00        01                10        11
10        76.984     ---             707.362   1.000
11      1785.545     ---             132.174    ---
20       104.692     ---             838.572  13.440
21      2645.850  10.440             151.926   5.000
30         1.000   1.000             165.264   8.440
31       156.858   1.000             119.148  12.000
40        11.776   1.000             108.220  18.000
41       148.878   3.000              95.476  21.440
50         ---      ---               5.000  17.440
51        42.935   4.000              1.904  12.000

 CELL CONTENTS --
          C14:SUM
```

Exhibit 3. Computation of sums and sums of squares of weights.

		SENTENCE LIFE		DEATH	
BLACK VICTIM	I	51.8	I	3.0	I
	I	148.9	I	3.0	I
WHITE VICTIM	I	39.6	I	21.2	I
	I	95.5	I	21.4	I

$$\text{LOGCPR} = \ln \frac{51.8 \times 21.2}{3.0 \times 39.6} = 2.23$$

$$\text{STD ERR} = \sqrt{\left(\frac{148.9}{51.8^2} + \frac{3}{3^2} + \frac{95.5}{39.6^2} + \frac{21.4}{21.2^2}\right)}$$

KEY

I	SUM OF WEIGHTS		I
I	SUM OF SQUARED WEIGHTS		I

Exhibit 4. LogCPR computation for black defendants at aggravation level four.

The Modified Mantel-Haenszel statistic is based on the weighted average of the logCPR's, weighted by the reciprocals of their squared standard errors. In other words,

$$\text{Avg logCPR} = \frac{\text{Sum of logCPR/(Squared Std. Err.)}}{\text{Sum of (1/Squared Std. Err.)}}.$$

The standard error of the average logCPR is

$$\text{Std Err of Avg logCPR} = \frac{1}{\sqrt{\text{Sum of (1/(Squared Std Err))}}}$$

The Modified Mantel-Haenszel statistic is the z-ratio for the average logCPR, i.e.,

$$\text{MMH} = \frac{\text{Average logCPR}}{\text{Standard Error of Avg logCPR}}.$$

The computation of this statistic using MINITAB is shown in Exhibit 5. The MMH statistic is interpreted as a z-score.

For those familiar with logistic regression, the weighted average logCPR is an efficient estimate of the regression coefficient of the race of victim in the logistic regression of sentence (DSENTALL) on level of aggravation (AGGSCALE) treated as a categorical variable, race of defendant, and the interaction of level of aggravation and the race of defendant.

The data of Exhibit 3 were entered into columns 1 through 9 of the MINITAB worksheet.

(Sums of case weights)

C1	C2	C3	C4	C5
10	19.73	0.00	262.55	1.00
11	550.73	0.00	32.31	0.00
20	36.10	0.00	315.78	13.20
21	760.23	10.20	55.65	5.00
30	1.00	1.00	55.14	8.20
31	42.29	1.00	48.20	12.00
40	5.42	1.00	49.08	18.00
41	51.84	3.00	39.56	21.20
50	0.00	0.00	5.00	17.20
51	14.53	4.00	1.38	12.00

(Sums of squares of case weights)

C1	C6	C7	C8	C9
10	76.98	0.00	707.36	1.00
11	1785.54	0.00	132.17	0.00
20	104.69	0.00	838.57	13.44
21	2645.85	10.44	151.93	5.00
30	1.00	1.00	165.26	8.44
31	156.86	1.00	119.15	12.00
40	11.78	1.00	108.22	18.00
41	148.88	3.00	95.48	21.44
50	0.00	0.00	5.00	17.44
51	42.93	4.00	1.90	12.00

```
MTB > LET C10 = LOG(C2*C5/(C3*C4))
MTB > LET C11 = SQRT(C6/C2**2+C7/C3**2+C8/C4**2
                                    +C9/C5**2)
MTB > NAME C1 'TABLE' C10 'LOGCPR' C11 'STD ERR'
MTB > PRINT C1 C10 C11

ROW    TABLE       LOGCPR       STD ERR
  1       10      6.62212       100.006
  2       11      2.83601       141.422
  3       20      9.62171       100.001
  4       21      1.90152         0.595
  5       30     -1.90566         1.476
  6       31      2.35405         1.106
  7       40      0.68700         1.225
  8       41      2.22574         0.705
  9       50      1.23547       141.422
 10       51      3.45281         1.240

MTB > LET C12 = 1/C11**2
MTB > MULT C10 C12 C13
MTB > SUM C13 K1
   SUM = 13.605
MTB > SUM C12 K2
   SUM = 7.4294
MTB > LET K3 = K1/K2
MTB > PRINT K3
    K3 = 1.83118
MTB > LET K4 = 1/SQRT(K2)
MTB > PRINT K4
    K4 = 0.366880
MTB > KET K5 = K3/K4
MTB > PRINT K5
    K5 = 4.99122
```

Exhibit 5. Computation of modified Mantel-Haenszel statistic.

References

Bishop, Yvonne M.M., Stephen E. Fienberg, and Paul W. Holland (1975). Discrete Multivariate Analysis. MIT Press.

Cressie, Noel (1982). "Playing safe with misweighted means," Journal of the American Statistical Association, no. 380, 77:754-759.

Goldberger, Arthur (1981). "Linear regression after selection," Journal of Economics, 15:357-366.

Grizzle, J.E., C.F. Starmer, and G.G. Koch (1969). "Analysis of categorical data by linear models," Biometrics, 25:489-504.

Manski, C.F. and S.R. Lerman (1977). "The estimation of choice probabilities from choice-based samples," Econometrica.

Manski, C.F. and D. McFadden ( ). "Alternative estimates and sample designs for discrete choice analysis," in Structural Analysis of Discrete Data, edited by C. Manski and D. McFadden, pp. 2-50. Cambridge: MIT Press.

Wilks, Samuel S. (1962). Mathematical Statistics. John Wiley and Sons.

Woodworth, George G. (1982). "Y-stratification in linear regression," Technical Report no. 84. University of Iowa, Department of Statistics.

Some Consequences of Convenience Samples in Criminal Justice Research

Richard A. Berk, *Department of Sociology*
University of California, Santa Barbara (1)

I. Introduction

Convenience data sets are endemic in criminal justice research. Whether because of budget constraints, practical difficulties, or design errors, probability samples from known populations are relatively rare. Popular instances often rest on official sources of various kinds: UCR extracts, arrest statistics, published clearance rates, and others.

It is widely recognized that causal models estimated from such data bases are generalized at some risk; external validity is always in some doubt. Less well known is that causal models estimated from convenience samples often rest on weak internal validity unless the means by which the sample was selected are explicitly taken into account. And this follows even if one is prepared to limit one's conclusions to the available data. In other words, one cannot define away the problem by making one's results conditional upon the data that can be obtained.

Under the rubric of sample selection bias (Heckman, 1979), I will briefly summarize here some recent econometric literature on problems that can develop when causal modeling is undertaken with convenience samples. (2) Potential remedies will also be discussed along with an illustration from a real data set. It cannot be overemphasized, however, that my exposition is necessarily superficial and that in addition to many primary sources (e.g., Tobin, 1958; Amemiya, 1973; Heckman, 1976, 1979; Goldberger, 1981), there now exist a number of textbook-level discussions (e.g., Judge et al., 1980: 609-616; Berk, 1983; and especially Maddala, 1983). My goal is to alert law and justice researchers to an important problem that is typically overlooked.

II. Some consequences of convenience data sets

Perhaps the best way to gain an initial understanding of the consequences for causal modeling of non-probability samples is to consider Figures 1 through 4. In each figure, there is an initial population of interest represented by a schematic scatter plot shaped like a parallelogram. (3) However, for reasons that will soon be considered, one cannot observe either the endogenous variable (Y) or the exogenous variable (X) for a non-random subset of cases. In Figure 1, cases with Y-values below some horizontal threshold are not available for study. One implication is that the proper regression line (labeled "before") is replaced by an attenuated regression line (labeled "after"). In other words, when a least-squares regression line is constructed from the subset of cases, the impact of X on Y is underestimated.

FIGURE 1

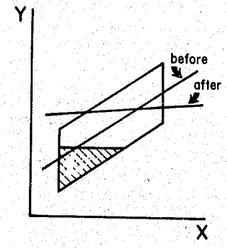


FIGURE 2

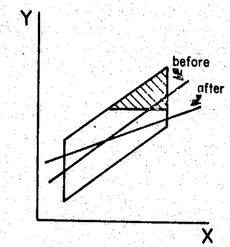


FIGURE 3

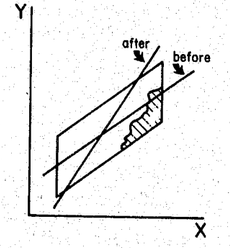


FIGURE 4

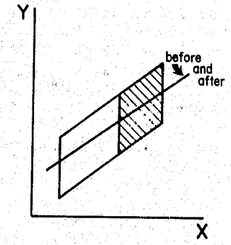


Figure 2 shows what happens when observations on X and Y are eliminated from above. Now Y-values falling above some horizontal threshold are lost. Again, the after regression line is attenuated compared to the before regression line.

Figure 3 represents a more complicated pattern. Once again, part of the scatter plot is lost, but not through a single, horizontal threshold on the endogenous variable. Note that in this instance, the after regression line is inflated relative to the before regression line.

Finally, consider Figure 4. There is now a vertical threshold for the exogenous variable, and cases with X-values to the right of this threshold are not available for study. However, the before and after regression lines now correspond.

To briefly summarize, it should first be apparent that when a non-random subset of observations is excluded, a least-squares regression line built on the remaining data will not correspond to the regression line from the original data set. (4) Second, the difference between the before

and after regression lines will be a function of the number and location of observations lost. Exclusion via a vertical threshold on the exogenous variable does not distort the regression line as long as the true relationship is linear over the full range of the exogenous variable. Other kinds of exclusions are more problematic. Third, given differences between the before and after regression line, external validity is clearly in jeopardy. One would be foolish to generalize from the subset regression line to the regression line based on the full data set.

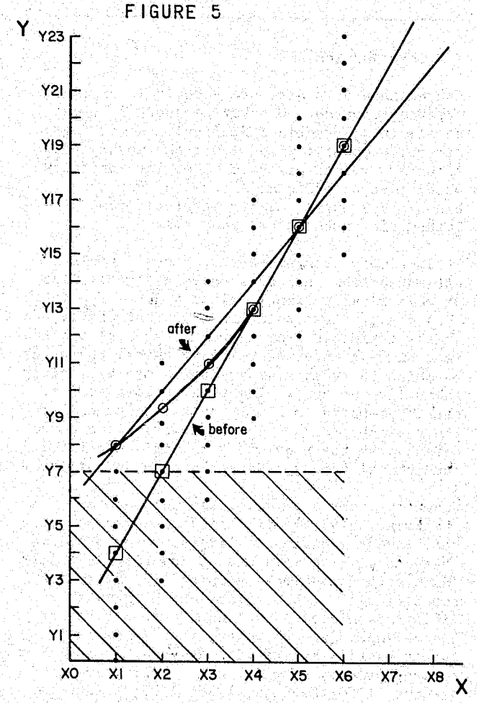
But what about internal validity? Suppose that one were only interested in regression results for the non-random subset. Would not the after regression line then suffice? To answer this question we must examine the before and after regression lines a bit more carefully.

Consider Figure 5. Suppose that Y is a measure of the seriousness of incidents of wife battery and X is the number of prior wife battery incidents. Should we take a random sample from the population of such altercations, we might obtain a scatter plot much like the one represented by the dots in Figure 5. If we assume, for purposes of illustration, that the number of prior incidents is the only systematic cause of incident seriousness and that for each observation the expectation of the disturbance term is zero, an estimated linear regression function capturing the impact of X on Y will (within sampling error) pass through the means of Y for each value of X. (5) Unbiased estimates of the slope and intercept follow. In Figure 5, the conditional means for Y are indicated by boxes, and the regression line is labeled "before."

Now suppose that police only make an arrest in wife battery incidents if the dispute exceeds some threshold of seriousness (Berk et al., 1983). Then, if the data come solely from arrest reports, less serious incidents below the threshold will be excluded. In Figure 5, observations in the shaded area below Y7 are missing. One important consequence is that the conditional means of Y for low values of X are substantially altered. For example, when X equals X1, the new mean of Y is Y8. When X equals X2, the new mean of Y is Y9.5. These new conditional means are indicated by circles.

Focusing on the pattern of new conditional means for Y, it is apparent that the original least-squares regression line no longer fits. Indeed, no straight line can pass through these conditional means because the conditional means now follow a non-linear path. If one attempted to fit a straight line to the data, the "after" regression line might result. It is less steep than the original regression line but, more importantly, does not pass through the conditional means of Y. This implies that the disturbances are correlated with X. In short, one has misspecified the estimated relationship between Y and X while confounding the impact of X with the impact of the disturbances; one has a biased and inconsistent estimate of the slope and intercept. And the estimate is biased even if one only wants to model the impact of X on Y for wife battery incidents falling above the seriousness threshold.

FIGURE 5



It is easy to demonstrate more formally what is illustrated in Figure 5. Consider the usual bivariate regression model with disturbances meeting all of the requisite assumptions.

$$Y_i = \beta_0 + \beta_1 X_i + U_i \qquad (1)$$

$$U_i \sim IN(0,\sigma^2)$$

Suppose that for a case to be included in the analysis, Y must be equal to or exceed some threshold C. This implies that

$$\beta_0 + \beta_1 X_i + U_i \geq C \qquad (2)$$

or

$$U_i \geq C - \beta_0 - \beta_1 X_i.$$

It then follows that

$$E(U_i | U_i \geq C - \beta_0 - \beta_1 X_i) \neq 0. \qquad (3)$$

Equation 3 indicates that the conditional expectation of the disturbances for each case cannot equal zero when cases falling below the Y threshold are excluded. Moreover, the conditional expectation is a function of X.

Figure 5 is an example of "explicit selection" (Goldberger, 1981). Cases for study are chosen if, in this instance, they fall on or above some threshold defined on the endogenous variable of interest. Figures 1 and 2 are also based on explicit selection; there exists a horizontal threshold operating on the scatter plot to be analyzed.

Matters become far more complicated when "incidental selection" is at work (Goldberger, 1981). There is no longer a horizontal threshold for the endogenous variable of interest. Rather, the exclusion of certain observations comes about through processes involving another endogenous variable. The shaded area in Figure 3 could well result from incidental selection.

Consider the following formulation, simplified from Heckman's work (1979). Suppose that one is interested in the length of sentences given to convicted burglars. For expositional purposes (and with no loss in generality), assume that the only systematic cause of sentence length is the number of prior felony convictions. Equation 4a shows the linear regression equation that would follow.

$$Y_{1i} = \alpha_0 + \alpha_1 X_{1i} + U_{1i} \qquad U_{1i} \sim IN(0,\sigma_{11}^2) \qquad (4a)$$

Burglary sentences can only be observed for individuals convicted of burglary. Following legal convention, assume that burglary convictions require proof of guilt beyond a reasonable doubt. One could then imagine a second regression equation with the strength of evidence as the endogenous variable. For simplicity (and with no loss of generality), assume that the strength of the evidence is a linear function of a single systematic variable, the number of eyewitnesses. Thus:

$$Y_{2i} = \beta_0 + \beta_1 X_{2i} + U_{2i} \qquad U_{2i} \sim IN(0,\sigma_{22}^2) \qquad (4b)$$

Finally, assume that there is some threshold on the endogenous variable that must be exceeded for a conviction to occur. With no loss of generality, assume that the guilt threshold is coded as zero. The threshold can then be formulated as requiring that

$$\beta_0 + \beta_1 X_{2i} + U_{2i} > 0 \qquad (5)$$

or

$$U_{2i} > -\beta_0 - \beta_1 X_{2i}$$

Now, recall that a properly specified regression function should correspond to the conditional expectation function for the endogenous variable. Under incidental selection this implies:

$$E(Y_{1i} | X_{1i}, Y_{2i} > 0) = \qquad (6)$$

$$\alpha_0 + \alpha_1 X_{1i} + E(U_{1i} | U_{2i} > -\beta_0 - \beta_1 X_{2i})$$

Equation 6 indicates that if the disturbances in equations 4a and 4b are correlated, the disturbance term in 4a will not have an expectation of zero. Furthermore, the expectation will be a function of the regressor in 4b (i.e., the number of eyewitnesses). Thus, if equations 4a and 4b are linked through correlated disturbances and one applies ordinary least squares to equation 4a as usual, one will obtain biased and inconsistent estimates of the intercept and slope. In addition, the bias will increase with the size of the correlation. Stated a bit more concretely, the processes determining whether an observation will be available for study will be confounded with the processes of substantive interest. In our example, factors altering the probability of conviction for burglary will be confounded with factors influencing sentences given to convicted burglars. The larger the correlation between chance process affecting the two outcomes, the more complete the confounding.

More generally, equations 4a and 4b are called "seemingly unrelated" in the econometrics literature (e.g., Judge et al., 1980: 245-257) and are likely when two or more regression equations represent two or more processes unfolding in the same setting, at about the same time, and/or with the same actors. Under such circumstances, random perturbations are likely to simultaneously affect each of the endogenous variables. In our example, the demeanor of the suspect, conceptualized as a random perturbation, would affect the likelihood of a guilty finding and the severity of any sentence that might follow. Ordinarily, seemingly unrelated equations can be estimated one at a time or, if greater efficiency is desired, as a set. In both instances, unbiased estimates will result. However, if the seemingly unrelated equations capture part of a process by which some observations are lost, bias will be introduced into estimates based on any non-random subset of cases.

III. When are convenience samples a problem?

There can be no disputing that all convenience samples risk biased regression estimates, even

if one is only interested in the data on hand. But biased estimates are not inevitable. Under incidental selection the key is whether the disturbances between the selection process (e.g., equation 4b) and the substantive process (e.g., equation 4a) are correlated. When the correlation is zero, one has (with respect to estimation bias) the equivalent of a simple random sample. Unfortunately, even if one is able to formulate the selection processes as a set of structural equations, there is no direct way to estimate the correlation between disturbances across equations in an unbiased (or even consistent) manner. In the absence of any straightforward empirical diagnostics, therefore, one must rely on previous research or social science theory. However, there is to date very little work on sample selection bias in criminal justice settings and virtually no formal theory on when cross-equation disturbance correlations are likely to be problematic. In short, there is very little guidance to be found.

For explicit selection, the diagnostic issues are far more simple. Since a single equation is involved (i.e., the substantive and selection processes involve the same endogenous variable), the degree of bias depends solely on the proportion of observations lost. It is even possible under some circumstances to formally derive the direction and size of the bias (Goldberger, 1981).

As a means to make these ideas more concrete, consider the following examples.

1. Suppose one is collecting data from insurance companies on the value of household property lost through theft or burglary. However, for the insurance companies in question, virtually all policies are sold with the requirement that losses must exceed $100 before claims can be made. This means that claims below $100 will not be available for study. There is, therefore, a threshold on the endogenous variable of interest, and one has an instance of explicit selection. Regression estimates of the effects of exogenous variables on the dollar loss will be biased, even if one is concerned solely with losses over $100.

2. Suppose one is interested in modeling the number of crimes committed by young males during the previous year. However, one only has access to incarcerated populations. To the degree that chance factors affecting the risk of incarceration are correlated with chance factors affecting the number of crimes committed in the year prior to imprisonment (which is almost certain), incidental selection will be at work. Regression estimates of the causes of criminal activity will then be biased, and the bias will be more severe if the correlation is large.

3. Suppose that one is attempting to interview a random sample of adults about victimization. Unfortunately, there are a number of refusals leading to a response rate of 75 percent. Then, any causal analysis of the victimization data risks incidental selection bias insofar as chance factors affecting the likelihood of refusing to be interviewed are correlated with chance factors affecting victimization.

4. Continuing the victimization example, suppose one is interested in the severity of incidents reported in a single year. The questionnaire is constructed to reflect the fact that one must be victimized before questions on severity are asked. That is, appropriate skip patterns are built in. Unfortunately, insofar as chance factors affecting the likelihood of any victimization are correlated with chance factors affecting severity, a regression analysis of victimization severity will be jeopardized by incidental selection bias.

5. Suppose that one is doing a study of the causes of recidivism and is following a cohort of ex-prisoners for two years after their release. Perhaps one important measure of success is the amount of time after release until a new offense occurs. However, since the follow-up period is only two years, failures after that period will not be recorded. In other words, there is a threshold on the endogenous variable of interest, which, if exceeded, means that time-to-failure cannot be observed. Such problems are often called "right-hand censoring," but can also be viewed as examples of explicit selection. Biased estimates follow under the usual regression procedures.

6. Suppose that one is doing a study of the processing of reported felonies in rural counties within a given state. A simple random sample of reported felonies is selected, therefore, excluding felonies from urban counties. Insofar as county is treated as an exogenous variable in any analyses to follow, the exclusion of cases from urban counties will not bias regression estimates. However, the sample of reported felonies from rural counties is surely not a random sample of all felony incidents in rural counties (given the ways in which crimes are reported) so that there is a risk of incidental selection bias, but from a second source.

The last example raises a more general and disturbing problem. Even with random samples from known populations, there is the prospect of incidental selection bias. A given population is almost surely a non-random subset of some larger group, which means that the potential for bias is carried along into the random sample. In other words, the possibility for sample-selection bias is virtually universal. (6) Whether the potential for bias is realized depends on how the selection and substantive processes are linked. Within the formulation we have been using, the key is the cross-equation correlations between the disturbances.

IV. Solutions

There are a number of ways to conceptualize solutions to the problems we have been discussing. Perhaps the easiest can be found in Heckman's observation (1979) that bias introduced by non-random selection actually results from an omitted variable. Recall that, within the usual regression formulation, if one omits a variable from a regression equation that is correlated with variables that are included, biased and inconsistent estimates of the regression coefficients follow. (7) Heckman (1979) shows that one must take into account the likelihood of exclusion for each observation available for study. Roughly stated, one must adjust the regression results for the risk of not being included in the sample. Phrased still another way, one must include a variable measuring the consequences of selection to control the confounded selection and substantive processes. (8)

How does one do this? The first step is to distinguish between situations in which data are unavailable on both the endogenous and exogenous variables and situations in which data are only unavailable on the endogenous variable. The former is often called truncation, while the latter is often called censoring (Maddala, 1983: 1-6). Typically, the estimation problems are most difficult under truncation. Here, the focus will be on censoring, which occurs in a wide variety of criminal justice situations. (9)

Second, one must formulate a model that includes both the substantive processes of interest and the selection method of the cases on hand. Explicit selection must be distinguished from incidental selection, and in either instance, structural equations need to be specified.

Third, the implications of the selection process must be built into the substantive structural equation(s). Often this means including a new variable in the substantive equation(s) that captures the likelihood of exclusion from the sample. Within Heckman's formulation (1979: 156), this new variable is a "hazard rate," which represents the "instantaneous probability" of exclusion from the sample conditional upon the probability of being a risk to exclusion (See Berk, 1983, for a more thorough and grounded discussion).

Finally, one must choose among several estimation procedures. Until Maximum-Likelihood software is more widely available, Heckman's (1979: 157) two step approach may well suffice. (10) Hence, under incidental selection (with a single selection equation), one begins by estimating a probit equation for each case's chance of being included in the final data set. All potential cases are included in this equation. Thinking back to our illustration of sentences for convicted burglars, the probit equation would model the likelihood of conviction (with conviction coded as "1" and no conviction coded "0"). Then, predicted values from the probit equation are used to construct the hazard rate for each case. Finally, the substantive equation (assuming one substantive equation) is estimated with the hazard rate inserted as a new regressor. Ordinary least squares may be employed, although generalized least squares is technically preferable. (11)

To make this more concrete, consider the following "real" example. For reasons that need not concern us here (Berk and Shih, 1982), there was interest in finding out how people called for jury duty reacted to the experience. In Santa Barbara County, California, the full population of people called for jury duty over a period of several months was mailed a short, self-administered questionnaire with the usual expectation that many people would not respond. In other words, incidental selection (through non-response) was anticipated. By design, however, other data from the Jury Commissioner's Office were collected on the full population. These data included a few important biographical attributes (e.g., sex, age) and rather rich material on the experiences of each person called (e.g., whether the person was selected to serve on a jury). In short, the data on the full population allowed for the possibility of correcting for selection bias when the returned questionnaires were analyzed.

Of the 498 questionnaires mailed, 69 percent were returned. By the usual standards of survey research, the response rate was quite high. Moreover, a probit analysis for the likelihood of returning the questionnaire did not reveal a large number of systematic factors affecting who responded. The major determinant was age; older individuals were more cooperative.

However, when the hazard rate was constructed and inserted in equations analyzing the questionnaire items, there was clear evidence of selection effects. Table 1 shows the results for one of the items. In the first two columns, the regression coefficients and t-values for the uncorrected results are reported. The second two columns report the regression coefficients and t-values for the corrected results. A careful comparison will indicate that in the absence of the corrected equation, one false positive and three false negatives would have been interpreted. Individuals who had served on a jury would have been incorrectly seen as more satisfied with their experience. Missed would have been the findings that individuals subjected to a more lengthy jury selection process were less satisfied while individuals who were called for the first time and who were given more notice were more satisfied. Note that each of these errors could have important policy implications should one care about making jury duty more palatable.

It is also apparent that the hazard rate has a statistically significant impact and adds 3 percent to the explained variance. Individuals who are less likely to return the questionnaire are less satisfied; complainers are less likely to respond. (12)

It cannot be overemphasized, however, that the credibility of these or any other set of results depends on the credibility of the equations being estimated. In particular, we have been assuming that, prior to selection, the substantive and selection processes are properly modeled; there are no specification errors. Consequently, specification errors (in the usual econometric sense) must be corrected first.

"All in all, how would you rate your experience of being called for jury duty?"

Very satisfied = 3	Somewhat satisfied = 2	Somewhat dissatisfied = 1	Very dissatisfied = 0
39.9% (135)	37.6% (127)	15.7% (53)	6.8% (23)

	Uncorrected		Probit correction	
Variable	Coeff.	t-Value	Coeff.	t-Value
Intercept	1.47	5.31	2.26	5.91
Hazard Rate	----	----	-1.26	-2.97
Female	0.28	2.91	0.26	2.67
Employed	0.03	0.32	0.15	1.44
White	0.16	1.18	0.12	0.92
Served	0.37	2.07	0.21	1.13
Served X Criminal Trial	0.11	0.73	0.10	0.70
Served X Length of Trial	-0.01	-0.68	0.00(a)	0.25
Served X Defendent Won	0.07	0.40	0.21	1.20
Length of Jury Selection	-0.17	-1.07	-0.33	-1.98
1st Time Called	0.16	1.57	0.20	1.99
# Days Notice	0.11	1.85	0.12	1.98
Does Not Drive	0.05	0.31	0.07	0.48
	R^2 = .10		R^2 = .13	
	F = 3.41		F = 3.94	
	P = < .001		P = < .001	

(a) Positive, but less than 0.00

Table 1. Ordinary Least Squares Analysis of Overall Dissatisfaction

Otherwise, one risks finding pseudo-selection effects that are actually artifacts produced by variables that have been omitted from the substantive and selection equations.

V. Conclusions

The potential for selection bias is virtually universal, but whether the potential for selection bias becomes a reality depends on how the selection and substantive processes are related. Unfortunately, there are no direct ways to determine the degree to which such relations exist. Hence, it is probably wise in practice to proceed as if selection bias were present and make, when possible, appropriate adjustments. The corrected results may or may not be more credible than the uncorrected results, but one then has the option of working from either (or both).

One implication is that probability samples should be collected whenever possible. While this may not eliminate the potential for selection bias (given that membership in the original population was not determined through probability sampling), it will often reduce the risk dramatically. Another implication is that research designs should anticipate possible selection biases and include plans to collect data that may be used to model the selection process. A final implication is that the case for randomized experiments is strengthened. If random assignment is undertaken after the sample is selected, treatment dummy-variables will be (within sampling error) uncorrelated with all omitted variables, including those reflecting the selection process. (13) Unbiased estimates of treatment effects follow. In the absence of randomization, one's results are only as good as one's model.

Notes

1. While working on this paper I was supported in part by a grant from the NIMH Center for Studies in Criminal and Violent Behavior (grant No. R01 MH34616-03). The data used, however, was collected in an earlier study funded by the National Institute of Justice (grant No. 80-IJ-CX-0037). Finally, the figures were done by Christine Allen.

2. One can consider the sample selection formulation to be a special case of limited dependent variables. See, in particular, Maddala (1983).

3. For simplicity of exposition, Figures 1 through 4 can be viewed as capturing relationships between Y and X for some population of interest and showing what happens when a non-random subset of observations from that population cannot be observed. Following Goldberger (1981), we can thus postpone worrying about estimation issues while an intuitive appreciation of the problem is being developed.

4. The same sort of results follow from a non-linear regression function or any formulation summarizing the expectation of an endogenous variable conditional upon values of one or more exogenous variables.

5. There are, of course, other assumptions that are usually made when ordinary least squares is applied, but these need not concern us for now.

6. For reasons that will soon be apparent, randomized experiments eliminate the potential bias in the regression coefficients.

7. Since the omitted variable's effects will be included among the disturbances, the regressors will be correlated with the disturbances. This in turn is the direct cause of the biased estimates.

8. A bit more technically, one must include a variable to capture the non-linear part of the relationship shown on Figure 5.

9. Readers interested in estimation under truncation should consult Maddala's excellent text.

10. There is some debate about proper estimation procedures and the robustness of each procedure's underlying assumptions (e.g., Maddala, 1983: 178-194). The practical implications of such concerns are still being sorted out.

11. GLS is formally superior because the OLS disturbances are necessarily heteroscedastic (Heckman, 1979: 157). In my experience, however, the OLS and GLS results are substantively indistinguishable.

12. Recall that the hazard rate captures the chances of exclusion from the sample.

13. The treatment variables will also be uncorrelated with the included regressors, within sampling error.

References

Amemiya, T. "Regression Analysis When The Dependent Variable is Truncated Normal." Econometrica 41: 997-1016 (1973).

Berk, R.A. "An Introduction to Sample Selection Bias in Sociological Data." American Sociological Review 48: 386-98 (1983).

Berk, R.A., and A. Shih. Measuring Citizen Evaluations of the Criminal Justice System: A Final Report to the National Institute of Justice. Santa Barbara, California: SPRI (1982).

Berk, R.A., S.F. Berk, D.R. Loseke, and D. Rauma. "Mutual Combat and Other Family Violence Myths." In D. Finkelhor, R.J. Gelles, G.T. Hotaling, and M.A. Straus (eds.), The Dark Side of Families. Beverly Hills: Sage Publications (1983).

Goldberger, A.S. "Linear Regression after Selection." Journal of Econometrics 15: 357-66 (1981).

Heckman, J.J. "The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimation for Such Models." Annuals of Econometric and Social Measurement 5: 475-92 (1976).

Heckman, J.J. "Sample Selection Bias as a Specification Error." Econometrica 45: 153-61 (1979).

Judge, G.G., W.E. Griffiths, C.R. Hill, and T. Lee. The Theory and Practice of Econometrics. New York: John Wiley (1980).

Maddala, G.S. Limited-Dependent and Qualitative Variables in Econometrics. New York: Cambridge University Press (1983).

Tobin, J. "Estimation of Relationships for Limited Dependent Variables." Econometrica 26: 24-36 (1948).

The Civil Litigation Research Project: Lessons for Studying the Civil Justice System*

Herbert M. Kritzer, *University of Wisconsin-Madison*

Abstract

The Civil Litigation Research Project was a major study of the civil justice process, funded by the United States Department of Justice. This paper summarizes major findings of the Project and discusses the implications of those findings and the more general experience of the Project for future research on the civil justice system.

Introduction

Research on court systems in the United States has been an ongoing phenomenon for much of the twentieth century; one need only look to the work of Charles Clark (e.g., Clark and Shulman, 1937) or to the crime surveys of the 20's and 30's (see Nardulli, 1978: 3-39) for prominent examples of such work. Over the last two decades, the attention of scholars and court reformers has focused heavily on the criminal side of the justice system. Much of this research began as what might be called "gap" studies (see Nelken, 1981; Sarat, 1983), in which scholars and/or commentators discovered substantial gaps between the ideal (or ideology) of the system, and the way the system worked on a day by day basis. The gap study is most clearly represented in much of the work on plea bargaining (see particularly the series of articles by Alschuler, 1968, 1975, 1976), though more recent work on plea bargaining (see Utz, 1978; Feeley, 1979) has suggested that the perceived gap may have reflected a misunderstanding of both the basis and the goal of the bargaining process. The gap study is particularly important for justice system policy because such studies often lead to proposals for reform. At the same time, however, a recent critical evaluation of criminal justice reform (Feeley, 1983) has suggested that many of the reform efforts have sought to oversimplify a complex phenomenon and have often intervened in situations that might have been better left alone.

In the last few years, court reformers have begun to turn their attention to the civil justice system. Spurred on by comments of the Chief Justice (see, for example, the Washington Post, January 25, 1982, p. 8), reformers have sought to channel disputes typically processed by the civil courts to other forums (e.g., see the report of the Pound Conference, 70 F.R.D. 79 [1976]; Johnson and Schwartz, 1978; Sander, 1982; Hensler, Lipson, and Rolph, 1981; Lind and Shapard, 1981). Dissatisfaction with what has become known as "discovery abuse" has led to proposals to alter rules of civil procedure to prevent such problems. Concern about problems of delay have resulted in new rules to require federal judges to impose a timetable on pretrial preparation. Problems of high cost have led to suggestions for new procedures to handle "modest" cases (Epstein, 1981; Rosenberg, Rient, and Rowe, 1981) and to a consideration of modifications of the American rule regarding costs in litigation (Rowe, 1982a, 1982b).[1] Burgeoning case loads have led judges to expand their efforts to settle cases short of trial as a means of relieving the trial burden (Kritzer, 1982; Galanter, 1983).[2] The fundamental question for applied civil justice research is whether these reform efforts reflect (1) an accurate understanding of the condition of the civil justice system, (2) an understanding of the actual working of the system, and (3) an understanding of the purposes of the users of the system.

The amount of research on the civil justice is system small in comparison to that which has resulted from the massive resources that have been poured into studying the criminal justice system. Hurst's recent review (1980-81) of this (and other) research identifies a number of research themes and a number of relatively small scale research projects that have dealt with questions of civil justice. Yet the relatively narrow findings reported by Hurst reflect the absence of substantial support for research in this area. One can identify only a handful of large scale civil justice research projects (e.g., the Columbia University School of Law Project for Effective Justice of the early 1960's; largely unreported civil justice aspects of the American jury project of the 1950's; the district court study project of the Federal Judicial Center; the Civil Litigation Research Project; and the Rand Corporation's civil jury study). The role of these large-scale projects is of central importance because (1) they have the ability of putting the questions they pose into a fairly broad context, something which small projects are often (if not usually) unable to do; (2) they provide data sets which can be the basis of secondary analyses of many questions and issues; and (3) they can often provide insights into the issues and problems of studying the civil justice system.

The purpose of this discussion is to share with you some insights from the Civil Litigation Research Project (CLRP). All of the three points above are applicable to CLRP: results of the project provide broad contextual background that has been generally lacking in discussions of civil justice; the research experience of the project staff provides important lessons for future research on civil justice; and the data we produced is already being mined by a number of other researchers (e.g., scholars at Purdue and Ohio State Universities, and analysts at the Rand Corporation and the Federal Judicial Center). My comments will focus on some of the major substantive and methdological insights of the Project. First, let me give you a little background on the way the Civil Litigation Research Project came into being and the nature of the research that we carried out.

What CLRP was and what it did

The Civil Litigation Research Project was a child of the Carter Department of Justice (DOJ). Attorney General Griffin Bell created within DOJ a planning and research office which was known as the Office for Improvements in the Administration of Justice (for some background on this see Sarat, 1981). Within this office (which had the fond acronym of OIAJ) was housed the Federal Justice Research Program (FJRP--FeJeRP). FJRP was a very modest program with a budget of $2,000,000 per year. In 1978, the Assistant Attorney General in charge of OIAJ, Daniel Meador, and the administrator of FJRP, Harry Scarr, decided to begin to establish a research agenda on civil justice. The major (though not the only) research to come out of this decision was CLRP. In August 1978, DOJ published an RFP (Request for Proposals) describing a broad, ambitious study of disputing behavior inside and outside the courts, with a particular focus on the costs of dispute processing. Through a competitive process a consortium of the University of Wisconsin Law School's Dispute Processing Research Program and the University of Southern California's Program for Dispute Systems Research obtained the contract. The final research project differed in many ways from the original conception of DOJ. I will not here trace the evolution from DOJ's conception to what we actually did, though that in itself is an interesting story. Needless to say, we believed that the final research design reflected a coherence and focus that was not present in the original RFP.

What was that final design?[3] The focus of the research was on dispute processing behavior in the courts, in alternative dispute processing institutions, and in what we termed "bilateral" dispute processing (i.e., negotiated resolution outside any third party institution). We selected samples of disputes from each of these modes of processing. Disputes were identified either through institutional records (in the case of courts and alternatives) or through screening surveys of households and organizations. For the institutional cases, we examined the record of the dispute and coded substantial information from those records; we also identified the major participants in the case (i.e., the disputants and their lawyers). From the screening surveys, we tried to obtain some very basic information about the dispute and to identify the participants. The disputes were drawn from five federal judicial districts (Eastern Wisconsin, Eastern Pennsylvania, South Carolina, New Mexico, and Central California); a total of 2,912 disputes were included in the overall "sample," over 1,650 of which involved law suits.

Once the participants were identified, we sought to interview as many of them as we could find; as I will discuss later, we were very successful in finding and interviewing the lawyers and much less successful either in finding or interviewing the disputants (except in the cases identified through the screening surveys). The content of the interviews included questions about the way the case was processed, the nature of the case (in terms of what was at stake), the negotiation process, relationships among the parties to the dispute, relationships between lawyers and clients, time and cost of processing the dispute, and reactions to the processing experience. All of the information from the institutional records and from the participant interviews was combined into a single public use data base; we also constructed public use versions of organizational and household screening surveys. All of these data are now available through the Interuniversity Consortium for Political and Social Research (ICPSR).

Let me now turn to a discussion of some of our empirical findings. In this discussion, I will limit my remarks to analyses dealing directly with the courts. I will not discuss our analyses dealing with general disputing experience and behavior (see, for example, Miller and Sarat, 1980-81).

Important findings

Many of our findings are summarized in a paper (Trubek et. al., 1983) to be published in the fall 1983 issue of the UCLA Law Review. The title of that paper, "The Costs of Ordinary Litigation," captures what is probably the most important single finding of our research: the "discovery" of "ordinary litigation." The typical discussion of civil litigation sees cases that are relatively substantial in scope; take for example the comments by Wayne Brazil in two of his discussions of the discovery problem:

> Most of my exposure to litigation has been in urban San Francisco with cases of moderate size involving claims ranging from $20,000 to $50,000 (1978: 1310).
>
> The sample group was also well balanced with respect to the median dollar size of the cases on which the attorneys worked. All were asked to estimate the median dollar value of the cases in which they had been involved over the preceding five years. The answers ranged from four to eight digits, the median being approximately $150,000. The figures also show that the group was well balanced on the extremes: 39 [of 180] attorneys indicated that the median value of their cases had been $25,000 or less, while 42 lawyers reported $1,000,000 or more (1980: 220).

As Table 1 shows, only 36% of federal cases and only 11% of state cases come up to Brazil's "small" case level of $25,000; I should note here that we excluded from our samples cases where the claim was for less than $1,000. The median case, as of 1978, in the state courts of general jurisdiction involved approximately $4,500. While the median federal case involved substantially more (around $15,000), the vast majority of civil cases (something between 95 and 98%) are filed in the state courts, and thus it is the state median that represents the most typical case. One possible reaction to this finding is that these cases represent the ones that seldom get to trial; that is, cases that go to trial are typically much more substantial. However, the Rand Chicago Jury Verdict study (Peterson and Priest, 1982: 22-24) has shown that the median jury verdict (even after eliminating those cases in which the defendant

wins and the verdict is $0) is on the order of $7,000, which is generally in line with our notion of ordinary litigation.

Stakes	Federal Cases	State Cases
0-5,000	26%	55%
5,001-10,000	15	18
10,001-25,000	23	16
25,001-50,000	17	7
50,001 & up	19	4
	100%	100%
(N)	(448)	(411)

Table 1. Distribution of Lawyers' Perceptions of Stakes

The nature of the everyday civil case is also to be seen in the nature of the activity reported in cases. We counted the number of discovery events, the number of nondiscovery-related motions, and the number of briefs in each case in our sample. Except for motions, the "typical" state case had no discovery and no briefs; if one discounted "motions to dismiss" because of settlement, it is likely that the typical case would also have zero motions.

	Discovery Events		Motions		Briefs	
	Federal	State	Federal	State	Federal	State
Number						
0	52%	62%	13%	28%	42%	88%
1-5	31	30	79	69	46	11
6-10	10	5	7	3	8	1
11&up	7	3	1	0	4	0
	100%	100%	100%	100%	100%	100%
(N)	(809)	(840)	(809)	(840)	(809)	(840)

Table 2. Activity level in court cases

The idea of "ordinary" litigation is of crucial importance in the study of civil justice because it should tell us to distinguish between the ordinary and the "extra"-ordinary when we consider legal, procedural, and administrative reform. As I noted above, reform is typically considered in the context of an image of the civil justice system that does not focus on the "ordinary" case, and reform that is thinking in terms of the "big" case may have very negative implications for the ordinary case. The source of this big case bias is simply that it is the big case that catches the attention of the observer and the participant; the message of our research is simply not to lose sight of the everyday, case which makes up the bulk of a court's docket. I would suggest that the recent changes to Rule 16 of the Federal Rules of Civil Procedure, whereby judges are to issue a scheduling order in every civil case on their docket, is an example where the ordinary case would be better served by a more informal tickler system (cf., Flanders, 1977: 20).

A second major message of our research, one that is closely related to our first, is that one can view much of what goes on in the civil justice system within a framework that is not particularly complicated. We devoted substantial effort in our design of questionnaires to being able to handle complex cases involving multiple parties, numbers of changes in perceptions of what was at stake, intricate negotiations, complex relationships among the participants, and multiple dispute processing forums. While cases do occur where some (or even all) of these would apply, most cases are relatively modest in what is at stake, are handled in a relatively straight-forward fashion that involves relatively little court activity (either by the court or by the lawyers), and are relatively simple in terms of the issues and relationships involved. Certainly every case has its unique aspects, but from the viewpoint of civil justice planning and administration, one can safely simplify many (perhaps even most) cases into categories of size, complexity, substance, etc.

A third major finding, again related to the message of modesty that I am presenting, is that lawyers spend relatively small amounts of time on most cases. Table 3 shows the amount of time lawyers reported spending on the cases in a part of our sample. In the median case, lawyers spend

Total Hours	Percent of Cases
0 - 8	13
9 - 24	28
25 - 40	19
41 - 80	19
81 -120	9
over 120	12
	100
Median: 30.4	N = 719

Table 3. Distribution of lawyer hours per case

less than 4 working days on the case, and in 60% of cases, lawyers spend less than a week. The message of this is that for most cases reforms that seek to reduce lawyer time might have a substantial effect because a small reduction in absolute time can often represent a substantial portion of the time devoted to a case; however, it may be difficult to make such reductions because lawyers are already spending relatively little time and there may be little that can be cut out. This finding also suggests that reforms that require lawyers to do additional things (e.g., appear in court for scheduling conferences) may significantly increase the amount of time that must be spent on cases; obviously, if the additional tasks increase the quality of the results

that lawyers achieve for their clients, the additional effort may be justified, though judgments of quality tend to reflect the eye of the beholder. One danger of many reform activities is that they seek to improve the situation from the viewpoint of the bench or from the viewpoint of the court administrator; while this may be a valid goal, one should keep in mind the potential costs to the consumer of court services.

The next finding I want to mention is directly related to the preceding point regarding the relatively modest amounts of time involved in court cases. Not surprisingly, we found that the lawyer's fee represents virtually the entire out of pocket cost of processing a case from the viewpoint of the litigant. When one factors in the value of the disputant's time (or the value of employee time for organizational litigants), the lawyer's fee represents about 80% of the total cost in the typical (median) case. Furthermore, in the median case only 8% of the lawyer's bill represents the lawyer's non-time related expenses (e.g., travel, duplicating, copying, etc.). Thus, changes that affect the amount of time a lawyer must devote to a case directly affect the cost to the litigant (assuming that the litigant is paying the lawyer on an hourly basis). If one can reduce the lawyer's time by 25%, one is going to reduce the litigant's cost by almost 25%.

The last finding I will discuss goes off on a somewhat different direction. It has become popular in discussions of civil justice in the United States to emphasize the role of settlement in disposing of cases. At a recent conference in Madison, one of my colleagues from the Law School commented that perhaps the traditional sign above the court house door, "Equal Justice Under Law," should be replaced with "Let's Make a Deal." While knowledgeable persons have known for many years that very few civil cases ever come to trial (and I will comment on the ambiguity of the notion of "coming to trial" in a few minutes), we should not underestimate the role of adjudication in disposing of cases. By adjudication, I mean an authoritative decision by a judge or an inferior court officer (e.g., a magistrate). There are two ways in which adjudication directly disposes of cases other than through trial.[4] First, a judge may directly dispose of a case through a ruling on a substantive or procedural motion (e.g., granting a motion for summary judgment). In one subset of cases, we found that only 5% of either federal or state cases were tried, but another 26% of federal and 16% of state cases were terminated through motions and involuntary dismissals (see Kritzer and Anderson, 1983). Second, a judge may indirectly dispose of a case by a ruling on a motion that answers the central dispute in a case (e.g., ruling on a point of law or on the admissibility of a piece of testimony); once that central question has been authoritatively resolved, the parties can proceed to arrive at a settlement. We have no explicit data on how often this happens, but in my discussion of "what is a trial," which I will turn to shortly, I will suggest that this is not infrequent.

There are many other findings that I could mention (on the likelihood of disputes leading to lawsuits, on how lawyers spend their time, on what accounts for the amount of time a lawyer spends on a particular case, on the background and expertise of litigating lawyers, on the nature of the outcomes of litigation, on what accounts for those outcomes, on what accounts for hourly rates charged by lawyers, on a comparison of the courts and the American Arbitration Association vis a vis case processing, and so on), but I think I have touched on the highlights that you would find most interesting (though I will be happy to talk about other areas during the discussion if anyone has specific interests).

I should also note that we have only scratched the surface of our data to date. We are now working on analyses of the nature and import of non-monetary stakes, of what lawyers think about when they consider the alternative forums where they might take a given case, and of the negotiation process. Most of our analyses have been based on interviews with lawyers and on data from court records; we have done relatively little with data from individual and organizational disputant surveys, and we haven't even touched data from a separate government lawyer survey (partly because of a relatively small number of interviews). Likewise, most of our analyses to date have focused on court cases. While we have looked at cases from the American Arbitration Association, we have not looked systematically at the cases from the ten other alternative institutions included in the study, and we have not looked at all at the bilateral cases included in our disputant surveys. We hope to be able to obtain funding to conduct some of these analyses in the future, though we know that many of them will have to be left to other researchers working with our data.

Lessons for the research community

As is true of any large study (and perhaps of any study regardless of size), I wish that I knew four years ago what I know now. There are many things I would do differently if I had to start over again. When we first heard that we had been awarded the contract, one of my colleagues said that he felt like he had been big game hunting and had shot an elephant, and that the elephant had just been delivered to his front yard. I now know how to cut up the elephant to put it into the freezer. I would like to share some of the lessons I learned with you in the hope that when and if the elephant carcus arrives on your yard, you do not have to repeat many of the mistakes that we made.

The first lesson, which should already be clear from what I said previously, is to avoid overcomplication. We tried to be very sophisticated in our understanding and research collection strategy. We would have been better off if we had ignored some of the complications that we knew existed and had sought to simplify the design. We would have missed some subtleties, but in the end, we found that there was relatively little that could be done analytically with those subtleties. Furthermore, to the degree that one is concerned with administration of civil justice, one has to simplify in order to administer; if everything is

deemed to be unique, then there is relatively little to administer. To the student of any organizational process (and that is what the civil justice process is), the goal must be to identify the similarities in order to structure the process while providing mechanisms to allow the unique aspects to be considered. In our research, we tried to include too much in the "common" component, and the result was that we later had to simply ignore some of the complexities that we captured.

The second lesson, one from elementary research methods, is that there should have been a pilot study prior to undertaking the main study. This was not a problem that we were unaware of. We sought to convince the Department of Justice of the importance of a pilot, but time and cost factors did not allow us to undertake a pilot prior to the main study. Given that there had never been a study anything like the one that DOJ wanted done, it would have made much more sense for the initial RFP to have been for a research design and pilot study (ideally, perhaps two designs and pilots by two different research organizations). That this was not done reflected the fragility of OIAJ and its need to quickly produce some research product that would legitimize its function.

Third, and this is something that we should have learned from the pilot that never existed, contrary to at least one recently reported research experience (Danet, Hoffman, and Kermish, 1980), lawyers are extremely cooperative. We were able to conduct long (averaging about an hour) and detailed (the survey instrument was over 100 pages long) interviews with most of the lawyers we contacted. Only 17% of the lawyers we contacted actually refused. An additional 17% indicated that they had not had sufficient involvement with the case to be able to discuss it in detail with us; certainly, some of these were polite refusals, but we believe that most of them represent another problem which is most clearly seen in our efforts to get information from organizations about their dispute experiences.

In order to obtain interviews with organizational litigants, we sought to identify within each organization we contacted a particular person to talk to, someone whom we referred to as the Key Organizational Decisionmaker (or, in Project lingo, the KOD). In designing the organizational questionnaire, we went to some length to allow for situations where we would need to talk to several persons in the organization to get the whole picture of what happened; to my knowledge, we never actually spoke to more than one person. The problem that this reflected was that when we designed the questionnaire, we had in mind the big complex dispute of popular concern, not the modest, routine, everyday dispute represented in our sample. The biggest problem that we encountered with organizations, and the one I believe that was indicated by the responses of lawyers who said that they had little or no involvement in cases in our sample even though they were listed as the attorney of record in the court file, was that of organizational memory. The problem with ordinary cases is that they are not particularly memorable;[5] in fact, the processing of routine cases is likely to be dispersed through an organization with no one person having any significant memory of a particular case. Case files may be hard to find, and even if they can be found they may contain little information on the way the case was processed (they probably contain primarily factual information about the dispute). This lack of institutional memory makes it extremely difficult to study organizational processing of disputes using the kind of retrospective design that we were using.

Regarding individuals, we encountered a different problem: finding them. Court records contain very little information on how to contact litigants; this is not surprising since, after the initial complaint is served, service of papers is through the lawyers (in large part because the lawyers are easier to contact than are the litigants). We tried to solicit information from lawyers about how to get ahold of their former clients, but most individual clients are "one shot" players, and the lawyers have no reason to maintain contact. Furthermore, geographic mobility is such that any information on where a person lives has a high probability of quickly becoming out of date; this is particularly true for individuals involved in divorce cases, since the termination of the marriage itself is likely to lead to relocation.

The theme that has been most dominant in this discussion has been the need to avoid overcomplication. I would like to close my discussion by suggesting at least one area where one should avoid oversimplification. One totally unexpected finding has to do with the ambiguity of terminology. For example, what is a "trial"? Given the formality, as defined in rules of civil procedure, of the civil justice process, we expected there to be relatively little ambiguity, at least in the formal proceedings. We noted in our data obtained from the court records whether there was an indication that a trial had taken place. In our interviews with the lawyers, we asked them whether a trial had taken place. We were shocked by the level of inconsistency between the court records and the lawyers' recollections.

Was there a trial according to the court record?	Was there a trial according to the lawyer? Yes	No
Yes	119	32
No	109	1117

Table 4. Is a trial a trial?

There are at least two interpretations to the inconsistencies shown in Table 4. First, one source of data, most likely the interview data, is unreliable. No doubt there is some reporting inaccuracies in the interview data, but we do not think that this is the major problem reflected by the table. The second explanation is that the notion of a trial may be much more ambiguous than we had realized. When we first encountered this inconsistency, we took a sample of the cases and went to the raw data to see if we could figure out why the inconsistency was there. What we found was that in many cases where the lawyer reported a trial but the court record did not there was a hearing on a major issue in the case which resulted in a ruling by the judge; after the ruling the case was settled. We also found that in a number of cases where the record showed a trial but the lawyer did not recall there being one, the trial may have been a hearing that simply ratified a previously worked out agreement among the parties; for one reason or another the parties either needed (e.g., in a divorce case) or wanted some formal decision from a judge to legitimize or to make enforceable their agreement.[6]
A second way in which we encountered this kind of ambiguity was in our coding of court records. Our field staff frequently encountered motions they had never heard of. When they discussed this with the local court clerk, they were told that it was standard procedure for lawyers to simply file a motion for "Y" either if "Y" fell between "X" and "Z," which were provided for in the rules, or if "Y" was something they needed or desired even if it was not provided for in the rules. Needless to say, this creativity made for problems in coding and analyzing information on the individual events that make up the litigation process.

The implication of this ambiguity is that one has to be careful in adopting for research purposes the formal categories of the civil justice process. The categories may serve to disguise the process in important ways. It is better from a research standpoint to define analytic categories and then to frame questions around those categories. For example, instead of asking about whether a "trial" occurred, we could have asked whether a decision by a judge or other court officer served to resolve the major issue in dispute; if the answer were affirmative, we could have followed up with a question concerning whether the presentation of the issue to the judge involved primarily questions of law, questions of fact, or a combination of law and fact. This would have provided a more useful analytic category scheme than simply asking whether there was something that was labeled a "trial."

Conclusion

I am very excited by the prospects for research on the civil justice system. There is much to be learned and there is much to be done. I believe that the Civil Litigation Research Project was an important beginning. We have learned a lot about the civil justice process (and there is much more yet to be learned from the data that we have in hand), and we have learned a lot about how to carry out research on this part of our court system. The civil justice system is an extremely important topic for research since it accounts for substantially more of the business of the courts than does the criminal side of justice.

Notes

* This research was funded by the U.S. Department of Justice under Contract No. J01A-79-C-0040, with supplemental funding from the University of Wisconsin Law and Graduate Schools. The senior research staff included, in addition to the author, David M. Trubek, William L. F. Felstiner, Austin Sarat, and Joel B. Grossman.

1. I suspect that current discussions of fee shifting also reflect a desire to create disincentives as a means of reducing the number of law suits.
2. Some of the interest in settlement reflects a belief that a resolution through settlement is "better" than a resolution through adjudication.
3. The research design is described in some detail in Kritzer (1980-81).
4. Adjudication is also a fundamental part of the settlement process in that the adjudication of other cases sets the context through which a given case is ultimately settled.
5. Last summer, I conducted field work in Toronto involving interviews with corporate lawyers and corporate officials. I found that, while corporations had many small cases, when I mentioned litigation to persons in these kind of positions, they immediately thought in terms of the "big" case. The "everyday" case simply gets lost within the routine of the organization or the legal practice.
6. This may be the civil equivalent of what is referred to as a "slow plea" in the study of criminal trials.

References

Alschuler, Albert W. 1968. "The Prosecutor's Role in Plea Bargaining." 36 University of Chicago Law Review 50.
_____. 1975. "The Defense Attorney's Role in Plea Bargaining." 84 Yale Law Journal 1179.
_____. 1976. "The Trial Judge's Role in Plea Bargaining, Part I." 76 Columbia University Law Review 1059.
Brazil, Wayne D. 1978. "The Adversary Character of Civil Discovery: A Critique and Proposals for Change." 31 Vanderbilt Law Review 1295.
_____. 1980. "Views from the Front Lines: Observations by Chicago Lawyers About the System of Civil Discovery." 1980 Amercian Bar Foundation Research Journal 217.
Clark, Charles E. and Harry Shulman. 1937. A Study of Law Administration in Connecticut. New Haven: Yale University Press.
Danet, Brenda, Kenneth B. Hoffman, and Nicole C. Kermish. 1980. "Obstacles to the Study of Lawyer-Client Interaction: The Biography of a Failure." 14 Law & Society Review 905.
Epstein, Norman L. 1981. "Reducing Litigation Costs for Small Cases." 20 (2) The Judges' Journal 8.
Feeley, Malcolm M. 1983. Court Reform on Trial: Why Simple Solutions Fail. New York: Basic Books.
_____.(ed.). 1979. "Special Issue on Plea Bargaining." 13 Law & Society Review 197.
Flanders, Steven. 1977. Case Management and Court Management in the United States District Courts. Washington: Federal Judicial Center.

Galanter, Marc. 1983. "The Emergence of Judges as Mediators in Civil Cases, 1930-1980." Paper presented at meetings of the Law & Society Association, Denver Marriott Hotel, Denver, Colorado, June 2-5.

Hensler, Deborah R., Albert J. Lipson, and Elizabeth S. Rolph. 1981. Judicial Arbitration in California: The First Year. Santa Monica, California: The Rand Corporation.

Johnson, Earl, Jr. with Elizabeth Schwartz. 1978. A Preliminary Analysis of Alternative Strategies for Processing Civil Disputes. Washington: National Institute of Law Enforcement and Criminal Justice.

Kritzer, Herbert M. 1982. "The Judge's Role in Pretrial Case Processing: Assessing the Need for Change." 68 Judicature 28.

_____. 1980-81. "Studying Disputes: Learning from the CLRP Experience." 15 Law & Society Reveiw 503.

Kritzer, Herbert M. and Jill K. Anderson. 1983. "The Arbitration Alternative: A Comparative Analysis of Case Processing Time, Disposition Mode, and Cost in the American Arbitration Association and the Courts." 8 Justice System Journal 6.

Lind, E. Allan and John E. Shapard. 1981. Evaluation of Court-annexed Arbitration in Three Federal District Courts. Washington: Federal Judicial Center.

Miller, Richard E. and Austin Sarat. 1980-81. "Grievances, Claims, and Disputes: Assessing the Adversary Culture." 15 Law & Society Review 525.

Nardulli, Peter F. 1978. The Courtroom Elite: An Organizational Perspective on Criminal Justice. Cambridge: Ballinger Publishing Company.

Nelken, David. 1981. "The 'Gap Problem' in the Sociology of Law: A Theoretical Review." Windsor Yearbook of Access to Justice 35.

Peterson, Mark A. and George L. Priest. 1982. The Civil Jury: Trends in Trials and Verdicts, Cook County, Illinois, 1960-1979. Santa Monica: The Rand Corporation.

Rosenberg, Maurice, Peter F. Rient, and Thomas D. Rowe, Jr. 1981. "Expenses: The Roadblock to Justice." 20 (3) The Judges' Journal 16.

Rowe, Thomas D., Jr. 1982a. "Attorney Fee Arrangements and Dispute Resolution." Paper prepared for the National Conference on The Lawyer's Changing Role In Resolving Disputes, Harvard Law School, Cambridge, Massachusetts, October 14-16.

_____. 1982b. "The Legal Theory of Attorney Fee Shifting: A Critical Overview." 1982 Duke Law Journal 651.

Sander, Frank E.A. 1982. "Alternatives to the Courts." Paper prepared for the National Conference on The Lawyer's Changing Role In Resolving Disputes. Harvard Law School, Cambridge, Massachusetts, October 14-16.

Sarat, Austin. 1983. "Injustice in the Name of the Law." Plenary address at the Annual Meeting of the Law & Society Association, Denver Marriott Hotel, Denver, Colorado, June 2-5.

_____. 1981. "The Role of Courts and the Logic of Court Reform: Notes on the Justice Department's Approach to Improving Justice." 64 Judicature 300.

Trubek, David M., Austin Sarat, William L.F. Felstiner, Herbert M. Kritzer, Joel B. Grossman. 1983. "The Costs of Ordinary Litigation." 31 UCLA Law Review (forthcoming).

Utz, Pamela. 1978. Settling the Facts: Discretion and Negotiation in Criminal Courts. Lexington, Massachusetts: Lexington Books.

An Examination of U.S. Appellate Court Opinions

Alan E. Gelfand, *The University of Connecticut*[1]
W. Robert Stephenson, Iowa State University

I. Introduction and Statement of the Problem

In a common-law system such as ours, the written, published appellate court opinion is the most important body and source of legal doctrine. While the court is charged to establish whether or not "reversible error" has occurred in a lower court decision, its opinions often supply more than a critical review. These opinions have a directive force based on their acceptance and use as precedents. The reported opinions of appellate courts are what lawyers and judges study so that they may decide, predict, and counsel. Some appellate opinions change the course of commerce, government, and other social activity. Much of Anglo-American jurisprudence has been a series of colloquies about the meaning of appellate opinions.

In many aspects of legal regulations, the appellate opinion becomes the law applicable to widespread geographical areas covering millions of persons and many important matters. This is especially true of the intermediate federal appellate courts, the twelve[2] courts of appeals. Federal appellate courts make the bulk of the law in the federal system and because of diversity jurisdiction may have some fair impact on state law.

The federal Supreme Court has certiorari jurisdiction--the power to decide not to review a case--and, by granting certiorari in less than two percent of the petitions presented, actually provides full written opinions of only some 150 to 175 cases a year. Thus the importance of the courts of appeals to lawyers and litigants and to those concerned with legal regulation may, in many areas of law and for many purposes, be greater than that of the Supreme Court. Litigants, lawyers, and the federal appellate judges know that the probability of reversal or even of a case being heard by the Supreme Court is so slight as to be insignificant. The courts of appeals, however, lacking certiorari jurisdiction, must decide every appeal. In 1960 these courts decided fewer than 4,000 appeals per year; by 1977 that number had reached approximately 20,000.

Table 1, derived from data in the annual reports of the administrative office of the U.S. courts, details this dramatic increase in case load on a per judge basis for four of the busier circuits. It is even more striking when one notes that in each of these circuits the number of judges has increased over this period as well; the total number of judges rose by more than a third, from 61 to 82.

Circuit	1962	1970	1973	1976
2nd	61	131	162	216
4th	52	161	239	191
5th	65	126	191	210
9th	52	117	165	198

Table 1. Cases Terminated Per Judge

The appellate courts have the same basic task as that of the Supreme Court, which is (1) to effectively decide controversies that have sharply divided legislators, lawyers, and the public; (2) to understand and elucidate the complexities of issues precisely and persuasively in their expositions; and (3) to give direction to the law. This task cannot be accomplished adequately with a large backlog of cases awaiting appellate review. "The indispensable condition for the discharge of the Court's responsibility is adequate time and ease of mind for research, reflection and consultation in reaching a judgment...."[3]

Some proposals to ameliorate this burden have suggested that the solution lies with the creation of more circuits and/or the expansion of the present circuits. Such growth has been criticized in that it may lead to intercircuit conflict and in that larger courts pose the "danger of losing the quality of collegiality, losing time for conference, time for deliberation, [and] time for the slow maturation of principle."[4]

Much of the difficulty in handling the large number of appeals facing the courts is that the complicated and often "political" process of opinion writing is considered the single most laborious task of federal appellate judges. Perhaps up to thirty percent of all judges' time is devoted to opinions.

Many federal appellate judges believe that there is no need for an opinion at all in a substantial number of cases. "Where in a given case [it] is the

considered judicial judgment of three Judges comprising a panel [that a written opinion is unnecessary], then it is perfectly obvious that the now limited and precious judicial resources can be husbanded by a procedure which eliminates that unnecessary opinion."[5] If the number of written opinions could be reduced, judges could spend more time on the fewer opinions that have to be written. Such decisions might then be more thoroughly reasoned statements of the law. Moreover, as noted above, many judges feel capable of discerning whether a particular case merits a written opinion. Of course, litigants and lawyers who invest time and money in the development of an appeal would feel entitled to a written opinion on that appeal. However, the dramatic increase in the number of appeals suggests the possibility that many routine, perhaps even frivolous, appeals are being brought before the appellate courts. Accordingly, by 1980 all but the Third Circuit had adopted local rules allowing, in certain circumstances, judges to decide cases without a written opinion.

The Fifth Circuit with its Local Rule 21[6] was the first circuit to adopt such a rule, and for a 28-month period between August 1970 and November 1972 it was the only circuit with such a rule.[7] It is the focus of this article to investigate how effectively the Fifth Circuit judges have utilized this rule.

One might envision a variety of controlled experimental designs to study such performance. For example, the judges could have continued to write opinions in all cases, designating, however, those not meriting a written opinion. After a period of time, comparisons could be made. Alternatively, cases could have been randomly assigned to two groups: one in which the rule was operative; the other, not. Again after a period of time, comparisons could be made. The latter design raises constitutional and ethical issues. In any event, the urgency of the problem--an apparently unmanageable workload--led to the implementation of the rule in the absence of such controlled experimentation. Fortunately, no other sharp or sudden operative changes appear to have been made in the federal appellate court system during the aforementioned 28-month period, which allows for a somewhat "controlled" investigation.

Table 2 records the number of opinions disposed of by Rule 21 as well as the total number of dispositions in the Fifth Circuit from 1970 to 1976 as reported by the Clerk of the Fifth Circuit. After utilization of the local rule peaked in 1974, it has since 1976 been applied to between 25 and 30 percent of cases. This may not imply that the proportion of nonmeritorious opinions has declined. Rather, it may be a response to the vociferous complaints of affected attorneys and litigants. Table 3 shows that, as a result of Rule 21, the opinion-writing load of the Fifth Circuit judges on a per judge basis has been brought to a level comparable with that of all other circuits (AOC).

II. Data

Six different samples labeled A through F of sizes indicated in Table 4 were drawn at random from the *Federal Reporter*.[8] The earliest samples, A and D, were drawn nearly ten years prior to the implementation of Rule 21, while the second pair, B and E, were drawn during the period shortly prior. This sampling enables comparison of case mix and decision mix over this decade. The third samples, C and F, were drawn during the 28-month period when only the Fifth Circuit had a discretionary rule in effect.

Table 5 reports the mix of cases and decisions encountered in the six samples. With regard to decision mix, we see that the frequency of affirmations has remained quite stable over the samples. With regard to case mix, there are differences. The Fifth Circuit samples show more civil cases and fewer administrative and criminal cases than the other circuits taken together (AOC). However, within the Fifth Circuit data, if samples A and B are joined and compared against sample C (see Table 6), there has been no significant change in case mix. Of separate interest is the very strong relationship between case mix and decision mix.

We next turn to the "physical" characteristics of the written opinions. For each case sampled, we recorded the length of the opinion, measured by the number of columns in the *Federal Reporter*. We also recorded the number of citations in each opinion, defined for this study to include all full-case references whether in a "string cite" of introduced by any kind of signal and whether in the text of the opinion or in a footnote. Table 7 presents the total incidence of and average value of these characteristics for samples B, C, E, and F. Examination of these samples suggests that, for a given case, the number of columns of opinion and the number of citations per opinion both follow a Poisson distribution. Under such a model we find that, for the Fifth Circuit, the post-Rule 21 sample, C, has significantly longer opinions with more citations per opinion than the pre-Rule 21 sample, B. In the AOC samples, the reverse is true. A possible inference from these comparisons is that Rule 21 allowed the Fifth Circuit

judges more time to write longer and more detailed opinions while the judges in the other circuits in response to increased loads were writing shorter, less detailed opinions.

III. How Good Were the Judges?

We may finally address the question of how effective the judges were in selecting cases to be affirmed without written opinion by Local Rule 21. Recall the text of Rule 21.[9] In order for it to be applicable it must be the case that "[t]he Court . . . determines that an opinion [written] would have no precedential value." Taking this as the performance criterion, how may we assess the ability of the judges to determine nonmeritorious cases? To do so requires specification of how to measure the precedential value of a case.

With regard to judicial opinions, given the assumption that "cited" is equivalent to "used," we think precedential value means at least that the cited opinion had an effect on later citing opinions. It does not seem an unwarranted hazard to assume that the frequency of citation to an opinion reflects the influence--the precedential importance--of that opinion.

The dispute whether citation counts are a useful measure of quality continues in the natural and social sciences. Some critics take an "absolute" view, disputing the theory that the importance of a published paper--considered here as roughly equivalent to a reported case--may be measured by the number of citations to it. This often means that only in a retrospect of historical dimensions can the importance of a paper be assayed. But the more "social" measure of quality, which also does not require the consensus of historians of a general discipline or a subspecialty, is that of citation frequency. This latter method is based on the premise that if those whose business it is to use papers do in fact use some more often than others, that use is largely due to their greater importance. Thus, precedential value is regarded as meaning, even necessarily implying, that a set of reported opinions having that characteristic is more useful to jurists and will be cited more often.

There do not appear to be any "landmark" cases amongst those sampled, i.e., cases that foreclose issues, thus resulting in no later litigation on these issues and rendering later citation unnecessary. Similarly, an appellate opinion heard by the federal Supreme Court might, by such review, have its impact terminated or be given greater force according to whether the higher court reversed or affirmed the opinion. We would anticipate an inconsequential number of sampled opinions to be heard by the Supreme Court.

Therefore, utilizing *Shepard's Federal Citations*, we recorded the number of subsequent cases that cited each sampled opinion.[10] For each subsequent citation, we also recorded how long after the date of the sampled opinion this citation occurred. This latter data is of particular use in characterizing the active lifetime of an appellate opinion with respect to subsequent citation.[11]

It may be argued that such counting fails to gauge the quality of an individual reference, that each subsequent citation receives equal weight. We would respond by arguing that frequency of use is assumed to measure importance and that variability in the nature of subsequent citations would average out.

It may be further argued that opinions on complicated or unique cases having unusual associated facts might be disadvantaged by our measure of precedential value. That is, if "string-citing" is becoming an easier and more common practice, it may, in more straightforward cases, encourage perfunctory citation of many opinions where only a small subset truly establish precedent. We would argue that during the period of observation this may be less of an issue since prior to 1976 there were few automated opinion retrieval systems available, which likely made extensive string-citing less frequent. Moreover, in a comparative sense, there is no reason to suggest that such complicated or unique cases are occurring with different relative frequency across our samples, so again any such disadvantage should average out.

Experimental modification relevant to the above arguments and worthy of further investigation would be to count a subsequent citation only if it quotes text directly from the sampled opinion or at least if it refers to the sampled opinion in its own text rather than footnotes.

For another argument against our frequency count, one might conclude that, with an increasing number of written opinions over time, an important case from a later sample is apt to receive more citations over the same amount of time than an equally important case from an earlier sample. The same argument, however, suggests that over time an increasing number of important cases will arise, for there seems to be no reason to expect a limit on the number of important opinions. Moreover, in our samples we have seen relatively little change in either the length of opinions or in the number of citations per opinion. Thus,

it is likely that, with more important opinions competing for comparably more citations, no adjustment over time is required.

While some of our sampled opinions (in A and D) were available for as much as 15 years of subsequent study, others (in C and F) allowed fewer than five years. To render comparisons fair, we utilize for all sampled opinions only the first 48 months following the opinion. Unfortunately, during this interval we may anticipate seeing only a third of subsequent citations.[12] However, in a comparative sense, this should create no problem.

Table 8 presents what appears to be a disconcerting picture. The Fifth Circuit judges seem to be doing worse than would be expected under random application of Rule 21. In fact, the difference between the post- and pre-Rule 21 samples, while perverse in direction, is not statistically significant. But the more important issues are the following:

(1) Why is a cited/not cited dichotomy appropriate? Since we argue that frequency of subsequent citation measures importance, why not a dichotomy into, e.g., three or less/four or more? Indeed, why is any simple dichotomy appropriate? Why not look at the distribution of subsequent citation in more detail?

(2) Local Rule 21 is applicable to affirmations only. Therefore, only sampled affirmations ought to be compared. By looking at all cases in aggregate, might we not be masking sample differences with regard to affirmations?

(3) Suppose that application of Rule 21 at a rate of at most 30 percent implies that the Fifth Circuit judges are voluntarily underutilizing it. If, in addition, the incidence of routine appeals after Rule 21 runs at a higher rate than that prior to Rule 21, the effectiveness of the judges would not be expected to be revealed in such a tabulation.

In Tables 9-12 we respond to issues 1 and 2. Whether issue 3 is of consequence cannot be evaluated under our experiment and in any event seems difficult to assess.

Table 9 addresses issue 1 providing a more detailed distribution of subsequent citation for the Fifth Circuit samples. The limitations of Table 8 are revealed.[13] In fact the post-Rule 21 sample seems to be receiving more frequent subsequent citation. Again the differences are not statistically significant.

Table 10 returns to the dichotomy of Table 8, but in response to issue 2 separates affirmations and nonaffirmations. The size of the affirmation samples within C and F were each increased by 50 after the original sampling to provide adequate sample size for significance testing. This results in larger total C and F samples with different aggregate figures for "percent cited within 48 months" than those in Table 8. The Fifth Circuit affirmations in the post-Rule 21 sample show an increase in the proportion cited within 48 months compared with the pre-Rule 21 sample. Again this difference is not significant. We have no explanation for the dramatic decline (16.3%) in proportion cited within 48 months for the Fifth Circuit in nonaffirmations (C vs. A and B).

Table 11 considers issues 1 and 2 together. Using the categories in Table 9, we see that the post-Rule 21 affirmations in the Fifth Circuit record significantly more subsequent citations as measured by (1) comparison of mean number of subsequent citations under a Poisson arrival model, (2) comparison of the "three or less" with "four or more" groups as a 2x2 table, and (3) comparison of the incidence of heavily cited ($\geq$ 10) cases, nearly twice as many arising in the post-Rule 21 sample.

Lastly, in Table 12, using the other circuits (where, again, no discretionary rule was operative during the "post" sampling period) as a control, under the identical categories as in Tables 9 and 11, no differences are seen between the pre- and post-Rule 21 samples.

We are thus led to conclude that, under our assumptions, one can demonstrate effective selection by the Fifth Circuit judges of appeals meriting written opinion; Rule 21 seems to be enabling reduced opinion-writing loads while fulfilling its judicial intent.

	1971	1972	1973	1974	1975	1976
Rule 21's	210	488	629	633	764	664
Total dispositions	1748	1877	2092	1763	2244	2181
Percent	12.0	26.0	30.1	35.9	34.0	30.4

Table 2. Incidence of Rule 21 in the Fifth Circuit

	1970	1971	1972	1973	1974	1975	1976
Number of cases disposed of after hearing and submission/judge AOC	57.2	71.4	81.2	91.8	81.6	83.3	87.4
Number of cases disposed of after hearing and submission/judge 5th Circuit	96.4	110.7	121.7	133.9	121.1	133.4	128.6
Number of cases adjusted for Rule 21	96.4	96.7	89.1	91.9	78.9	82.5	84.3

Table 3. Comparison of dispositions between circuits adjusted for Rule 21

	Time Periods		
	1960-62	1967-70	1971-72
5th Circuit	100 (A)	101 (B)	99 (C)
Other Circuits	100 (D)	295 (E)	108 (F)

Table 4. The sampling frame

		Criminal	U.S. Civil	Private Civil	Administrative	Bankruptcy	Other	Totals
A	Affirmed, etc.	12	14	27	7	3	0	63
	Reversed, etc.	4	9	13	5	1	0	32
	Dismissed	0	1	0	0	0	0	1
	Other	0	1	2	1	0	0	4
	Totals	16	25	42	13	4	0	100
B	Affirmed, etc.	11	10	27	10	1	0	59
	Reversed, etc.	6	3	20	0	0	0	29
	Dismissed	0	1	0	0	0	0	1
	Other	0	1	8	2	1	0	12
	Totals	17	15	55	12	2	0	101
C	Affirmed, etc.	17	7	28	6	0	0	58
	Reversed, etc.	4	11	15	5	0	0	35
	Dismissed	0	0	0	0	0	1	1
	Other	0	2	3	0	0	0	5
	Totals	21	20	46	11	0	1	99
D	Affirmed, etc.	9	14	25	8	1	0	57
	Reversed, etc.	4	5	11	4	2	0	26
	Dismissed	0	0	4	1	0	0	5
	Other	2	0	7	3	0	0	12
	Totals	15	19	47	16	3	0	100
E	Affirmed, etc.	46	26	76	35	4	0	187
	Reversed, etc.	15	10	37	10	2	1	75
	Dismissed	0	1	0	3	0	0	4
	Other	4	4	14	7	0	0	29
	Totals	65	41	127	55	6	1	295
F	Affirmed, etc.	26	8	24	15	0	0	73
	Reversed, etc.	3	4	12	5	0	0	24
	Dismissed	2	0	1	0	0	0	3
	Other	3	2	3	0	0	0	8
	Totals	34	14	40	20	0	0	108

Table 5. Classification by decision and case type for the six samples

	Criminal	U.S. Civil	Private Civil	Admin-istrative	Other
A,B (201)	33 (16.4)	40 (19.9)	97 (48.3)	25 (12.4)	6 (3.0)
C (99)	21 (21.2)	20 (20.2)	46 (46.5)	11 (11.1)	1 (1.0)

Table 6. A comparison of case mix for the Fifth Circuit samples

		Number (average number) columns of opinion	Number (average number) citations per opinion	Average citations / Average columns
(FC)	B (101)	559 (5.53)	1221 (12.09)	2.18
	C (99)	631 (6.37)	1308 (13.21)	2.07
(AOC)	E (295)	1923 (6.41)	3483 (11.61)	1.81
	F (108)	622 (5.76)	1036 (9.59)	1.67

Table 7. "Physical" characteristics of appellate court opinions

5th Circuit	Percent cited within 48 months	Others	Percent cited within 48 months
Pre-Rule 21 (A,B) (201)	77.6	Pre-Rule 21 (D,E) (395)	79.5
Post-Rule 21 (C) (99)	70.7	Post-Rule 21 (F) (108)	79.6

Table 8. All cases--citation within 48 months

	0	1-3	4-6	7-9	≥ 10	Mean number of citations
A & B (201)	45 (22.4)	86 (42.8)	29 (14.4)	19 (9.4)	22 (10.9)	4.4
	65.2		34.7			
C (99)	29 (29.3)	27 (27.3)	21 (21.2)	8 (8.1)	14 (14.1)	5.0
	56.6		43.4			

Table 9. Fifth Circuit written opinions--subsequent citation within 48 months

Fifth Circuit	Percent cited within 48 months	Others	Percent cited within 48 months
Affirmations (A,B)	71.7	Affirmations (D,E)	74.3
Nonaffirmations (A,B)	85.2	Nonaffirmations (D,E)	86.4
All (A,B)	77.6	All (D,E)	79.5
Affirmations (C)	76.0	Affirmations (F)	76.5
Nonaffirmations (C)	68.9	Nonaffirmations (F)	79.5
All (C)	73.8	All (F)	77.2

Table 10. Subsequent citation--affirmations and nonaffirmations

	0	1-3	4-6	7-9	≥ 10	Mean number of citations
A & B (113)	32 (28.3)	51 (45.1)	13 (11.5)	7 (6.2)	10 (8.8)	3.4
	73.4		26.6			
C (54)	15 (27.8)	13 (24.1)	12 (22.2)	5 (9.3)	9 (16.7)	5.2
	51.9		48.1			

Table 11. Fifth Circuit affirmations with written opinion--subsequence citation within 48 months

	0	1-3	4-6	7-9	≥ 10	Mean number of citations
D,E (395)	70 (17.7)	151 (38.2)	71 (18.0)	36 (9.1)	67 (17.0)	5.3
F (108)	21 (19.4)	42 (38.9)	16 (14.8)	11 (10.2)	18 (16.7)	5.4

Table 12. Other circuits' written opinions--subsequent citation within 48 months

Footnotes

1. This research owes its nascence to Philip Shuchman, Professor of Law, Rutgers University, who collaborated with the first author on the paper, "The Use of Local Rule 21 in the Fifth Circuit: Can Judges Select Cases of 'No Precedential Value'?" *Emory Law Journal* (1) 29 (1980), 195-230, from which much of this article was drawn.

2. At the time the data for this research was gathered, there were ten circuit courts in addition to the D.C. circuit. Subsequently, the Fifth Circuit, which will be the focus in the sequel, was split into two circuits.

3. Report of the Study Group on the Caseload of the Supreme Court, 57 FRD 573, 606 (1972).

4. Bork, Dealing with the Overload in Article III Courts, 70 FRD 231, 233 (1976).

5. NLRB v. Clothing Workers Local 990, 430 F.2d 966, 971 (5th Cir. 1970).

6. Local Rule 21 provides: When the Court determines that any one or more of the following circumstances exists and is dispositive of a matter submitted to the Court for decision: (1) that a judgment of the district court is based on findings of fact which are not clearly erroneous; (2) that the evidence in support of a jury verdict is not insufficient; (3) that the order of an administrative agency is supported by substantial evidence on the record as a whole; and the Court also determines that no error of law appears and an opinion would have no precedential value, the judgment or order may be affirmed or enforced without opinion. In such case, the Court may in its discretion enter either of the following orders: "AFFIRMED. See Local Rule 21," or "ENFORCED. See Local Rule 21." 5th Cir. R. 21.

7. On November13, 1972, the Tenth Circuit adopted its Local Rule 17, which is nearly identical to the Fifth Circuit's Local Rule 21.

8. The following volumes of the *Federal Reporter*, Second Series were used: For samples A and D, volumes 282, 285, 287, 290, 293, 297, 300; samples B and E, volumes 371, 379, 385, 392, 399, 407, 415, 424; samples C and F, volumes 433, 441, 448, 455, 468.

9. See note 6, *supra*.

10. The samples contained too few subsequent state court citations to impact on our findings.

11. See Shuchman and Gelfand, note 1, *supra*, for a brief discussion. Gelfand (1981), "Am Estimation Problem for Poisson Processes," *Australian Journal of Statistics* 23 (2):1-8 contains a formal mathematical formulation and solution for this sort of estimation problem illustrated with the sampled data. Gelfand and Stephenson (1979), "A Model for Precedential Value of Appellate Court Decisions," *Proceedings* of the Social Statistics Section, 1979 ASA Meetings, pp. 559-563, provides preliminary discussion on the modeling of the active lifetime of a U.S. appellate court opinion.

12. See discussion in references given in note 11, *supra*.

13. The number of and definition of categories is arbitrary and a similar picture would emerge under other choices.

Data Analysis Problems at the State Level: Data Problems and Data Collection

Michael H. Rabasca, *Delaware Statistical Analysis Center*

I think the best way to begin to describe data problems and data collection for the state-level practitioner is to recall for you the first time state officials came to me for assistance in early 1977. The Statistical Analysis Center (SAC) was newly formed, and I was engaged in trying to identify both data sources and potential users. There was considerable turmoil in the prison system. Overcrowding was the issue. It still is. Someone suggested that prison population projections might be useful in promoting what was being heralded as the Corrections Master Plan. My assignment: project the prison population to the year 2000. My data base? End-of-year populations for 1975 and 1976--two years of data to project twenty-three years into the future. I indicated that if I could get that two years converted into twenty-four data points I might take a shot at projecting next month's population. That was O.K., they said; they had already decided to build a 360 bed prison. All they wanted were some statistics to support the decision.

It was then that I learned the first law for criminal justice statistics--give them what they want or they'll stop asking. They got a projection; the prison was built, is currently full, and double bunking begins shortly.

In a nutshell that means, first and foremost, that the criminal justice system is a political being and that it is policy, not numbers, not crime, not arrests, not populations, that drives the system, and any agency that wishes to meet the primary goal of a bureaucracy--survival--must recognize that fact. Establishing credibility is the only way to survive. It's the best friend a Statistical Analysis Center can have.

You probably noticed that I didn't mention the time-honored tenets of Reliability and Validity It is not that these are not important; they most assuredly are; however, they often do not exist. Credibility comes from making logical presentations, sometimes without statistical proofs, that customers, whoever they are, will accept. That might be called lucking out at the 95% level. It also comes from meeting head-on the problems of data collection and analysis which plague us at the operating level.

There are several very formidable constraints which make the classic scientific approach a difficult, if not impossible goal.

Not the least of these is <u>time</u>. The ability to design and conduct a specialized research project or experiment over an extended period of time is a luxury which most of us must forego.

It is an extraordinarily rare event to receive a request for data or data analysis that is part of a long-range planning process. Normally, the requester wants it immediately and, preferably, over the phone.

We find ourselves lacking in <u>resources</u>, both people and funds. As I have just said, a sizable portion of our requests from other agencies for data analysis is knee-jerk, reactive to an immediate issue. Data collection, as you are aware, costs money and uses people. Planning and committing funds a year in advance for ad hoc requests that may never come is a tough concept to sell to budget-writing committees. As it turns out, we are, more often than not, forced to use data in the form in which it is produced by the agency that collects it. Data collection at the criminal justice system operating level is not always a very sophisticated process. Each agency collects data that measures its activity and supports its continued survival. The thought that data would be captured to support some other part of the system is a new idea. Even with the advent of computerized record keeping and data collection we find that better data is not available. Manual tasks are simply converted to computerized tasks.

Next, we have the data itself, or themselves, whichever you prefer. It comes in many forms--normally non-normal, non-existent, erroneous, missing, fabricated, discontinued, illogical, non-transactional, and ignored.

The fourth constraint is the people from whom we collect this data. I don't recall who said, "You can't reason someone out of something they were not reasoned into," but he was right. The problem of collecting data, reporting data, and analyzing data must be dealt with in the light of the parochial viewpoint of those from whom we collected it and the bureaucratic paranoia that somehow you will use the data against them or that someone else might use it if they get their hands on it or that you might have something to say about it.

At this point I'm reminded of the story of the traveling salesman driving along a country road. Near a farmhouse he spots a pig hobbling around on a wooden leg. Fascinated by this sight, he stops and inquires as to why a pig might have a wooden leg. The farmer replies, "We all love that pig. Why, one day on the back forty, I hit a stump with my tractor. The tractor tipped over and pinned me underneath. That pig ran home and got help. He saved my life. Then one night the house caught fire and that pig ran from room to room wakin' everyone up and leadin' them out to safety. He's a fine pig."

"But, why the wooden leg?" asks the salesman

"We love that pig," replies the farmer. "You don't eat a pig like that all at once."

Well, the criminal justice system is a pig that, like it or not, some of us must eat all at once. The option of selecting just one part is not available to those of us who find ourselves in the criminal justice system information business. We have to deal with the whole, and focusing time and resources to correct data collection and reporting problems and bureaucratic inertia at one point necessarily neglects another. Not that we don't do this; we do. It's a risky enterprise. It causes us to narrow our system perspective, and if we can't provide a broader picture of what's happening in the system than the system components themselves, we're out of business. Some political savvy and finger-tip access to data, research findings, and informed opinion on the criminal justice issues are assets as important to the state-level analyst as are statistical skills.

I've dealt with some generalities; let me be a little more specific about some of my problems in system-wide data collection and analysis.

Everyone's bread and butter seems to be the Uniform Crime Report (UCR). In Delaware, responsibility for the collection and reporting of data belongs to the State Bureau of Identificiation (SBI), part of the State Police. I might point out here that Delaware has only three counties (only one of which has a police force), a corrections and judicial system operated only by the state, and no jails. Everyone goes to prison. The only criminal justice agencies below state level are local police departments and a municipal court and prosecutor's office in our largest city.

Through agreement with the State Police, we handle all UCR data requests outside the police community. The UCR is the only collective measure of crime and response to crime that we have. We aren't part of the National Crime Survey and simply don't have the capability to conduct our own.

Now for the problems. The UCR is the classic case of the tail wagging the dog. Data reported to the State Bureau (the SBI) is converted to data that meets FBI reporting requirements and then into our computerized incident-based system. The data then immediatly loses its identity as Delaware crime. For example, Delaware had no crime called "aggravated assault", which is one of the seven Index Crimes. Assaults in Delaware are by degree, first and second being felonies, third degree a misdemeanor. Assaults are not differentiated by weapons use and only partially by physical injury. Differentiation in degree is by state of mind of the accused. Thus, first, the charge and, second, the UCR classification are highly judgmental.

In 1978, the State Bureau Director unilaterally elected to record all domestic quarrels as aggravated assaults in order to highlight the problem. He succeeded. We thought a roving band of maniacs was loose in the state. The problem, once discovered, was that no one had the time or resources to go back to the source documents and change the classification. Besides, they were reported, and they are official. The outcome of the UCR classification is that there exists a rate of aggravated assault for Delaware that is comparable to rates of aggravated assault in other states but that is essentially erroneous and meaningless within the state.

Speaking of comparable rates, we all know that the incidence of crime is related to a myriad of agreed - and disagreed - upon factors and that comparing the crime rates of various jurisdictions or the frequency of crime in jurisdictions with similar populations can be an invalid and unfair comparison. But you simply cannot publish the crime rates of City A and City B or State C and State D without inviting comparisons, particularly in the press. Did you know that the <u>reported</u> crime rate of Manchester, New Hampshire was higher in 1981 than that of Chicago, Illinois? Now where would you rather walk down the street at night? The point is that if an issue were made that Manchester is a new U.S. crime center, you can bet the rate will come down. People are sensitive to those things that are official measures of their success or failure. If the security of a police chief's job is dependent on the crime rate, it'll be low -- until budget time.

Back to the tail wagging the dog. UCR's divide larcenies into classifications of less than $50, $50 to $250, and over $250. In Delaware the dividing line between grand and petty theft is $300. Now, I can get the data given the time and money. But, in the meantime, the state is routinely collecting, and distributing to every agency in the state, data that is produced to meet FBI rather than local needs.

The one last word about the reporting and recording of crime data that has serious implications on reliability and validity is the universal lack of auditing. No one seems to have the wherewithal to tackle it. As a result, when recording the map-grid location of a crime, police officers occasionally make one up or use the same one over and over, one they have memorized. Consequently, Delaware crimes might be reported in grids located somewhere in Iceland, or we discover a major crime wave in someone's backyard. It appears that magistrates are committing persons to prison for crimes over which they lack jurisdiction. They are not doing it, it is simply that the crime codes are in error. Persons disappear from sentenced-inmate reports one month, only to mysteriously reappear the next.

The relative severity of crime is something that most researchers and practitioners have dealt with. All of our data about offenses and offenders is collected according to the most serious offense. Most offenders who enter the system have more than one charge against them. If one charge is murder and the other is overparking, there is no problem. Problems occur when the differences are subtle and when procedures for determining the most serious charge are non-standard. If two offenses are in the same felony class, which is more serious? Is a crime that requires mandatory incarceration more serious than one that does not? In the absence of

agreement, we can only assume that the most serious is reported. But when someone is imprisoned for contempt of court, violation of probation or parole, resisting arrest, or the offense of miscellaneous, of which we have several per year, we simply don't know if these offenses were more serious than the one that brought him to court in the first place. Incidentally, those commitments account for 25% of all admissions. For a very long time, corrections input operators were selecting the offense with the lowest code number as the most serious when in reality those who originally designated the codes had grouped the crimes by type of offense without regard to severity or statutory classification.

Similar to these data collection problems are the internal decisions of criminal justice agencies that supply us data, decisions to change or discontinue the data collection process.

One of the principal activities we have been involved in is sentencing reform. I'm almost convinced it's an impossible task. We collected five years of sentencing data that was produced annually by the staff of the State Superior Court. We developed a methodology, using standard scores, that allowed us to evaluate the average performance of each judge over six measures of possible disparity. We had to use aggregated data simply because of the time and resources problem again. With a five-year data base in hand we set out to compare 1981 and 1982 to it. As you might guess, as a cost-saving measure, the data was no longer being aggregated. The court had no need for the data in that manner. I, of course, was welcome to go through the 8,000 court dockets manually. Sampling is an obvious and reasonable suggestion. However, consider that there are eleven judges and several hundred types of crimes. There would hardly be a sample large enough for each cell. The only answer is to deal with selected crimes, which we are doing, to the detriment of the Reform Commission's timetable.

We are constantly looking for clues about what is happening to prison populations. By the time input and turnover information is collected, however, major policy decisions have occurred that change everything. Looking over eighty-nine months of population data, we, through serendipity, discovered what appeared to be a cyclical pattern. Using time-series analysis and classical economic-decomposition theory, we described fairly regular periods of alternating growth and stability. The pattern indicated that the current level of growth should shift to a period of stability in October 1982. Naturally, in October 1982, Corrections changed its reporting format and is currently months behind and is reporting the data differently. So it may be some time before we can learn whether we could describe what might be the underlying cause of the cycle, if we, indeed, even find that the cycle continues.

The whole business of projections is enough to challenge sanity. Delaware's response to the efforts of Mother's Against Drunk Driving was to enact a very tough driving-under-the-influence law. This law was the product of a task force that among other things tried to project the impact of the new law. Would increased sanctions deter drinking? Would new emphasis result in more arrests? longer incarceration? Unbeknownst to anyone, the State Police had devised and kept secret a road-block plan that they launched immediately after enactment. Arrests skyrocketed and, suddenly, driving-under-the-influence is the single greatest cause of incarceration in the State.

I would like to take a moment to tell you how valuable the research performed by people like yourselves is to people like me. As I indicated earlier we have not the time, the resources, or, especially in my case, the talent to dig into some of the research questions you undertake. As one who often must provide some analytical insights about what appeared in this morning's newspaper, I can respond to those who ask, having the benefit of your research and wisdom. The news reporter won't wait for me to perform a research project on why crime is up or down. He has a deadline. If I can't meet it, there is no story--no story, no exposure--no exposure, eventually no SAC. The same holds true when I talk to a legislator who is looking for both research and statistical support for his proposed bill. He needs it today. As a voracious reader of what the research community finds, I can at least articulate some of the major considerations and relate them to Delaware data.

Some words about the politics of criminal justice: after seven years of collecting, analyzing, and disseminating data in the system, I read of the decision to build a new state prison in the morning paper two months ago. That was a political decision. The announcement and the size of the prison were dictated by the political environment, not on a projection of future needs, because we don't know our future needs, and even if we did, then size would still be based on what the political market would bear.

I respect decisions like those. They reflect the reality of the system. Living with them makes one live, and think about data, in the world as it is.

I'd like to give you another example of a political response to a criminal justice issue. It was proposed in the Delaware Legislature that the crime of rape in the first degree be punished by mandatory life imprisonment without parole. Rape was (and is) increasing, and this was thought to be the logical deterrent. Now it just so happens that the existing penalty was also mandatory life but with parole. First eligibility comes in about thirteen years.

Opponents of the change noted that the new penalty was the same as that for murder, so, they argued, rapists would murder their victims after raping them since they had nothing to lose. Proponents, of course, touted the deterrent effect.

We were asked to provide a legislative impact statement. We discovered first, that parole currently occurred in less than one in four rape cases and, second, that 85% of all rapes were committed by persons under age thirty. Our projection showed that only three persons currently in prison would be under thirty if paroled at the one in four rate. We also pointed out that it

seemed unlikely that a person who would currently commit rape and risk life in prison with a less than one in four chance of parole would be deterred from the crime because the chance was decreased to zero in four. We added that the cost of keeping all persons for life despite the very low probability of recidivism exceeded a half-million dollars additionally per non-parolable inmate. The ultimate outcome had to be a political compromise. The bill was amended to prohibit parole for the first twenty years, and it passed like a Flag Day resolution. It, of course, has had no measurable impact on the occurrence of rape, but that wasn't the real purpose. The real purpose was to show the constituency a get-tough-on-crime posture.

My final comment deals with our audience. My job is to service that audience, the public, the media, the private agencies, the criminal justice agencies, and the policymakers. They are very naive about statistics. They are interested in frequency, how much more or less than a year ago. They tend to reach their own conclusions about data. I could give them a mountain of statistical proofs, but I don't, because they won't read them. I tell them what the data tells me; they respect it but may not believe it. A great deal of legislation and criminal justice policy is the result of conventional wisdom and gut feelings. Data must be presented to the audience in a manner they'll understand and in a form they will accept. There is a fine line between keeping people informed and bombarding them with unsolicited and untimely information. You know you've arrived when they come to you. If I do not respond, they will stop asking. When there is a perception that there is no further need for the service, we will cease to exist. So the bottom line is, whatever the data, it's all we have and we'll use it. Improvements come slowly because analysis for planning, which is really what we ought to deal in, is not perceived to be as important as operational data. Fighting fires and crisis management are a large part of daily life in the criminal justice system. We in the data business have to be prepared to respond to those needs. That scientific methodology and data cleaning sometimes suffer goes without saying.

Lest I be accused of being insensitive to the principles of management, I am very aware that Delaware has developed a bottom-up criminal justice information system. We have numerous applications at the operating level. It's when tactical or strategic questions are asked that we have very disparate data. The data problem is not always the absence of data but sometimes too much data that is irrelevant. We tend to collect that which is easiest to collect whether or not we really need it. To paraphrase Russell Ackoff, the noted management thinker, "Wisdom is the ability to control what you can and not fret over what you can't."

Policy Relevance in Criminal Justice Research

Daryl Fischer
Paul Stageberg
Statistical Analysis Center, Office for Planning and Programming

Introduction

One of the major challenges in social science research is to make analytic products policy relevant. Despite good intentions, data presentations frequently end up being too general or too superficial to directly inform policy deliberations on complicated issues. In other cases, researchers "overkill" the problem by writing reports that only their colleagues can read. As a result, much of the material generated by social science research organizations, both public and private, lies unused in large stacks "out of sight" and "out of mind."

At the Iowa Statistical Analysis Center, we've become acutely aware of this problem and have adopted a hard and fast rule to guide our research and report writing efforts: unless a research project can potentially and realistically have an impact on the operation of criminal justice or on important policy deliberations, we won't do it! As a result, most of our projects deal with the needs of the Governor, the Legislature, and operational agencies. Since legislators, administrators, judges, parole board members, and other "non-researchers" have little time or patience to offer in consuming our products, we strive for reports and presentations that are relatively short, hard-hitting, and to the point.

Obviously, as research professionals, we want our products to be sound methodologically and to show a high level of expertise and sophistication. Yet this desire must be carefully balanced and tailored to allow those unfamiliar with research methods to assimilate the material. In this paper, we would like to discuss some of the approaches we have taken in Iowa to identify the important research issues in criminal justice and to conduct the type of quality, policy-relevant research that can inform those issues.

It is our view that the survivors in social science research during this period of funding parsimony will be those who can best balance superior statistical/technical capabilities and a working appreciation for justice system problems and needs. The successful researcher will need to wear a number of hats, as a quality statistician, computer scientist, writer, public speaker, and politician. The days of grace are over for social science research.

Identifying the Important Issues

When the Statistical Analysis Center began operation in early 1978, most of the incoming staff had previously been involved in criminal justice research and planning in other state agencies. As a result, the staff came in with an appreciation for many of the important criminal justice issues facing the state.

One of the major issues in Iowa, as in other states, has been the size of the prison population. Each year, the legislature has debated the construction of new prison facilities and whether other alternatives should be considered, such as changes in sentencing and/or parole laws or an expansion in community-based alternatives to incarceration. As is normal in legislative bodies, the debate has dealt mainly with questions of costs and funding, but issues of public protection and punishment have also played a role. Law enforcement advocates have emphasized the need to protect and punish and to find solutions to the apparently growing crime problem in the state. Others have cautioned against fiscal irresponsibility, while yet a third group has pushed for more humane and effective treatment of offenders in community programs.

Very early on, the Statistical Analysis Center recognized that a major emphasis in its research efforts should be directed to assisting the state in avoiding the unnecessary construction of a new state prison. Both the Governor and the Legislature had expressed displeasure with the thought of the massive expenditures required for such a project. Further, various studies have suggested that rising prison populations may be a temporary phenomenon associated with the "age bulge" of post-WWII babies moving through the crime-prone years.

The approach we took in Iowa to analyze the prison population problem differs in one primary respect from the approaches used by many other states. While many others have accepted current admission and release practices as "givens"--resulting in the development of multi-year prison population forecasts and projections--we have, instead, devoted the majority of our time to analysis of the types of persons sent to prison in Iowa in an effort to identify those who either don't need to be there at all or who could be released in shorter periods of time. Underlying our work is the realization that the size of a state's prison population is policy-driven, the result of decisions determining who needs to be punished, who needs to be rehabilitated, and who needs to be locked up for public protection.

Necessarily, the issue of public protection arises in the context of efforts to avoid prison overcrowding. Reducing the prison population would obviously involve the release of offenders who would otherwise be incarcerated, with the resulting question of a possible threat to society. In a more general context, public protection is an important issue in and of itself. Legislators and releasing authorities are continually looking for ways to provide better public protection from dangerous and recidivistic offenders. Research directed to this end could yield benefits both for the justice system and for society as a whole.

Other issues of importance include the debate over the relative advantages of determinate and indeterminate sentencing; concerns with prosecution/sentencing/parole disparity by sex, race, or jurisdiction; the issue of punishment versus the other goals of sentencing and corrections; the issue of classification of offenders in and out of institutions for custody and supervision-level assignment; and the more general issue of the allocation of resources by correctional agencies--where should the money be spent to get the most bang for the buck?

Parole Guidelines

The Parole Guidelines System offers a prime example of the application of criminal justice research to real-world problems. In early 1981, the Statistical Analysis Center completed the development of a set of guidelines to assist the Iowa Board of Parole in determining how much time individual inmates should serve and when they should be released, in keeping with the goals of public safety, punishment, and other concerns. The system was formulated specifically to help the Board increase paroles and stabilize the prison population while simultaneously reducing violence among parolees.*

The guidelines reflect the results of two separate but coordinated research studies, one of time-served patterns among inmates released on parole in previous years and the second of the factors that best predict success and failure of parolees. In these studies, the Statistical Package for the Social Sciences (SPSS) was used to examine the experiences of several thousand offenders who moved through the adult correctional system in Iowa in the mid-to-late seventies. Several statistical techniques, including regression analysis, configural analysis, and unit weighting (the Burgess technique) were used to identify the best possible predictive combinations of independent variables in the data base.

The first study was directed to the formulation of descriptive guidelines to quantify past Parole Board practices, that is, a mathematical/statistical model of past parole decision-making. This study identified the fact that three factors explain much of the variation in time served among parolees: (1) the length of the inmate's sentence, coupled with the seriousness of the crime for which convicted, (2) the number and seriousness of previous adult felony convictions (emphasis on past prison record), and (3) current institutional misconduct.

Past time-served averages were calculated and displayed as a function of the logical combinations of these factors. This "statistical overview" of past parole decision-making afforded a clear formulation of existing parole policies and practices and thus gave us a "base" from which to estimate the impact of alternative policy modifications, such as increasing term lengths for certain offenders and decreasing them for others.

The second study was directed to an entirely different concern, i.e., the identification of the types of offenders who fail at the highest rates while on parole and/or pose the greatest threat to society after release from prison. The goal of the study was to identify through various "risk-assessment" techniques the characteristics of "high-risk" offenders and thus also of "low-risk" counterparts. The resulting prediction "guidelines" could also be called descriptive in that they describe the past performance of various categories of parolees. However, they are also highly prescriptive in that they provide useful information for the parole board in screening inmates for potential risk to society. It is this prescriptive function that holds the potential benefit both to the Board and ultimately to the state and to the general public in terms of reduced expenditures for corrections and better public protection.

Offender Risk Assessment

The risk assessment portion of the project was completed in mid-1980, following approximately five years of research at a cost of over 3,000 man-hours and $300,000. The final product entailed two predictions, one of recidivism in general and the second of new violence. Using the Mean Cost Rating (MCR), the Coefficient of Predictive Efficiency (CPE), and general percentages of correct prediction, the staff determined that the system was approximately 75% accurate in predicting both violence and general recidivism, establishing the Iowa system as one of the most, if not the most, accurate system in the country.

Following the completion of the recidivism research, the staff devoted the next several months to completing research on time-served patterns and on interrelating the results of the two studies. Essentially, a set of descriptive guidelines of past parole practices was devised and was then modified to reflect the application of risk assessment as a separate "dimension" of the decision to parole. Guideline terms of expected time-to-be-served were modified to allow for longer terms for the highest-risk offender and shorter terms for the lower-risk ones. The new "prescriptive" guidelines were carefully structured to achieve a definite anticipated impact on the prison population, in terms of increased paroles of lower-risk offenders, and on public protection, in terms of reduced recidivism and violence among parolees.

The structure of the current prediction/risk-assessment system is summarized in Figure 1. The process involves five steps, based on an analysis of four types of offender factors: (1) Age at Conviction, (2) Substance Abuse History, (3) Current Offense Classification, and (4) Prior Criminal Record. The emphasis of the system is on the identification of types of recent serious criminal history, coupled with other good predictors of recidivism such as an early age at the current conviction, a serious substance abuse history,

and conviction of an offense that is frequently repeated. We find, then, that the most dangerous or recidivism prone offender is typically a male in his late teens or early twenties who has a history of serious substance abuse, who has had a recent serious prior involvement with the justice system, and who is currently convicted of an offense such as burglary, robbery, motor vehicle theft, forgery, or aggravated assault.

In the new model, there are effectively three levels of risk, labeled Poor Violence Risk, Poor Property Risk, and Good Risk. Recently, SAC staff tested the accuracy of the new model on a sample of 722 ex-prisoners followed for approximately four years each. Figure 2 illustrates the high level of accuracy of the model on this sample, i.e., the Poor Risks constituted 35% of the sample yet accounted for 85% of the total threat of violence. These figures translate roughly into 85-90% accuracy in prediction.

Figure 1

The Iowa Risk Assessment Model

Step I - 4-Factor Scoring

- o Substance Abuse History
- o Current Offense Classification
- o Age at Conviction
- o Total Volume of Past Record

Step II - Offender Typing

- o Violent Offender/Non-Violent Offender
- o First Offender
- o Burn-out Factor

Step III - Special Risk Factors

- o Prior Felony History
- o History of Violence
- o Street Time Since Prior Offenses

Step IV - General Risk Assessment

- o Combination of 4-Factor and Special Factor Results
- o Application of First Offender and Burn-out Classifications

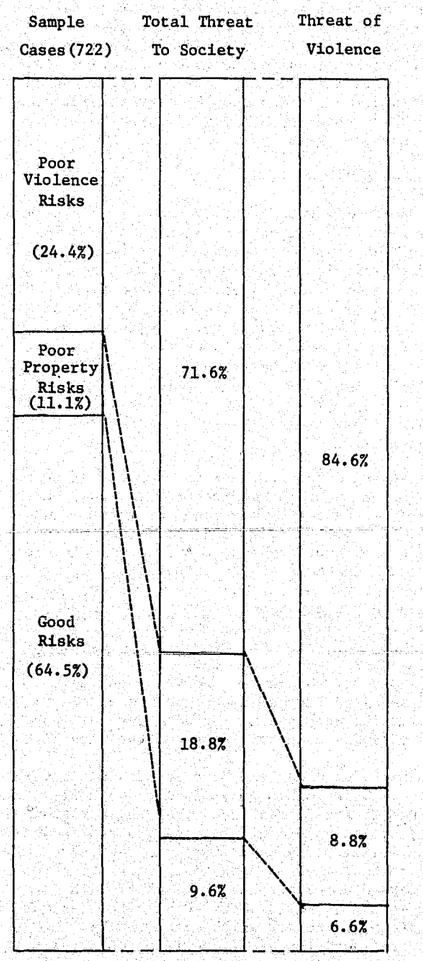
Step V - Violence/Property Risk Assessment

- o Combination of General Risk Assessment and Violent Offender/Non-Violent Offender Classification

Figure 2

The Iowa Risk Assessment Model

Predictive Accuracy



Guideline Implementation and Impact

In early 1981, the Parole Guidelines System was put into final form and was presented to both the Board of Parole and the Iowa General Assembly for consideration to aid in ameliorating the prison population problem at the time. Subsequently, the Board agreed to begin using the guidelines and was given further support for that decision by the passage of H.F. 849, which required the Board to consider statistical predictions of recidivism in making parole decisions. In April, 1981, the system was formally implemented, and since that time (to August, 1983), the Board has reviewed over 4,000 cases in which guideline information was available.

Early in the project, when the Board was still adjusting to the use of the guidelines, and when the SAC staff had yet to fully understand the complexity of the parole decision, the rate of agreement between recommendations and actual parole decisions was in the neighborhood of 40-60%. At the present time, however, we see about a 75-80% rate of agreement. Further, the Board members themselves have expressed the feeling that they've come to depend on the guidelines to help them make good decisions. It would appear that the "prescriptive" nature of the guidelines will soon become more "descriptive."

In H.F. 849, the General Assembly called for a final report to be submitted by January 1, 1983 on the progress of the Parole Guidelines System. The report, entitled The Impact of Objective Parole Criteria on Parole Release Rates and Public Protection, documented that the guidelines, combined with a legislatively imposed population limit, had contributed to a 52% increase in paroles during 1981-1982 over the previous two years, which served to prevent extremely serious overcrowding in state prisons. Further, with the aid of the violence-prediction portion of the guidelines, the Board of Parole was able to reduce the rate of violent crime among parolees by 35% and the general threat posed by a typical parolee by 17%. Despite the huge increase in paroles, the conclusion of the report was that public protection had not been compromised in the process.

Transferability of the Iowa Model

During the last several months, the Statistical Analysis Center has considerably simplified and improved the risk-assessment system, with the result that it is now judged suitable for use outside the state. During the next year, the Iowa SAC will be preparing a series of manuals and reports designed to interest other jurisdictions in either adopting the Iowa model or in conducting parallel research. This effort is being funded by the Bureau of Justice Statistics with support from the Criminal Justice Statistics Association.

The present simplified risk-assessment model has been shown to be approximately 88% accurate in predicting serious recidivism and violence. The Iowa SAC conservatively estimates that, with the Iowa model, prison populations nationwide could be reduced by 20% with no increase in threat to the general public. Further, with more emphasis on incapacitation of potential recidivists, the total threat to society posed by paroles could be reduced by 35% while simultaneously reducing prison populations by 5-10%. Various strategies/scenarios for implementing the Iowa model would lead to varying benefits in reduced prison populations and enhanced public protection. In any case, considering the apparent impact that recidivism/parole research has had on the prison population problem in Iowa, we strongly recommend that other states begin considering similar efforts.

The Classified Sentencing System

The Iowa Classified Sentencing System is another example of the application of research methods to the solution of a perplexing problem in criminal justice. Over the last several years in Iowa, as elsewhere, there has been a growing dissatisfaction with parole and the indeterminate sentencing system. Critics point to generally unfettered discretion vested in paroling authorities, the early release of sometimes dangerous criminals, an emphasis on rehabilitation and reintegration at the expense of punishment and deterrence, and the general lack of fairness and "justice" in the determination of criminal penalties.

In 1976, for example, the Iowa General Assembly enacted a group of mandatory sentence provisions that prohibited the granting of probation to certain classes of offenders and established minimum prison terms for others. Applicable to about 10-15% of prisoners, these provisions narrowed the gap between the present indeterminate system and a determinate model eliminating parole release discretion. However, in 1982, following a series of murders charged to parolees from state institutions, the legislature began deliberations on further limitations of parole discretion.

Testimony was received from many sources, including county attorneys, judges, and other public officials to the effect that the indeterminate sentencing system was not working and that the state should move to a determinate system in which judges would have greater control over time served. Some further advocated the move to a Sentencing Guidelines System to allow structured discretion for judges in setting prison terms.

During the latter half of 1981, in anticipation of legislative debate on sentencing reform during the upcoming session, the Governor established a body called the Interagency Council on Criminal Justice Planning. This body was charged with the development of alternatives and proposals for legislative debate on this issue, among others. The Statistical Analysis Center was asked to provide staff support for this effort and, as a result, undertook in-depth research into various sentencing models and systems.

In part, this involved the documentation of the relative advantages and disadvantages of determinate and indeterminate sentencing, including the potential impact on prison population, recidivism, and other facets of system operation of the move to a determinate model. Further, consideration was given to the possibility of a "hybrid" system incorporating aspects of both extremes. Specifically, a system was devised by SAC staff that would allow for continuation of the parole release function, only within definite limits to be set by law. A simple classification system based on current offense severity and past adult felony convictions was devised that would allow the legislature to set basic "ranges" of parole discretion associated with these factors. The more serious the past and present offenses, the less the discretion allowed for early release, and the greater the potential maximum penalty.

This Classified Sentencing System was structured in such a manner that it could be incorporated directly into the Code of Iowa yet retain considerable discretion at the parole level. Further, the Habitual Offender Classification System, which measures the seriousness of the past record, was structured to correlate strongly with the risk of recidivism as ascertained from recidivism research previously conducted in Iowa. In this manner, potentially longer terms could be allowed for the most dangerous offenders, while the early release of low-risk offenders would be further encouraged. The Classified Sentencing System would thus provide for greater potential impact of the previously discussed Parole Guidelines System on prison population and public protection.

Following extensive work with legislative leaders in ironing out the details of the system, a bill was introduced and passed by the Iowa House of Representatives on a 94 to 2 vote. However, because of other pressing matters, time ran out at the end of the session before the bill could be passed by the Senate. Further, many key legislators left the General Assembly prior to the next session, and despite the support of newly elected Governor Branstad, the bill died in committee in 1983 due to a lack of understanding of its ramifications.

Nonetheless, the Classified Sentencing System is now being considered by the Iowa Board of Parole for incorporation in their administrative rules. If adopted, the system would serve to establish "parole policy guidelines" to provide basic expectations on time served for the Board, for inmates, and for the general public. This move has the support of the Governor and may well deflate any further interest in a move to determinate sentencing in Iowa.

Sentencing and Community Corrections

The whole issue of sentencing policy and the use of community alternatives to incarceration has been at the forefront of criminal justice debate in Iowa for a number of years. Following the early success of the Des Moines Project, which received the first "exemplary" designation from the Law Enforcement Assistance Administration, the Legislature moved quickly in 1973 to fund a statewide community corrections network. Based on favorable evaluations of the Des Moines Project prepared by the National Council on Crime and Delinquency, eight locally administered community corrections projects were established across the state offering pre-trial release screening, pretrial supervision, pre-sentence investigation, probation supervision, and community residential corrections services.

With the new emphasis on the use of community corrections programs, the widely expressed expectations of corrections officials, the Legislature, and the Governor were that Iowa was entering a new era of de-institutionalization of adult offenders. Some officials even speculated openly that one of the two major prisons in the state could soon be closed with the expected decrease in the prison population.

In 1976, following a 22% increase in the population during 1975, a special "blue ribbon committee," officially entitled the Advisory Commission on Corrections Relief, was appointed to study the sources of the increase and to offer options to resolve the apparently growing prison population problem. This commission collected extensive data on trends in various correctional populations and on the characteristics of prisoners and community corrections clients.

The conclusion of the Commission, after a careful consideration of available evidence, was that the prison population problem was only temporary in nature and that no new prison construction was necessary. The basis for this conclusion was twofold. One, the Commission's analysis of a profile of state prisoners indicated that as many as 15-20% of committed offenders could, and should be diverted to community corrections programs with little additional threat to public safety. Two, the Commission observed that the community residential corrections programs had been implemented in only one of the state's eight judicial districts, and that establishment of these programs in remaining districts would reduce "unnecessary" prison commitments.

Since residential facilities were to be expanded statewide during 1977, the expectation of the Commission was that individuals of the type identified in the 15-20% group would no longer be sent to prison. This observation was further supported by the estimate of a public official that the existing residential program in Des Moines had operated as an alternative to imprisonment, rather than to probation or other programs, for 75% of its clients. In fact, the Commission developed a prison population projection that indicated a decline in the population in the foreseeable future based on the impact of the new residential programs.

However, the expectations of the Commission did not hold true, despite the full implementation of the community corrections model. In fact, the prison population continued on the previous upward trend as if nothing whatsoever had happened in the interim.

During 1977-1978, a second group called the Corrections Master Plan Task Force was formed to study the Commission's findings, to determine the sources of error in their estimates, and to make further appropriate recommendations. This group concluded that the major error in the Commission's calculations derived from a faulty estimate of the potential impact of the new residential programs on the prison population. By means of a careful statistical analysis of offender characteristics, the Master Plan staff concluded that instead of the 75% estimate used by the Commission, at most 20% of residential corrections clients were being diverted from the state prison system, and no more than 15% successfully. The Commission had vastly over-estimated the impact of the community corrections program on the prison population. The majority of residential clients were in fact young property offenders without serious prior records, most of whom would have been placed on straight probation had the residential programs not been

available. The full conclusions of the Master Plan group on this issue are detailed in the report Crime and Criminal Justice in Iowa: Volume IX - Prison Population from the Statistical Analysis Center.

The SAC report concluded further that the members of the 15-20% group identified by the Commission were not in fact better risks for release than were remaining prisoners, partially due to the fact that the Commission's analysis ignored factors associated with the risk of recidivism. However, this opened up a whole line of research by the SAC as to the identification of individuals committed to state prisons unnecessarily, i.e., who could safely be placed in community corrections programs without depreciating the seriousness of convicting offenses.

As in the Commission's analysis, the SAC examined this issue in relation to the problem of jurisdictional disparity in sentencing. It soon became apparent that many of the so-called "marginal" commitments to state prisons were coming from certain counties and judicial districts where more reliance was placed on the use of imprisonment as a sentencing alternative. In 1979, a committee was appointed by the State Supreme Court to study the sentencing disparity problem, with SAC staff providing research support for the project.

Following the study of approximately 17,000 felony sentences handed down in the state during 1974-1979, the SAC developed a statistical model for ascertaining the extent of sentencing disparity across state jurisdictions. The application of this model led to the conclusion by the Committee that sentencing disparity by jurisdiction did exist in Iowa, and that it was significant, but that it did not constitute a serious enough problem that remedial steps should be taken at that time.

Despite the conclusion of the committee, interest continued in the state in the sentencing disparity phenomenon, especially in light of the fact that the prison population was continuing to rise. With the prison population problem as a major source of inspiration, the SAC became heavily involved in the study of existing sentencing policies of judges and generally in a line of research parallel to the study of parole decision-making. In fact, a system of model sentencing guidelines, based on the same classification scheme as the parole guidelines, was formulated in late 1980. The reports Risk Assessment in Iowa and Offender Risk Assessment: Implications for Sentencing and Parole Policy from the SAC discuss the results of the sentencing research and the potential impact of implementation of the model guidelines.

It was the estimate of the SAC that both the number of commitments to state prisons and the burden of recidivism among community corrections clients could be reduced by 25% if judges would begin using the model guidelines in reaching sentencing decisions. Although the sentencing guidelines were not implemented in the same manner as the parole guidelines, they were tested in Polk County during 1981 with apparently favorable results.

In early 1982, following continued increase in prison population, the Department of Social Services, which operated the prison system and monitored community corrections, asked the SAC to undertake a study of the impact of jurisdictional sentencing disparity in the prison population and to look for means of solving the problem. As a result, detailed offender and offense information was collected from inmate files on 559 individuals committed to state prisons during 1981. The staff calculated risk assessments on all 559 offenders and otherwise developed measures of the seriousness of the committing offense, the prior record, and other factors affecting the decision to imprison. A model was then constructed to judge the "marginality" of a given commitment, i.e., the extent to which a given offender might have served as a good candidate for community corrections.

The conclusions of this study, which have yet to be published, were that (1) approximately 15-20% of committed offenders could instead have realistically been placed in community corrections programs without serious additional threat to society and (2) the bulk of these "marginal" commitments were coming from certain judicial districts and counties with higher than average commitment rates. This suggested an avenue of approach to solving the problem by working with judges in those jurisdictions in better identifying good candidates for release.

Further, SAC staff formulated a strategy for the identification of "marginal commitments" soon after prison admission to allow corrections officials to recommend such individuals back to the court for reconsideration of sentence and release on shock probation. It was envisioned that this process would also provide candidates for early parole. To date, the project has not gotten underway due to a problem with the timely generation of inmate files in the Parole Board Office. Nonetheless, the SAC fully intends to pursue this project as time permits.

Looking at Statistics in the "Right" Way

Much of what we have discussed so far has concerned rather sophisticated studies of recidivism and decision-making patterns and the structuring of guidelines and other tools for translating research into practice. However, there are a number of quite simple and straightforward functions that research organizations can undertake that can have an impact on criminal justice policy. Two examples of such application come to mind.

In late 1981, when the Interagency Council on Criminal Justice Planning was discussing system problems and needs, it was noted that the increase in the prison population during the preceding several years had been accompanied by a nearly corresponding increase in the level of reported violent crime in the state. This concurrence in trends had frequently been noted by legislators and other system critics, and served to convince many that a new state prison was necessary to meet the increase in violent crime.

However, in presenting the information to the Council, the staff noted that the increase in violent crime had, in fact, very little relationship to the increase in the prison population. The reason for this derived from a more careful examination of the crime data and of basic statistics on system operation. The SAC staff observed that almost all of the increase in violent crime in Iowa between 1975 and 1980 fell in one category--aggravated assault. The number of murders, rapes, and robberies had changed very little. Further, aggravated assault is the one violent crime that is unlikely to lead to imprisonment for several reasons. First, many are reduced to simple assaults, which can't possibly result in a prison term. Second, the commitment rate among persons convicted of aggravated assault is comparable to that for burglary and is much less than that for the other violent crimes. Finally, most aggravated assaults are charged at a level that can lead to at most a one- or two-year prison sentence, with the result that persons convicted of that crime generally serve shorter than average prison terms. All this is reflected in the fact that only a very small fraction of prisoners are committed for aggravated assault.

These observations established clearly that the increase in reported violent crime in Iowa during 1975-1980 was very misleading. The SAC staff demonstrated that much of the increase in the prison population during 1975-1980 was not due to an increase in violent crime, but rather to an increase in adult arrests for non-violent crimes such as burglary. This observation was vital to an understanding of what had indeed been happening in Iowa's criminal justice system and where the major problems lay.

A second situation arose in early 1981 during legislative debate on a solution to the prison population problem. At that time, the General Assembly was made aware that the prison population had jumped by 18% during 1979-1980. Further, the Department of Social Services had noted that prison admissions had increased by 12% from 1977-1978 to 1979-1980, and based in large part on this trend, the Department was projecting a further increase in the population through 1982. Since there is very little that can be done to reduce admission levels directly, indications were that new construction might be necessary to meet future populations.

However, in taking a closer look at prison population movement data, the Statistical Analysis Center observed that the 12% increase in admissions during 1979-1980 was actually an artifact of the advent of shock probation as a sentencing alternative for judges. In fact, 86% of the increase in admissions was of offenders who would end up serving no more than three months in prison. As a result, the impact on the prison population of increased admissions was far less than 12%. Further, most shock probationers are young property offenders without serious records who would have been placed in community corrections minus the shock probation option, i.e., the 12% admission increase was for the most part <u>not</u> of regular "prison types." Instead, the SAC documented that 79% of the increase in the prison population during 1979-1980 was due to a 29% reduction in paroles from 1977-1978. This apparently had gone unnoticed, or at least was not sufficiently emphasized in the analyses by the Department and its consultants.

As a result of the more accurate trend analysis provided by the SAC, the legislature recognized the real source of the problem at the time--reduced paroles--and subsequently enacted a prison population cap, placing most of the responsibility for population control with the Board of Parole.

Much of the work we have done in Iowa leads to the conclusion that many of the assumptions upon which the justice system has operated in the past are clearly in error. Study of the prison population and of sentencing practices has brought to light many of these false assumptions.

One example which is applicable in almost every state pertains to judicial sentencing practices. In discussions with criminal justice professionals, legislators, and concerned citizens, everybody agrees that judges send the "worst risks" to prison. After all, what judge would jeopardize the public by knowingly referring high risks to non-secure settings?

Our study of Iowa sentencing practices, combined with offender-based risk assessments, shows clearly that there is not a strong risk-related dichotomy between those sent to prison and those who remain in the community. While it is possible to develop a system of classification that would generally separate "prison types" from "Community-based types," such a system would not necessarily correlate strongly with risk. In Iowa, for example, the highest average risk for new property crime is found among those referred to community residential facilities, not among those sent to prison.

Another example is found in the dichotomy frequently made between those sent to prison for property offenses and those referred for violent crimes. In discussions about methods to avoid prison overcrowding in Iowa, it has not been unusual to hear support for the concept of releasing property offenders at earlier dates, while either maintaining or lengthening terms for violent offenders. Much of this debate apparently is based upon the belief that those committed to prison for violent offenses are significantly more dangerous to the public than those committed for property offenses.

Drawing such a simplistic dichotomy, however, shows little appreciation for either the nature of many violent crimes or for the development of criminal careers. Iowa research shows clearly that many of those committed to prison for violent crimes are incarcerated not because of lengthy criminal careers or dangerousness, but due to the severity of a single offense. While society may frequently feel the need to <u>punish</u> single offense violent offenders, our research indicates that many present very little threat to society if referred to community--based or other non-secure correctional programs.

On the other hand, many of the property offenders sent to prison constitute real threats to public safety, not only for new property crime, but for violent crime as well. This is true in part because of many of the property offenders sent to prison are committed due to lengthy criminal careers and a series of past failures in community-based programs. Many of them are incarcerated in their late teens and early twenties during very active portions of their criminal careers. Some, while they show only property offenses on their rap sheets, will ultimately "graduate" to armed robbery and other more serious violent offenses.

The lesson in this is that research may frequently be of assistance in examining the validity of the information and assumptions upon which policies are based. This is particularly true in the justice system, as all too frequently policies and procedures exist simply because people have gotten used to them. Many of the decisions made in the justice system historically have been made by "the seat of the pants."

We are starting to make some progress in this area, however, as practitioners and planners become more comfortable with basing their decisions upon good information. We researchers have an obligation to assist the policymakers in reaching informed conclusions.

Our experience with the Iowa Board of Parole is instructive. The Board wasn't quite sure about the utility of statistical risk assessments when they first became available to them two years ago, and at first they viewed the assessment somewhat suspiciously. After having referred to individual assessments since that time, however, they now report a need for the assessments to make informed parole decisions. The researcher who attends the needs of practitioners and policymakers can have significant impact on system operation, and, to our thinking, is perhaps the best hope for improved justice system operation in the future.

<u>Note:</u>

*We have space here to give no more than a general overview of the development, structure, application, and impact of the Parole Guidelines System. For a detailed discussion of these topics the reader should refer to several publications from the Statistical Analysis Center. In addition, two other reports provide useful references in understanding the Iowa research:

"Finding the Balance in Prison Sentencing," <u>Iowa Review Quarterly</u>, Office for Planning and Programming, Winter, 1983, Volume 2, Number 1.

Keon S. Chi, <u>Offender Risk Assessment: The Iowa Model</u>, Innovations Transfer Project, The Council of State Governments, July, 1983.

Is Crime Seasonal?*

Carolyn Rebecca Block, *Statistical Analysis Center/ Illinois Criminal Justice Information Authority*

Introduction

Researchers and policy makers often take for granted that seasonal fluctuation in crime is an established fact. To suggest otherwise goes against the grain of a long tradition in criminology. Indeed, Brearley (1932:161-199) begins his review of the literature on criminal seasonality with Hippocrates, and Wolfgang (1966) cites scholarly works dating from 1825.

Quetelet, a Belgian statistician and one of the earliest investigators of seasonal fluctuation in crime (1842:90;also see Sylvester,1982), states,

> The seasons, in their course, exercise a very marked influence: thus, during summer, the greatest number of crimes against persons are committed and the fewest against property; the contrary takes place during the winter.

The assumption that crime occurs seasonally continues to be made today. In heralding the "War on Crime" of the Johnson administration, the President's Commission on Law Enforcement and the Administration of Justice (1967:27) stated that

> Murder is a seasonal offense. Rates are generally higher in the summer, except for December, which is often the highest month and almost always 5 to 20 percent above the yearly average. In December, 1963, following the assassination of President Kennedy, murders were below the yearly average by 4 percent, one of the few years in the history of the UCR that this occurred.

One of the most influential basic criminology textbooks in the United States, Sutherland and Cressey (1978:82), states that

> statistical studies show very uniformly that crimes against property reach a maximum in winter months, and crimes against the person and against morals in the summer months.

In reality, the answer to the question, "Is Crime Seasonal?" is not nearly so straightforward as these quotes would suggest. Some types of crimes fluctuate with the seasons, while others do not. The same crime may show seasonal fluctuation in one geographic area, but not in another. The same crime in the same geographic area may fluctuate seasonally during one time period but not during another. In addition, the decision as to whether or not a particular series is seasonal depends upon the conceptual and operational definition of seasonality the decision-maker uses.

This report first reviews some analytical issues that must be considered in decisions regarding the seasonal fluctuation of crime. Second, using analysis that the Statistical Analysis Center has conducted over the years, the report answers the question, "Is crime seasonal?" for specific crime types, places, and time periods.

The report ends with a review of analyses, published and unpublished, that deal with seasonal fluctuation in crime, including a review of the findings of a previously unpublished survey of analyses done at the state level. Studies were not selected for inclusion in this review bibliography on the basis of good methodology or any other quality criterion. The report simply attempts to provide a comprehensive review of the existing literature.

Issues in the seasonality of crime

The question, "Is crime seasonal?" is not easily answered with a yes or a no. The answer depends upon the statistical criteria used to make the decision (see Block,1983). Methods of seasonal analysis are not completely objective. Their results must be interpreted, and researchers using the same method may come to differing conclusions. There is even disagreement among statisticians on the very definition of seasonality.

For a complete discussion of alternative conceptual and operational definitions, see Block (1983). For the purpose of this report, we will use Kallek's (1978:15) simple and straightforward definition:

> Seasonality refers to regular periodic fluctuations which recur every year with about the same timing and with the same intensity and which, most importantly, can be measured and removed from the time series under review.

Not all analysts concur with this definition. A case in point is Warren, et al. (1981), who found homicide to have a seasonal pattern that changes from year to year. That is, a "peak month" in some years is a "trough month" in other years. The authors conclude that homicide is seasonal, but "inconsistent." Since year-to-year consistency is implied in Kallek's definition, if we accept that definition, "inconsistent seasonality" is a contradiction in terms.

This section deals with several methodological and analytical issues that affect an investigator's decision as to seasonality. These issues may explain the seemingly contradictory results of some studies of seasonal fluctuation in crime.

Length of series. Generally, a number of years of data are necessary in order to answer the question, "Is this crime seasonal?" with confidence. A series shorter than seven years is considered too short for a definite decision about the presence of seasonality (see Block, 1983). The reason for this becomes clear if you consider that, in one year, you observe one instance of each month. In six years, you would have only six observations of Januaries, six observations of Februaries, and so on.

With an increasing number of observations (years), seasonal fluctuation is described more accurately. A few extremes will have less effect on the total analysis in a longer series than they will in a short series. Also, with a very short series, only strong seasonal fluctuation is likely to produce statistical significance. In general, the longer the series, the more likely that relatively weak seasonal fluctuation (that is, however, consistent over time) will be significant.

Table 1 shows the effect of length of series on a seasonal analysis, using the same method (Census X-11) and the same statistical criteria (F of stable seasonality and percent contributions of the seasonal and irregular components). The F value is an indicator of significance, while the percent contribution of the seasonal component is analagous to a measure of association.(1) The series, seven Index crimes in Illinois, begin in 1972, but extend either to 1977 (six years) or to 1981 (ten years).

Because we cannot assume that observations in a time series are independent, the stable seasonality F should be interpreted as an indicator of the degree of seasonality, not as an exact measure of significance. The "Plewes rule-of-thumb" (Block, 1983) uses the irregular contribution as a means of interpreting the F value. The Plewes criteria are: If the irregular contributes 30 percent or more of the total month-to-month variation, the decision should be "no stable seasonality," regardless of the F value. If the percent contribution is 25 to 29, the F value needs to be at least 15, and if the percent contribution is 15 to 24, the F value needs to be at least 2.41 for the series to be considered seasonal. An F value less than 2.41 indicates no stable seasonality, regardless of the irregular contribution.

One effect of increasing the length of the Illinois Index crime series was to increase the F value. This happened for murder, forcible rape, robbery, and burglary. However, the stable seasonality F values of aggravated assault, larceny/theft and motor vehicle theft were lower in the longer series.

The percent contribution of the seasonal component over a one-month span does not always increase when the length of the series increases. For example, for forcible rape, the seasonal contribution is 20.5% in the longer series, and 27.2% in the shorter series, although the F is higher in the longer series. This is also true of murder and robbery. Why should this be so?

It is, of course, possible that there is less seasonal fluctuation in the 1978-1981 years than in the previous six years. In that case, the addition of the 1978-1981 years to the series would decrease the seasonal fluctuation overall. However, if that were true, why would the F value increase? The answer to this apparent contradiction is that a longer series allows a more accurate description of seasonal activity. This more accurate description tells us that the seasonal contribution is less. Thus, it is possible that these violent crime series contain a weak degree of seasonal fluctuation, and that this weak fluctuation might become significant in series that are even longer than these.

The literature of crime seasonality contains numerous examples of seasonal analyses based on very short series, sometimes only one year (see "Review bibliography," below). It is not surprising that different analysts, analyzing the same crime but for varying time periods, would reach differing conclusions about the presence of seasonal fluctuation in that crime.

Local versus state or national. What is the appropriate level of geographic aggregation to answer the question, "Is crime seasonal?" Should we look for an answer at the local level, or at the state or national level? On one hand,

	Six-Year Series 1972-1977			Ten-Year Series 1972-1981		
Index Crime	Stable F	% Contribution Seasonal	Irregular	Stable F	% Contribution Seasonal	Irregular
Murder	1.5	30.5%	68.7%	2.8	29.0%	70.4%
Forcible rape	8.1	27.2	72.2	10.0	20.5	79.0
Robbery	7.1	34.4	60.5	10.9	32.3	65.0
Aggravated assault	49.5	63.1	35.7	45.7	59.6	38.5
Burglary	10.8	52.8	43.2	19.3	55.9	41.7
Larceny/theft	104.6	81.2	17.2	95.8	80.7	18.2
Motor vehicle theft	16.2	51.9	46.8	13.0	39.1	59.6

Table 1. Census X-11 results, Illinois Index crimes.

"using local data may lead to erroneous conclusions about the seasonal pattern because of the small numbers involved and the possibility of local intervention or prevention efforts" (Michael and Zumpe, 1983). On the other hand, many, if not most, administrative and policy decisions in law enforcement are made at the local level, and relate to local policy. If these decisions are based on information about seasonality at at the national level, they may lead to erroneous local-level conclusions (Coldren, 1980).

The same phenomenon may vary seasonally in one place but not in another. Some geographic areas, such as a college town, a tourist mecca, or the home of the state fair, have an influx of population during certain seasons of the year. In addition, if the weather is an underlying cause of seasonal fluctuation, areas with different climates may experience differing patterns of seasonality. For examples, see, in the "Survey of the states" below, Arizona, Delaware and Maine.

The argument that the effect of local intervention or other "error" is "equalled out" in nationally aggregate data is spurious. Error never disappears, though it may be hidden. At the local level, with knowledge of the local situation, there is a better chance to discover and then control for local effects.

The case of aggravated assault demonstrates that conclusions about seasonal fluctuation in a crime may be very different, depending on the local area concerned. In Illinois, the F of stable seasonality in Index assault from 1972 to 1981 varies by jurisdiction -- Chicago, Cook County (excluding Chicago), SMSA counties containing a central city, suburban SMSA counties not containing a central city, and a typical small city of less than 50,000 population (see chart 1). The less urban the place, the more aggravated assault known to the police appears to fluctuate seasonally.

Such a change in the F, a measure of significance, might be argued to be caused by a change in the number of observations. It is true that a longer series, or a series containing higher observations, is more likely to produce a high F, other things being equal, and there are fewer reported assaults in Quincy than in Chicago. However, the lack of seasonal fluctuation in Quincy is evident not only in the F statistic, but also upon inspection of the graph. Despite the difference in scale, Chart 2 and chart 3 show that there is clearly a pattern of seasonal fluctuation in Chicago, but no discernible seasonality in Quincy.

Does crime occur seasonally, or is it reported seasonally? Most homicides are, in their characteristics, a subset of another crime -- aggravated assault. They begin as a fight, brawl or argument. Although some homicides are precipitated by a robbery, a rape, a gangland "hit," or another non-assault circumstance, the great majority begin as a fight, brawl or

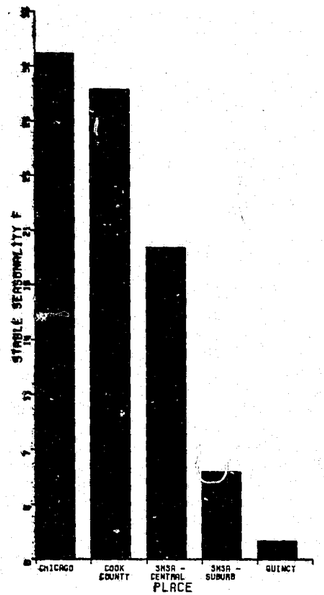


Chart 1. Seasonal fluctuation of aggravated assault: Illinois, 1972-1981 (Census X-11 additive adjustment).

argument, and escalate to murder. In Chicago, 68 percent of the homicides known to the police from 1965 to 1981 began as an assault, 17 percent began as a robbery, 2 percent began as a rape or burglary, and the rest involved other or unknown circumstances.(2)

Homicides that begin as fights or brawls differ from homicides that begin as robberies in a number of ways, including the weapon, the place and time of occurrence, the victim-offender relationship, and many other characteristics. In fact, homicides that begin as assaults are more similar in their characteristics to aggravated assaults than they are to homicides that begin as robberies (see Block, 1977). They can almost be considered to be separate types of crime. A homicide that begins as a fight or assault can be thought of as a type of aggravated assault, one in which the victim was injured so seriously that death resulted.

Because assault homicides are similar in most of their characteristics to all aggravated assaults, we would expect them to be similar in another characteristic, seasonal fluctuation. If more assaults occur in the summer months, for example, we would expect more assault homicides to occur in the summer. However, Chicago data (table 2) indicate that this is not necessarily true. Aggravated assaults known to the police fluctuate seasonally. Assault homicides do not. Why does this occur?

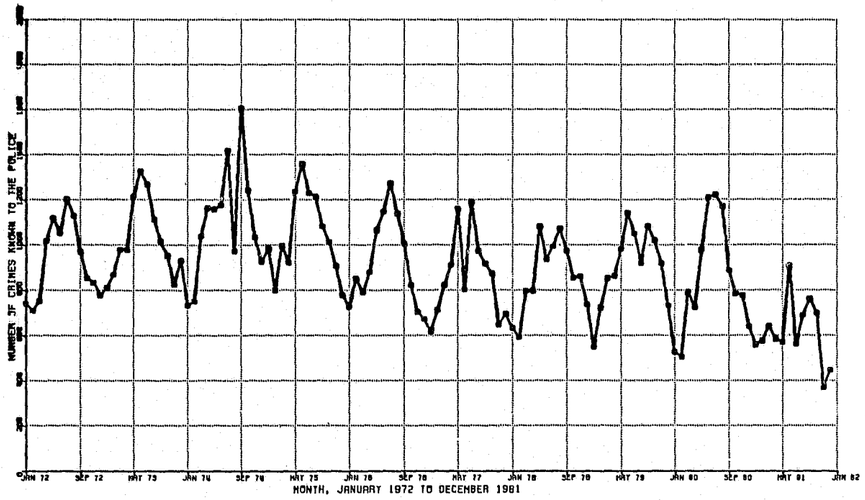


Chart 2. Chicago Index aggravated assault, 1972-1981.*

Source: SAC edition, Illinois Uniform Crime Reports offense data.

* Index aggravated assault includes aggravated assault, aggravated battery, and attempted murder.

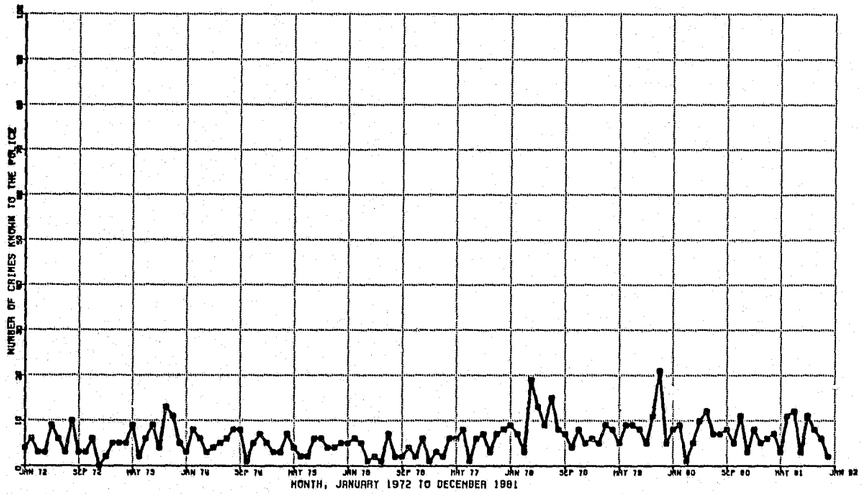


Chart 3. Quincy Index aggravated assault, 1972-1981.*

Source: SAC edition, Illinois Uniform Crime Reports offense data.

* Index aggravated assault includes aggravated assault, aggravated battery, and attempted murder.

Index Crime	Additive Assumption* Stable F	% Contribution Seasonal	Irregular	Multiplicative Assumption* Stable F	% Contribution Seasonal	Irregular
All weapon types						
Aggravated assault	109.1	70.0%	27.3%	104.5	70.3%	27.4%
Assault homicide	6.2	25.1	74.3	6.2	25.7	73.7
Firearms						
Aggravated assault	60.9	60.0%	35.4%	54.2	57.4%	38.7%
Assault homicide	4.3	24.2	75.2	4.2	24.1	75.3
Other weapons						
Aggravated assault	89.2	65.6%	31.5%	90.6	65.0%	32.4%
Assault homicide	3.8	29.3	70.3	3.8	24.2	75.5

Table 2. Aggravated assault and assault homicide: Chicago, 1967-1981.**

*Under an additive assumption, the seasonal and irregular components are independent; under a multiplicative assumption, they are not. For more detail, see Block (1983).

**Sources: Assault data set re-constructed by the Authority from two sources, each containing partial data: Chicago Police Department records, and the City of Chicago Municipal Reference Library. Time series data are in 13 "police periods" per year. To obtain estimates of months for the present analysis, we used a moving average. This probably decreased the amount of seasonal fluctuation. Assault homicide: see note 2.

In searching for an explanation, we hypothesized that, in fact, assault does not occur more often in the summer than in the winter months. However, it becomes known to the police more often in the summer months. More serious assaults, involving serious injury and possibly hospitalization, become known to the police with greater likelihood than less serious assaults. They are reported by medical personnel and hospital staff, and are more likely to be reported by the victims themselves (LEAA, 1972; Block and Block,1984). Less serious assaults, on the other hand, tend not to come to the attention of the police unless they are public. They are more likely to be public in the summertime. In the warm months, an assault is more likely to occur outside, and if it occurs inside, the windows are more likely to be open. Thus, we hypothesized that the explanation for the seasonality of assault and lack of seasonality of assault homicide is that neither, in fact, occurs seasonally, but that more assaults become known to the police in the summer months.

To test this hypothesis, we compared seasonal fluctuation in the number of aggravated assaults known to the police, and seasonal fluctuation in the number of aggravated assault victimizations, for the same place and time period: the United States from 1973 through 1979. The police data are Index aggravated assaults reported by police jurisdictions to the FBI through the Uniform Crime Reporting program.(3) The victimization data are aggravated assault victimizations (defined the same as Index aggravated assaults) estimated by the National Crime Survey.(4) We expected to find much more seasonal fluctuation in reported assaults than in assault victimizations. That is, in fact, what we found, as table 3 shows.

Having found that reported assault fluctuates with the seasons much more than assault victimization, at least for the United States as a whole, we looked at the same relationship for another violent crime -- robbery (table 3). Robbery victimizations, like assault victimizations, do not fluctuate with the seasons, but robberies known to the police do. This suggests that less serious and less public robberies do not occur more often in the summer, but are more likely to become known to the police in the summer months.

Because the hypothesis is grounded on the issue of "public" crime, and the degree to which a crime is public may be related to whether it occurs in an urban area or not, a better test of the hypothesis would compare seasonal fluctuation in victim and police data in urban areas only. If this were done, we would expect that the difference in Census X-11 results would be even higher. However, data are not currently available to do this.(5) Partial support is found in the above analysis of assault data (see "Local versus state or national"). Aggravated assault series are more likely to be seasonal the more urban the place. It is also supported by the seasonality of Chicago assaults by type of weapon, in table 2 above. The F of stable seasonality and the percent contribution of the seasonal component are higher for assaults without a gun than for assaults with a gun.

Index Crime	Additive Assumption Stable F	% Contribution Seasonal	Irregular	Multiplicative Assumption Stable F	% Contribution Seasonal	Irregular
Aggravated assault						
Known to police	94.2	83.7%	15.0%	131.3	85.0%	13.7%
Victim survey	7.5	31.0	68.4	7.4	29.9	69.5
Robbery						
Known to police	100.4	86.9%	10.6%	107.6	86.9%	11.0%
Victim survey	4.6	34.0	63.7	4.5	27.3	70.4

Table 3. Seasonality of victimization and reported crime: United States, 1973-1979.

Perhaps assaults without a gun are less serious, result in lesser or no injuries, and are thus more likely to become known to the police in public situations, especially in the summer months.

Is crime seasonal?

In this section, we summarize the findings of analyses of seasonality in crime that the Statistical Analysis Center of the Authority has conducted in the last few years, either in response to requests from users or in conjunction with a research project.

The reader's evaluation of the findings presented here, and indeed the evaluation of any analysis of seasonality, should take into account the method and criteria used, the place and time period, and the definition of the crime.

To facilitate comparison, all the analyses summarized here use the same method, the Census X-11, and the same criteria, F of stable seasonality and percent contributions of the seasonal and irregular components over a one-month span. However, we have repeated many of these analyses with a stochastic (ARIMA) model, and results are available on request to the author.

Crime definitions used are also consistent throughout the section: Index crimes, as defined by the Uniform Crime Reports. Even when we analyze victimization estimates from the National Crime Survey, we attempt to make NCS crime categories comparable to Index crime categories (see note 4, above, and Block and Block,1984). The section summarizes, in turn, findings for each of the four violent Index crimes and three property Index crimes. The eighth Index crime, arson, is not included, because, in general, consistently defined data have not been collected for the minimum seven-year period.

Place and time, however, change from analysis to analysis, and are noted for each.

Sources of the time series analyzed in this section are the following:

- United States 1970 to 1982: Uniform Crime Reports, FBI. See note 3.
- National Crime Survey: See note 4.
- Illinois, parts of Illinois: Authority version of the Illinois Uniform Crime Reports offense data or SHR data. See Miller and Block (1983).
- Chicago Aggravated Assault: data set reconstructed by the Authority from two sources, each containing partial data: Chicago Police Department records, and the City of Chicago Municipal Reference Library.
- Chicago homicide: Data collected from police homicide files, with the cooperation and assistance of the Chicago Police Department over a number of years and changes in administration. See note 2.
- California, Los Angeles: Authority time series version of data obtained from the Bureau of Criminal Statistics and Special Services, state of California.
- New York City: Lily E. Christ, John Jay College of Criminal Justice, City University of New York.
- Boston: Deutsch and Alt (1977); see "Review bibliography."
- Canada, Ontario: Time series files created by Craig McKie, Statistics Canada, from homicide data collected by the Law Enforcement Section, Canadian Centre for Justice Statistics, Statistics Canada. See Block, et. al., 1983.

Is violent crime seasonal?

Homicide. The preponderance of the empirical literature finds that homicide does not vary seasonally.(6) In the "Review bibliography" below, see Deutsch and Alt (1977), Hay and McCleary (1979), Lamp (1982), and Michael and Zumpe (1983). However, the prevailing point of view in criminology seems to be that homicide is seasonal. See the quote from the President's Commission in the Introduction, above, and Warren, et al. (1981).

As part of various research projects and in response to user requests, the Authority has analyzed hundreds of homicide series for the presence of seasonal fluctuation. Most of these series were types of Chicago homicide or types of Canadian homicide; for example, homicide committed with a gun, homicide committed with a knife, and homicide committed with a blunt instrument. In addition to the Chicago and Canadian series, we have analyzed Boston homicide, California and types of California homicide, Index homicide in the United States, Index homicide in New York City, and homicide and types of homicide in Illinois jurisdictions outside of Chicago. Table 4 summarizes Bell-Canada(7) results for the major Chicago homicide component series.

In addition to the types of homicide presented in table 4, we analyzed more detailed Chicago homicide types, such as multiple offender assault homicide with a gun. We also conducted more detailed analyses, using the Census X-11, on any series having a Bell-Canada F value of 2.41 or over. After analyzing over a hundred Chicago series, we finally found one that showed some seasonal fluctuation. This may be one of the few cases in which the maxim, "the exception proves the rule," is true. Of the dozens of series, the only type of homicide that fluctuates with the seasons is homicide occurring out of doors or in a vehicle. This type of homicide tends to occur half as often in January and almost twice as often in August as in an average month.

	Additive Assumption	Multiplicative Assumption
Total homicide	4.25	4.13
Firearm	2.86	2.75
Knife	2.17	2.07
Blunt instrument	3.15	*
Assault	5.38	5.20
Robbery	2.07	*
Family	1.23	1.26
Acquaintance	2.36	2.20
Stranger	2.58	2.04
Black victim	3.05	3.05
White victim	1.86	2.12
Male victim	5.34	5.00
Female victim	.39	.39
Offender age 15-19	1.43	*
Offender age 20-24	1.13	1.50
Offender other ages	4.15	4.10
Single offender	5.05	4.94
Multiple offenders	1.83	1.59
Inside a residence	1.28	.90
Inside a nonresidence	3.03	3.24
Out of doors, vehicle	18.60	16.56

Table 4. Bell-Canada F of stable seasonality: Chicago homicide 1965-1981.

*Because the series contains one or more zero values, a multiplicative adjustment is not applicable.

The Chicago analysis was duplicated with twenty years of Canadian homicide data, also categorized into component types. Consistently, each Canadian series showed a lower F value than the corresponding Chicago series (see table 5). This was not due to more observations in the Chicago data; the opposite is the case. The Canadian series is longer, and contains about the same number of homicides per month as the Chicago series.

None of the California homicide series, nor the twenty component California series, was seasonal (however, note that these series are short). The same is true in New York City and Boston.

The only exception to the general finding of very low F values for homicide is United States homicide from 1970 through 1982. There are several possible reasons for this. First, the relatively high F value may reflect the high number of cases, which ranges from 1087 to 1866 per month. Second, it may reflect some artifact in the data. These data were collected by the FBI from reports from local jurisdictions, and the number of jurisdictions increases over time (see note 3). Third, it is possible that homicide fluctuates seasonally in some United States localities that we have not analyzed, southern states for example. If so, seasonal fluctuation in these areas might override the lack of seasonal fluctuation in Illinois, New York City, Boston, and California to produce seasonality in the national aggregate.

Even though the United States Index murder series is significantly seasonal according to the Plewes criteria, the degree of seasonal fluctuation is very small. An indicator of this is the "final seasonal factors," a result of the Census X-11 analysis method. A seasonal factor can be interpreted as a monthly weight. If it is over 1.00, then the month tends to be high; if it is under 1.00, then the month tends to be low. For example, the January seasonal factors for Chicago homicides occurring outside or in a vehicle, discussed above, are all less than 0.50 over the seventeen years, and the August seasonal factors all approach 2.00.

In contrast, the seasonal factors for United States Index murder (multiplicative assumption) are close to 1.00 for all months. For 1982, they are the following:

January	.99	July	1.08
February	.90	August	1.07
March	.97	September	1.03
April	.92	October	.99
May	.99	November	.98
June	1.00	December	1.07

Seven or eight percent more Index murders tend to occur in July, August, or December than in the typical month. About ten percent fewer tend to occur in February. Otherwise, no month tends to be particularly high or low. However, because this weak pattern is consistent over the thirteen years of the series, and because the number of cases is so high, the F value is relatively high.

In summary, it is possible that certain types of homicide, such as those occurring out of doors or in certain parts of the country, vary with the season of the year. However, for most types of homicide in most places, seasonal fluctuation, if it exists at all, may be too weak to affect practical administrative or policy decisions.

Forcible rape. Forcible rape, another violent Index crime, is, like homicide, seasonal in the United States as a whole (table 6). However, it is not seasonal, by the Plewes criteria, in New York City or in any of the Illinois jurisdictions we have analyzed, including three categories of counties: Cook County, which contains Chicago, small-city counties (nonmetropolitan counties with a city of 25,000 to 49,999 population) and rural counties (all other nonmetropolitan counties).

For other analyses of forcible rape, see, in the "Review bibliography" below, Deutsch (1978), Edgerton et al. (1978), FBI (1981), Lamp (1983), Marshall (1977), Michael and Zumpe (1983). In

		Additive Assumption			Multiplicative Assumption	
Index Crime	Stable F	% Contribution Seasonal	Irregular	Stable F	% Contribution Seasonal	Irregular
Total homicide						
United States 1970-82	28.9	70.1%	28.2%	31.0	71.5%	27.1%
Canada 1961-80	1.4	13.7	86.0	1.4	16.7	83.0
Illinois 1972-81	2.7	29.0	70.4	2.8	29.4	70.0
California 1976-80*	5.1**	44.0	53.8	7.8	49.2	48.7
Ontario 1961-80	0.8	18.7	81.0	0.7	16.6	83.2
Chicago 1965-81	5.6	30.5	67.2	5.3	27.1	70.7
Other Illinois 1973-82	0.5	9.7	90.0	0.5	5.4	94.5
New York City 1973-82	7.2	45.2	54.5	7.1	41.0	58.7
Los Angeles 1976-79*	2.7	39.1	60.4	2.4**	35.8	63.7
Boston 1966-75	2.8	26.8	72.4	2.7	27.7	71.5
Homicide with a gun						
Canada 1961-80	2.1	19.0%	80.6%	2.3	19.1%	80.7%
Illinois 1973-82	2.0	26.9	72.5	2.1	9.0	90.3
California 1976-79*	1.5**	35.2	61.3	1.6	9.6	89.9
Ontario 1961-80	1.8	22.0	77.6	***	***	***
Chicago 1965-81	3.8	29.3	69.7	3.6	25.[illegible]	74.0
Other Illinois 1973-82	0.5	12.4	87.4	0.5	5.0	94.8
Los Angeles 1976-79*	1.4	41.8	57.7	1.3	31.0	67.7
Assault homicide						
Canada 1961-80	1.9**	22.0%	77.8%	1.7	26.6%	73.3%
Illinois 1973-82	3.5	23.8	75.8	3.5	10.5	89.4
California 1976-79*	3.0	29.8	69.3	3.0	25.8	73.4
Chicago 1965-81	6.2	21.6	77.6	5.7	22.2	77.2
Other Illinois 1973-82	1.1	17.5	82.1	1.0	10.8	89.0
Robbery homicide						
Canada 1961-80	1.9	25.1%	74.5%	***	***	***
Illinois 1973-82	1.3	19.8	77.3	***	***	***
California 1976-79*	0.5	9.0	90.4	0.5	9.6%	89.8%
Chicago 1965-81	2.4	9.1	80.5	2.4	19.1	80.5
Other Illinois 1973-82	0.7	13.6	86.0	***	***	***

Table 5. Census X-11 results in selected homicide series.

*Note that this series contains fewer than seven years. See "Length of series."

**Moving seasonality present at the one percent level.

***Because the series contains one or more zero values, a multiplicative adjustment is not applicable.

Index Crime	Additive Assumption: Stable F	Additive Assumption: % Contribution Seasonal	Additive Assumption: % Contribution Irregular	Multiplicative Assumption: Stable F	Multiplicative Assumption: % Contribution Seasonal	Multiplicative Assumption: % Contribution Irregular
United States 1970-82	110.8*	86.6%	12.1%	283.2	86.3%	12.6%
Illinois 1972-81	10.0	20.5	79.0	9.9	17.0	82.5
Cook County 1972-79	3.8	27.6	71.9	4.4	30.9	68.6
Chicago 1972-81	2.6	17.2	82.4	3.1	16.4	83.2
Small-city counties 1972-79	1.7	27.6	72.1	**	**	**
Rural counties 1972-79	12.7	36.2	63.4	11.2	25.4	74.3
New York City 1973-82	10.9	40.3	59.4	10.8	23.1	71.8

Table 6. Census X-11 results in selected forcible rape series.

*Moving seasonality present at the one percent level.

**Because the series contains one or more zero value, a multiplicative adjustment is not applicable.

the survey of the states, see California, Delaware, Kentucky, Maine, and North Carolina.

Robbery. Robbery is a violent Index crime that also involves the taking of property. We have suggested above ("Is crime seasonal or is reported crime seasonal?") that the seasonal fluctuation found in some robbery series may reflect a tendency for less serious robberies (attempts, for example) to become known to the police more often in the summer months. The analyses of Index robbery conducted by the Authority do not offer much illumination of this question (tables 7 and 8).

We would expect that robberies known to the police, especially less serious robberies occurring in cities, to fluctuate seasonally. The number of Index robberies reported to the FBI in the United States as a whole does fluctuate seasonally (table 7), as do Index robberies in New York City. However, none of the Illinois series we have analyzed are seasonal by the Plewes criteria, whether the state as a whole, Chicago, or various large and small jurisdictions (table 8). On the other hand, if we consider armed robbery to be relatively serious, we would expect that it would not fluctuate seasonally, and that is what the ten-year Boston data indicate.

Robbery victimizations (occurring to noncommercial victims aged 12 and over) do not fluctuate seasonally (table 7). Neither the weapon nor the relationship of the offender to the victim makes a difference in the seasonal fluctuation of robberies in the National Crime Survey.

For other analyses of robbery, see, in the "Review bibliography" below, Block (1979), Deutsch (1978), Deutsch and Alt (1977), Hay and McCleary (1979), FBI (1981), Ku and Smith (1977,1978), Marshall (1977), Michael and Zumpe (1983), US/BJS (1980). In the survey of the

Index Crime	Additive Assumption: Stable F	Additive Assumption: % Contribution Seasonal	Additive Assumption: % Contribution Irregular	Multiplicative Assumption: Stable F	Multiplicative Assumption: % Contribution Seasonal	Multiplicative Assumption: % Contribution Irregular
United States 1970-82	146.7	87.5%	10.4%	159.0	86.5%	11.4%
National Crims Survey 1973-79						
Total robbery	4.6	34.0	63.7	4.5	27.3	70.4
Robbery with a gun	4.3	38.4	61.2	4.6	30.8	68.9
Robbery without a gun	2.3	24.8	72.9	2.3	23.4	74.7
Robbery by a stranger	4.8	39.7	58.5	4.5	34.0	64.1
Robbery by acquaintance	0.4	31.6	68.0	0.5	25.8	73.8
Illinois 1972-81	10.9	32.3	65.0	10.9	28.9	69.0
Chicago 1972-81	6.4	26.3	71.6	6.6	21.7	76.8
New York City 1973-82	36.1	74.0	23.2	35.1	72.7	24.3
Boston 1966-75, armed	6.0*	36.8	60.0	5.6	41.2	55.4

Table 7. Census X-11 results in selected robbery series.

*Moving seasonality present at the one percent level.

City	Population	Stable Seasonality F: Additive Assumption	Stable Seasonality F: Multiplicative Assumption
Chicago	3,005,072	6.4	6.6
Springfield	99,098	4.2	4.6
Joliet	78,165	1.2	1.1
Champaign	57,176	1.1	1.0
E. St. Louis	54,966	1.1	1.2
Quincy	42,048	1.9	*

Table 8. Bell-Canada F of stable seasonality: Index robbery, selected Illinois cities, 1972-1981.

*Because the series contains one or more zero values, a multiplicative adjustment is not applicable.

states, see California, Delaware, Kentucky, Maine, and North Carolina.

Aggravated assault. Assault is repeatedly used as an example in "Issues in the seasonality of crime," above. We suggest there that assault victimization may not vary seasonally, but that less serious assault may become known to the police more frequently in the summer months because it tends to be more public. The findings for assault in the Authority's seasonal analysis file (tables 9 and 10) do not, generally, conflict with this interpretation. However, neither do they give it strong support.

Among the Illinois cities we have analyzed for the 1972 to 1981 period (table 10), assault fluctuates with the seasons in Chicago, but not in the smaller cities. If we analyze fifteen years of data for Chicago, the presence of seasonal fluctuation is even clearer (table 9). For Cook County and for Illinois as a whole, the irregular contribution is too high to make any definitive statement about seasonality.

Boston data for firearm assaults over ten years show no seasonal fluctuation at all. New York City Index assault data, which include aggravated assault by any weapon, have a high stable seasonality F value, but, like the Illinois data, also have an irregular that is too high for a definitive statement.

There seems to be no question, however, of seasonal fluctuation in assault victimization, whatever the weapon or the relationship of the offender to the victim (table 9). For all components of assault, F values are low and the irregular contribution is high.

For other analyses of assault, see, in the "Review bibliography" below, Deutsch (1978), Deutsch and Alt (1977), Hay and McCleary (1979), FBI (1981), Marshall (1977), Michael and Zumpe (1983), Pittman (1964), US/BJS (1980). In the survey of the states, see California, Delaware, Kentucky, Maine, and North Carolina.

Is property crime seasonal?

Burglary. In the jurisdictions we have analyzed, burglary generally has a low F of stable seasonality (though not usually as low as homicide) and a high percent contribution of the irregular component (tables 11 and 12). However, it is strongly seasonal in the United States as a whole and in New York City. The contrast between New York City and Chicago is striking. Chicago X-11 results contain no hint of seasonality, while the same measures indicate strong seasonality in New York.

Index Crime	Additive Assumption: Stable F	Additive Assumption: % Contribution Seasonal	Additive Assumption: % Contribution Irregular	Multiplicative Assumption: Stable F	Multiplicative Assumption: % Contribution Seasonal	Multiplicative Assumption: % Contribution Irregular
United States 1970-82	94.7*	83.2%	15.2%	237.8	83.9%	14.6%
National Crims Survey 1973-79						
Total aggravated assault	7.5	31.0	68.4	7.4	29.9	69.5
Assault with a gun	1.8	22.8	76.8	1.7	16.3	83.4
Assault without a gun	9.4	47.3	52.0	9.4	45.8	53.6
Assault by a stranger	4.8	18.8	80.4	4.6	20.6	78.7
Assault by acquaintance	6.2	29.2	70.4	6.4	34.9	64.6
Illinois 1972-81	45.7	59.6	38.5	42.9	58.9	39.5
Cook County 1972-82	46.5	52.4	45.7	44.6	52.6	45.8
Chicago 1967-81						
Total agg. assault	109.1	70.0	27.3	104.5	70.3	27.4
Assault with a gun	60.9*	60.0	35.4	54.2*	57.4	38.7
Assault without a gun	89.2	65.6	31.5	90.6	65.0	32.4
New York City 1973-82	58.1*	62.6	37.0	60.6*	59.4	40.1
Boston 1966-75, gun	2.8	24.3	74.8	2.9	12.8	86.6

Table 9. Census X-11 results in selected assault series.

*Moving seasonality present at the one percent level.

City	Population	Stable Seasonality F: Additive Assumption	Stable Seasonality F: Multiplicative Assumption
Chicago	3,005,072	32.3	28.9
Springfield	99,098	6.0	7.0
Joliet	78,165	2.0	1.5
Evanston	73,278	4.6	5.6
Champaign	57,176	4.8	4.8
E. St. Louis	54,966	4.6	*
Schaumburg	52,083	0.7	*
Quincy	42,048	1.2	*
Carbondale	26,144	0.9	*

Table 10. Bell-Canada F of stable seasonality: Index aggravated assault, selected Illinois cities, 1972-1981.

*Because the series contains one or more zero values, a multiplicative adjustment is not applicable.

City	Population	Stable Seasonality F: Additive Assumption	Stable Seasonality F: Multiplicative Assumption
Chicago	3,005,072	5.4	5.2
Springfield	99,098	1.3	0.9
Joliet	78,165	6.2	6.5
Evanston	73,278	3.6	3.4
Oak Lawn	60,358	3.2	*
Champaign	57,176	7.3	7.2
E. St. Louis	54,966	3.1	*
Schaumburg	52,083	1.4	1.2
Quincy	42,048	2.2	2.2
Carbondale	26,144	1.3	1.3

Table 11. Bell-Canada F of stable seasonality: Index burglary, selected Illinois cities, 1972-1981.

*Because the series contains one or more zero values, a multiplicative adjustment is not applicable.

For other analyses of burglary, see, in the "Review bibliography" below, Block (1979), Deutsch (1978), FBI (1981), Ku and Smith (1977,1978), Marshall (1977), Schneider and Sumi (1977), US/BJS (1980). In the survey of the states, see California, Delaware, Kentucky, Maine, and North Carolina.

Larceny/theft. Of all the crime types we have analyzed over the years, Index larceny/theft seems to have the most consistently seasonal pattern. The F of stable seasonality is high not only in Chicago (table 13), but also in smaller cities (although the irregular contributions are often high). The F value is extremely high in the United States as a whole and in Illinois as a whole (table 14). In the United States, Illinois and New York City, seasonality contributes 80 to 90 percent of the month-to-month variation in the number of offenses known to the police. Cook County and those nonmetropolitan counties containing a city of 25,000 to 49,999 population also show seasonal fluctuation in the number of larceny/thefts known to the police.

Such a strong seasonal pattern can be useful in practical situations. In our experience, a good prediction of larceny/theft can often be made by knowing the seasonal pattern and little else.

For other analyses of larceny/theft, see, in the "Review bibliography" below, Deutsch (1978), FBI (1981), Lamp (1983), US/BJS (1980). In the survey of the states, see California, Delaware, Kentucky, Maine, and North Carolina.

Motor vehicle theft. Index motor vehicle theft (table 15) is seasonal in the United States as a whole and New York City, but not in any of the other jurisdictions we analyzed. As with burglary, the contrast between the lack of seasonal variation in Chicago and the strong seasonal variation in New York City is striking.

Index Crime	Additive Assumption: Stable F	Additive Assumption: % Contribution Seasonal	Additive Assumption: % Contribution Irregular	Multiplicative Assumption: Stable F	Multiplicative Assumption: % Contribution Seasonal	Multiplicative Assumption: % Contribution Irregular
United States 1970-82	99.6*	89.8%	8.1%	130.6	90.1%	7.8%
Illinois 1972-81	19.3	55.9	41.7	19.0	53.6	44.2
New York City 1973-82	46.6	78.0	18.4	42.7	78.5	18.5
Chicago 1972-81	5.4	32.0	65.7	5.2	28.8	69.4
Champaign County 1972-79	5.1	40.2	57.9	5.3	43.0	55.0
Kankakee County 1972-79	4.4	30.3	66.4	4.1	23.8	72.9
Macon County 1972-79	2.4	16.7	80.9	2.4	19.6	78.0

Table 12. Census X-11 results in selected burglary series.

*Moving seasonality present at the one percent level.

City	Population	Stable Seasonality F: Additive Assumption	Stable Seasonality F: Multiplicative Assumption
Chicago	3,005,072	39.9	36.9
Springfield	99,098	18.6	15.4
Joliet	78,165	10.9	12.1
Evanston	73,278	15.5	18.6
Oak Lawn	60,358	2.2	1.9
Champaign	57,176	7.4	6.2
E. St. Louis	54,966	10.8	9.3
Schaumburg	52,083	7.9	8.0
Quincy	42,048	8.5	9.1
Carbondale	26,144	2.4	2.6

Table 13. Bell-Canada F of stable seasonality: Index larceny/theft, selected Illinois cities, 1972-1981.

For other analyses of motor vehicle theft, see, in the "Review bibliography" below, Deutsch (1978), FBI (1981), and US/BJS (1980). In the survey of the states, see California, Delaware, Iowa, Kentucky, Maine, and North Carolina.

Index Crime	Additive Assumption: Stable F	Additive Assumption: % Contribution Seasonal	Additive Assumption: % Contribution Irregular	Multiplicative Assumption: Stable F	Multiplicative Assumption: % Contribution Seasonal	Multiplicative Assumption: % Contribution Irregular
United States 1970-82	267.1	91.5%	6.7%	291.0*	91.0%	7.0%
Illinois 1972-81	95.8	80.7	18.2	98.2	77.2	21.4
New York City 1973-82	59.1	80.2	17.1	65.3	79.5	18.0
Chicago 1972-81	39.9	60.1	38.3	36.9	54.4	44.1
Cook County 1972-79	75.1	79.3	19.4	71.1	74.9	23.4
Small-city coun. 1972-79	46.5*	70.9	25.8	56.6	67.3	29.2
Rural counties 1972-79	23.9*	53.6	43.2	29.3	50.5	46.9
Quincy 1972-81	8.6	32.3	65.4	9.0	32.3	64.8

Table 14. Census X-11 results in selected larceny/theft series.

*Moving seasonality present at the one percent level.

Index Crime	Additive Assumption: Stable F	Additive Assumption: % Contribution Seasonal	Additive Assumption: % Contribution Irregular	Multiplicative Assumption: Stable F	Multiplicative Assumption: % Contribution Seasonal	Multiplicative Assumption: % Contribution Irregular
United States 1970-82	167.9	90.0%	8.9%	172.5	90.4%	8.5%
Illinois 1972-81	13.0	39.1	59.6	12.7	38.1	60.5
New York City 1973-82	46.0*	74.8	23.0	39.2*	72.6	25.1
Chicago 1972-81	5.0	17.9	81.8	5.0	17.2	82.5
Small-city coun. 1972-79	6.9	41.4	55.8	7.1	38.9	58.6
Rural counties 1972-81	12.8	38.0	59.3	10.7	30.7	68.6
Quincy 1972-81	4.0	21.5	78.0	4.5	9.3	90.4

Table 15. Census X-11 results in selected motor vehicle theft series.

*Moving seasonality present at the one percent level.

Review bibliography

This is a comprehensive listing of analyses of seasonal fluctuation in crime, published and unpublished, concentrating on research from the 1960's to the present. (For earlier research, see Wolfgang's (1966) review.) It also includes a review of the findings of a previously unpublished survey of the states, conducted in 1979, in which governmental officials were asked to describe existing analyses of seasonal fluctuation of crime in the state.

The review does not include those unpublished analyses that the Authority has conducted over the years, and that have been covered earlier in this report: homicide in California, Canada, the United States, and Illinois; Index crime in the United States and in New York City; certain crimes in Boston; National Crime Survey robbery and assault estimates for the United States; and Index crime in Illinois and cities within Illinois.

The studies have not been chosen for their quality, and are not necessarily methodolo-

gically sound. They have been chosen only because they each deal with seasonality in crime. The results given for each study represent the findings of its author or authors, not the opinion of the author of this review.

For each study, the review specifies title and author, type of crime analyzed, time period and geographical area, method and criteria used, as well as the findings. The included studies vary greatly in method, crime analyzed and its definition, and even in the author's definition of seasonal fluctuation. It is incumbent upon the reader, therefore, to take characteristics of each study into account when making a decision as to its contribution to our knowledge on the subject.

The reader should be especially cautious regarding the length of the time series analyzed in each study. As we discuss above (see the section, "Length of series") the results of any analysis of fewer than seven years should be read with skepticism. Many of the studies in this section utilize fewer than seven years of data, and some base their conclusions on one year only.

The reader may not be familiar with some of the terms used in the review bibliography under the categories "method," "criteria," and "results." These terms are explained in the report, How to Handle Seasonality (Block, 1983), which also contains, for those who want more detail, an annotated bibliography of the seasonal analysis literature. For the purpose of understanding the present report, however, the following may be helpful:

Census X-11	A method of analyzing seasonality, developed by the U.S. Census, and widely used from the 1950's to the present. It divides a series into three components: seasonal, trend/cycle, and irregular.
Stable Seasonality F	A statistic generated by the X-11: the ratio between the seasonal and irregular components. See "Length of series," above.
Contribution of the Irregular	A statistic generated by the X-11; the percent contribution of the irregular component to month-to-month variation in the series. Contributions of the three components add to 100%. See "Length of series," above.
Plewes rule-of-thumb	Criteria for the presence of seasonal fluctuation in a series. If the irregular contributes 30 percent or more of the total month-to-month variation, the decision should be "no stable seasonality," regardless of the F value. If the percent contribution is 25 to 29 the F value needs to be at least 15, and if the percent contribution is 15 to 24 the F value needs to be at least 2.41 for the series to be considered seasonal.
Bell-Canada	A short version of the Census X-11, used as a screener.
Stochastic model	An approach to seasonal analysis that emphasizes forecasts and the relationship of each observation to previous observations. Also known as "Box/Jenkins" and "ARIMA."
Autocorrelation	Describes the relationship of observations within a series. In a seasonal series, observations 12 months apart are correlated. In a successful model, the residuals will not be autocorrelated.
ARIMA model	A stochastic model of the form (p,d,q)(sp,sd,sq) where p = degree of autoregressive process, d = degrees of differencing, q = degree of moving average process, and sp, sd, and sq indicate seasonal autoregressive, seasonal differencing, and seasonal moving average process, respectively. If the second term is not in the model, it is not a seasonal model.

Block, Carolyn Rebecca
1979 Descriptive Time Series Analysis for Criminal Justice Decision Makers: Local Illinois Robbery and Burglary. Chicago: Statistical Analysis Center, Illinois Law Enforcement Commission.

CRIME: Index Burglary, Index Robbery.
TIME, PLACE: 77 Illinois jurisdictions (counties, cities over 25,000 population, planning regions) 1972-1977. Note: series are six years long.
METHOD: Bell-Canada screener; Census X-11.
CRITERIA: Plewes rule of thumb for stable seasonality F and % contribution of irregular.
RESULTS: No seasonality in any series.

Block, Carolyn Rebecca and Richard L. Block
1980 Patterns of Change in Chicago Homicide: The Twenties, The Sixties, and the Seventies. Statistical Analysis Center, Illinois Criminal Justice Information Authority.

CRIME: Eighteen types of homicide (homicides attributed to white, black or Latin offenders, to young or older offenders, to male or females; homicides of victims by age, sex and race categories; homicides with a gun or another weapon; homicides precipitated by a fight or robbery; and combinations of these).

TIME, PLACE: Chicago, 1965-1976.
METHOD: Bell-Canada screener, Census X11, stochastic model.
CRITERIA: Plewes rule of thumb for stable seasonality F and % contribution of the irregular; simplest model with no significant autocorrelations in residuals.
RESULTS: No seasonality in any series.

Block, Carolyn Rebecca, Craig McKie, and Louise S. Miller
1983 Patterns of change over time in Canadian and United States Homicide. Policy Perspectives, Spring.

CRIME: Thirteen types of homicide (gun vs. not gun; offender age 15-24 vs. other; fight vs. robbery; family vs. acquaintance vs. stranger; "native" vs. other for Canada; black vs. other for Chicago; female vs. male victim).
TIME, PLACE: Chicago, 1965-1981; Canada 1961-1980, also each Canadian province for same time period.
METHOD: Bell-Canada screener, Census X-11.
CRITERIA: Plewes rule of thumb for F of stable seasonality and % contribution of irregular.
RESULTS: No seasonality in any series. Highest F of Canadian series, 2.18. Highest F of Chicago series, 4.23 (irregular contribution 60%).

Block, Carolyn Rebecca, Louise S. Miller, Richard Block, Douglas Hudson
1981 Explaining patterns of change over time in Chicago homicides with a gun. Manuscript. Statistical Analysis Center, Illinois Law Enforcement Commission.

CRIME: Homicide with a firearm (same types as above), handgun, rifle and shotgun registrations, gun and not-gun suicides, gun and not-gun aggravated assaults; also number of general assistance (GA) and aid to dependent children (ADC) cases.
TIME, PLACE: Homicide, Chicago 1965-1978; registrations, Chicago 1969-1980; suicide, Cook County 1963-1979; assault, Chicago, 1965-1978; GA and ADC, Cook County 1965-1979.
METHOD: Bell-Canada screener, Census X-11.
CRITERIA: Plewes rule of thumb for stable seasonality F and % contribution of irregular.
RESULTS: No homicide series was seasonal. Rifle and shotgun but not handgun registrations seasonal. Assault seasonal. Neither gun nor not-gun suicide seasonal. ADC seasonal after 1975, not before. GA not seasonal.

Deutsch, Stuart Jay
1978 Stochastic models of crime rates. International Journal of Comparative and Applied Criminal Justice 2(2):127-151.

CRIME: 7 Index crimes: homicide, forcible rape, robbery, aggravated assault, burglary, larceny/theft, and motor vehicle theft.
TIME, PLACE: Ten cities: Jan. 1966 to July 1975 for St. Louis, Portland, Los Angeles, Kansas City, Atlanta, Cleveland, Boston and Denver. Jan. 1970 to July 1975 (5 1/2 years) for Dallas and Cincinnati.
METHOD: Stochastic modelling.
CRITERIA: Simplest model with no significant autocorrelations in residuals.
RESULTS: Homicide and rape not seasonal in any city. ARIMA model for all other crimes in all cities was (0,1,1)(0,1,1).

Deutsch, Stuart Jay and Francis B. Alt
1977 The effect of Massachusetts' gun control law on gun-related crime in the city of Boston. Evaluation Quarterly 1(4, November):543-568.

Deutsch, Stuart Jay and Lu Ann Sims
1978 An identification algorithm for dynamic intervention modeling with application to gun control. Series no. J-78-29, Georgia Institute of Technology, Atlanta 30332. Mimeographed.

Hay, Richard A., Jr. and Richard McCleary
1979 Box-Tiao time series models for impact assessment: A comment on the recent work of Deutsch and Alt. Evaluation Quarterly 3(2,May):277-314.

Deutsch, Stuart Jay
1979 Lies, damn lies and statistics: A rejoinder to the comment by Hay and McCleary. Evaluation Quarterly 3(2, May):315 328.

CRIME: Homicide, assault with a gun, armed robbery.
TIME, PLACE: Boston, Jan. 1966 to Oct. 1975 for Deutsch and Alt, Hay and McCleary; Jan. 1966 to Sept. 1977 for Deutsch and Sims.
METHOD: Stochastic modelling.
CRITERIA: Simplest model with no significant autocorrelation in residuals.
RESULTS: All analysts agree that homicide is not seasonal. ARIMA models for assault and robbery found by Deutsch and Alt and Deutsch and Sims:(0,1,1)(0,1,1) Hay and McCleary, using the same method and data, find a (0,1,1)(0,0,1) model for assault, and log (0,1,1)(0,0,1) for robbery.

Edgerton, Julie, Linda Phelps, Karen Boley-Chang, Constance Osgood
1978 Ecology of Rape, Kansas City Metropolitan Area: Summary Report of the Rape Data Bank. Institute for Community Studies, University of Missouri, Kansas City. Report Prepared prepared for the Metropolitan Organization to Counter Sexual Assault.

CRIME: Rapes known to the police.
TIME, PLACE: Kansas City, Missouri, Kansas City, Kansas and Independence, Missouri, 1971 and 1975 (two years).
METHOD: Inspection of monthly data.
CRITERIA: Not specified.
RESULTS: "No definite seasonal pattern."

Federal Bureau of Investigation
1981 Crime in the United States: Uniform Crime Reports 1980. U.S. Department of Justice, Washington, D.C. 20525.

CRIME: Index offense rates, agencies reporting all 12 months of each year.
TIME, PLACE: 1971-1980, United States.
METHOD: Quarterly data, expressed as a relative crime rate (proportion of the first quarter, 1971). Ratio-to-moving-average for each quarter, averaged over ten years into a "seasonal index."
CRITERIA: Not specified.
RESULTS: Seasonal indices (see method), by quarter, are:

	Q1	Q2	Q3	Q4
Murder:	95.0	95.4	104.2	105.5
Forcible Rape:	85.9	101.5	116.8	95.8
Robbery:	97.5	89.7	101.0	111.7
Agg. Assault:	88.6	103.4	110.8	97.2
Burglary:	96.1	94.6	104.8	104.4
Larceny/Theft:	88.5	102.5	109.1	99.9
Motor Veh. Theft:	91.7	98.0	106.6	103.8

Ku, Richard and Bradford Smith
1977 First Year Evaluation of the Illinois Urban High Crime Reduction Program: Final Report. Manuscript. Abt Associates, Inc., Cambridge, Massachusetts.
1978 Second Year Evaluation of the Illinois Urban High Crime Reduction Program: Final report. Manuscript. Abt Associates, Inc., Cambridge, Massachusetts.

CRIME: Residential burglaries and robberies known to the police.
TIME, PLACE: Jan. 1972 to June 1978 in three Illinois cities: Peoria, Champaign and Joliet.
METHOD: Ratio-to-moving-average.
CRITERIA: Not specified.
RESULTS: Not specified, but subsequent analysis uses seasonally adjusted data.

Lamp, Rainer
1983 Jahreszeit und Kriminalitat. (Time of year and criminality) Paper presented at the International Congress on Criminology, Vienna. Max-Planck-Institut, Freiburg.

CRIME: Total crime, theft, rape, murder.
TIME, PLACE: Federal Republic of Germany, 1971-1980.
METHOD: Census X-11, stochastic modelling.
CRITERIA: % contributions of three components, accuracy of 1981 forecasts, simplest model with no significant autocorrelations in residuals.
RESULTS: Both methods agree: Murder not seasonal, but theft, rape and total crime are seasonal (strong moving average effect.)

Marshall, Clifford W.
1977a Application of Time Series Methodology to Crime Analysis. The Polytechnic Institute, 33 Jay St. Brooklyn, 11201. Law Enforcement Assistance Administration Grant #76-TA-99-0028
1977b The State Space Forecasting Technique Applied to Reported Crime Data. Supplement to 1977a, above.

CRIME: Robbery, aggravated assault, burglary, rape.
TIME, PLACE: "District 1" of Cincinnati, 1968-1974.
METHOD: Census X-11.
CRITERIA: Stable seasonality F of 2.41, "reasonable" graphs of three components.
RESULTS: Robbery seasonal with December consistently high. Aggravated assault seasonal (F = 2.78). Burglary and rape show no seasonal effects.

Michael, Richard P. and Doris Zumpe
1983 Sexual violence in the United States and the role of season. The American Journal of Psychiatry 140(7,July): 883-886.

CRIME: Rates of forcible rape, aggravated assault, robbery, murder.
TIME, PLACE: 1975-1979 (5 years): Alabama, Arizona, Georgia, Honolulu, Illinois, Los Angeles, Maine, New Mexico, North Carolina, Oregon, Puerto Rico, San Francisco, South Carolina, Tennessee, Texas, Utah.
METHOD: Harmonic analysis of crime rates and monthly mean temperature, separate analyses for each crime and place. Locations ranked by latitude.
CRITERIA: Significance of relationships between crime, latitude and temperature, Spearman rank correlation.
RESULTS: Assault seasonal in 12 locations, rape in 14, robbery in 5, murder in one. Latitude had no effect, but temperature was significant for assault and rape.

Pittman, David J. and William Handy
1964 Patterns in criminal aggravated assault. Journal of Criminal Law, Criminology, and Police Science 55:462-470.

CRIME: Aggravated assaults known to police.
TIME, PLACE: St. Louis, 1961. Note: one year only.
METHOD: Crosstabulations.
CRITERIA: Chi square.
RESULTS: No seasonal pattern in zero-level tables, no relation between indoor-outdoor location and season.

Schneider, Anne L. and David Sumi
1977 Patterns of Forgetting and Telescoping in LEAA Survey Victimization Data. Institute of Policy Analysis, 777 High Street, Suite 222, Eugene, Oregon 97401.

CRIME: Burglary victimization. National Crime Surveys city survey data.
TIME, PLACE: Eighteen cities (all available cities), ten months of 1972, 1973 or 1974 (one year per city).
METHOD: "Seasonal Distinctiveness" (range of mean monthly temperature for year of survey), recall patterns (telescoping and forgetting), and season in which survey was fielded.
CRITERIA: Correlations, inspection of graphs.
RESULTS: Seasons influence memory. More crimes are reported to victim survey interviewers in the summer. This is especially true of crimes that the victim had not reported to the police.

Stein, Donald P., Jay-Louise Crawshaw and Algrid R. Barskis
1967 Computer-Aided Crime Prediction in a Metropolitan Area. Technical Reports 1-202 and 1-202-A, The Franklin Institute Research Laboratories, Philadelphia.

CRIME: Part I offenses, five percent sample.
TIME, PLACE: Philadelphia SMSA, 1966. Note: one year only.
METHOD: Multiple regression. Other variables: weather, time of day, day of week, phase of the moon.
CRITERIA: Probability that a certain type of crime would occur, given that some crime did occur.
RESULTS: Criteria render results irrelevant.

United States, Bureau of Justice Statistics
1980 Crime and Seasonality. National Crime Survey Report SD NCS-N-15,NCJ-64818. Report written by Richard W. Dodge and Harold R. Lentzner, Crime Statistics Analysis staff, Center for Demographic Studies, U.S. Bureau of the Census.

CRIME: Household larceny (total and over and under $50), Personal larceny without contact (total and over and under $50), Residential burglary (total, forcible entry and unlawful entry), Motor vehicle theft, Assault (total, aggravated and simple), Noncommercial robbery. Crimes occurring in "series" (recurring offenses) not included. National Crime Survey incident-level data.
TIME, PLACE: United States, 1973-1977 (5 years). Persons aged 12 and over.
METHOD: Census X-11.
CRITERIA: Stable seasonality F of 10 or more. F 2.34 to 9.99 considered "merely indicative of seasonality." Percent contribution of seasonal component, and seasonal factors.
RESULTS: Household larceny F=51.75 (under $50, F=22.61;over $50, F=27.65); Personal larceny F=25.08 (under $50, F=43.50, over $50, F=15.36); Residential burglary F=22.98 (forcible entry F=7.04, unlawful entry F= 35.96); Motor vehicle theft F=7.09; Assault F=7.52 (aggravated F=5.80, simple F=4.79); Robbery F=2.15.

Warren, Charles W., Jack C. Smith and Carl W. Tyler
1981 Seasonal variation in suicide and homicide: A question of consistency. Unpublished manuscript. Public Health Service, U.S. Centers for Disease Control, Atlanta, 30333.

CRIME: Homicides, except for deaths resulting from legal intervention, and suicides of people over age 14, National Vital Statistics Mortality files.
TIME, PLACE: United States, 1969-1978.
METHOD: Periodic regression analysis (PRA). See Bliss (1958, 1970).
CRITERIA: Goodness of fit of PRA model.
RESULTS: Both homicide and suicide have significant monthly patterns within each year. For homicide, however, the pattern differs from year to year. December, for example, is a "peak month" in some years and a "trough month" in other years. Suicide is consistently high in the spring and low in the winter.

Wolfgang, Marvin E.
1966 Patterns in Criminal Homicide. New York: John Wiley & Sons

CRIME: Homicide, by race and sex of victim. Number of multiple suspects arrested for criminal homicide.
TIME, PLACE: Philadelphia, 1948-1952 (5 years); United States, 1930 and 1950.

METHOD: Proportion of total homicides, aggregated over the five years, occurring in each month.
CRITERIA: Inspection of table.
RESULTS: Rejected "hypothesis (of) a relationship between monthly or seasonal changes and variations in criminal homicide." For the United States, 1930 and 1950, September is a high month. Arrests of multiple suspects show greater seasonal fluctuation.

Survey of the states, 1979

In late 1979, the Illinois Law Enforcement Commission (predecessor to the Illinois Criminal Justice Information Authority) surveyed the states, asking the director of each criminal justice planning agency the following:

> Has anyone analyzed reported Index crimes in your state to determine whether they increase and decrease with the seasons? If so, we would appreciate a copy of the publication, or a summary of the findings if they have not been published.

We received responses from seventeen states and the District of Columbia. Six of these states said that they had not analyzed seasonality. The findings of the other respondents appear below.

Arizona

CRIME: Property crimes (burglary, larceny/theft and motor vehicle theft).
TIME, PLACE: Arizona, by county, 1975, 1976 and 1977 (three years).
METHOD: Mean across three years of number occurring in each month. Northern counties and Maricopa and Pima Counties analyzed separately.
CRITERIA: Inspection of graph.
RESULTS: Number of property offenses varies with the tourist season (northern counties high in summer months, Maricopa and Pima high in winter months).

California

CRIME: Each Index crime (willful homicide, forcible rape, robbery, aggravated assault, burglary, theft, motor vehicle theft).
TIME, PLACE: California, 1974-1978 (five years).
METHOD: Graphs of number of offenses known to police each month over the five year period.
CRITERIA: Inspection.
RESULTS: Homicide: decrease each February, many fluctuations from month to month. Rape: high July-October, low January-February. Robbery: high December, low May June. Assault: peak in summer or fall. Burglary: high in fall and winter, lowest in June. Theft: September low in all years, but pattern inconsistent for other months. Motor vehicle theft: no consistent pattern for all years.

Delaware

CRIME: Each Index crime.
TIME, PLACE: Total Delaware, and Sussex County, 1976-1978 (three years).
METHOD: Ratio-to-moving-average.
CRITERIA: Not given.
RESULTS: Homicide, burglary, and motor vehicle theft not seasonal. Rape, robbery, assault and larceny seasonal. Sussex County, which has a lot of tourism, has large seasonal fluctuations.

Iowa

CRIME: Motor vehicle theft, by type of vehicle.
TIME, PLACE: Iowa, 1975-1978 (four years).
METHOD: Graph of each type of motor vehicle theft.
CRITERIA: Inspection of graph.
RESULTS: Decrease of total motor vehicle theft in winter, due entirely to a decrease in theft of motorcycles.

Kentucky

CRIME: Each Index crime.
TIME, PLACE: Total state, 1973-1978 (six years). For homicide, 1978 only.
METHOD: Comparison of high and low months.
CRITERIA: Inspection.
RESULTS: 1978 Homicide high in late spring and summer. Rape consistently high in summer, low in January. Robbery highest in final quarter of year. Assault high May to September, low January, February. Breaking and entering slightly more frequent in second half of year. Larceny consistently high June to August. Auto theft high July to August, low Jan-February.

Maine

CRIME: Each Index crime.
TIME, PLACE: Maine, 1975-1978 (four years).
METHOD: Graphs of number of offenses known to police per month.
CRITERIA: Inspection.
RESULTS: Homicide, rape, robbery not seasonal. Assault high May-October. Burglary high October-November. Larceny/theft strongly seasonal, high June-October. Motor vehicle theft high July-October. Seasonality may be an artifact of population fluctuation. Total Index crime rates, counting residents plus transients in the denominator, is not seasonal.

New Jersey

CRIME: Violent Index crime (homicide, rape, robbery, assault); nonviolent Index crime (burglary, larceny/theft, motor vehicle theft).
TIME, PLACE: New Jersey, 1975-1977 (three years).
METHOD: Graphs of crimes known to the police.
CRITERIA: Inspection.
RESULTS: Violent crime not seasonal. Nonviolent crime high June-September, low January-February.

North Carolina

CRIME: Each Index crime.
TIME, PLACE: North Carolina, 1977-1978 (two years).
METHOD: Graphs of crimes known to the police.
CRITERIA: Inspection.
RESULTS: Murder and motor vehicle theft are not seasonal. Rape and aggravated assault: high in summer. Robbery, burglary high in December; larceny/theft high in August.

Virginia

CRIME: Personal Index crime (homicide, rape, assault); Property Index crime (robbery, burglary, larceny/theft, motor vehicle theft).
TIME, PLACE: Virginia, 1975-1978 (four years).
METHOD: Graphs of crimes known to the police.
CRITERIA: Inspection.
RESULTS: Personal crimes high June-August; Property crimes high July-August.

Washington

CRIME: Homicide.
TIME, PLACE: Washington, 1958-1962 (five years).
METHOD: Graph.
CRITERIA: Inspection.
RESULTS: No seasonal pattern.

Washington, D.C.

CRIME: Violent and Property Index crime.
TIME, PLACE: District of Columbia, 1977 and 1978 (two years).
METHOD: Graphs.
CRITERIA: Inspection.
RESULTS: Violent crime not seasonal. Property crime peaks in August.

Footnotes

* An earlier version of this paper was presented at the Max Planck Institut, Freiburg, Germany.

1. For a more complete discussion, see Block (1983) and the studies listed in its annotated bibliography. For a brief review of the statistics in table 1, see the introduction to the "Review bibliography."

2. Source: Chicago Police Department police files, collected by the Authority, the Center for Studies in Criminal Justice of the University of Chicago Law School, and Loyola University. Time series files maintained by the Authority.

3. These unweighted UCR data are the total offenses reported by those agencies that reported to the FBI in every month of the year in question. The number of reporting agencies increased over the 1973-1979 period from 7,106 (representing a population of 174,249,026) to 11,782 (representing 196,836,371). (Note that the added agencies were mostly small jurisdictions, serving cities and counties with smaller populations.) Thus, these data cannot be used to examine trends over time. However, because Census X-11 results indicate the behavior of the seasonal component of the series as opposed to the behavior of any "trend," whether real or artificial, the data can be used as one indicator of the presence of seasonal fluctuation. For another analysis of UCR data, see FBI (1981) in the "Review bibliography."

4. These victimization estimates from the National Crime Survey, provided by Richard L. Block and Wesley Skogan, represent the number of victimizations occurring in the NCS sample, corrected for underrepresentation of various population groups. The sample includes only noncommercial victims aged 12 and over. Victimizations in which the month of occurrence is unknown, such as "series" victimizations, are not included. These missing data account for fewer than one percent of the robberies or assaults in any month. Dodge and Lentzner conducted an analysis of NCS data for a shorter (five-year) period, but did not attempt to analyze crime categories comparable to police categories. See US/BJS in the "Review bibliography." As this report went to press, 1980 and 1981 NCS data became available. We plan to continue this analysis with more recent data, as they are available.

5. Victim survey data are not available for a seven-year period in any city, and there are some problems in utilizing the NCS to examine metropolitan areas.

6. Although suicide is not usually considered to be a crime, it is often discussed in conjunction with homicide. For a review of the literature on seasonal fluctuation in suicide, see Lester (1972). The Authority has analyzed Cook County suicides from 1963, and has found no evidence of seasonal fluctuation, either in suicides with a gun or without a gun. F values range from 0.5 to 1.8.

7. The Bell-Canada is a quick screener for seasonality, similar to the Census X-11 but allowing no user options. The Bell-Canada F value can be interpreted as the Census X-11 F (see "Length of series").

Bibliography

Bliss, C. I. (1958). Periodic regression in biology and climatology. Connecticut Agricultural Experimental Station Bulletin 615:1-56.

Bliss, C.I. (1970). Statistics in Biology. Vol. II. New York:McGraw-Hill.

Block, Carolyn Rebecca (1983). How to Handle Seasonality: Introduction to the Detection and Analysis of Seasonal Fluctuation in Criminal Justice Time Series. Statistical Analysis Center, Illinois Criminal Justice Information Authority.

Block, Carolyn Rebecca and Louise S. Miller (1982). Manual for the Pattern Description of Time Series. Chicago: Statistical Analysis Center, Illinois Criminal Justice Information Authority.

Block, Richard L. (1977). Violent Crime. Lexington, Massachusetts: D.C. Heath and Co.

Block, Carolyn Rebecca and Richard L. Block (1984). Crime definition, crime measurement, and victim surveys. Social Issues, Spring.

Block, Carolyn Rebecca, Craig McKie and Louise S. Miller (1983). Patterns of change over time in Canadian and United States homicide. Policy Perspectives 3 (Spring,2):121-180.

Brearley, Harrington Cooper (1932). Homicide in the United States. Chapel Hill: The University of North Carolina Press.

Coldren, James (1980). Aggregation problems in the analysis of Illinois statewide criminal justice data. Illinois Criminal Justice Information Authority.

Kallek, Shirley (1978). An overview of the objectives and framework of seasonal adjustment. Pp. 3-25 in Zellner (ed) Seasonal Analysis of Economic Time Series. Proceedings of the Conference on the Seasonal Analysis of Economic Time Series, September 9-10, 1976, U.S. Department of Commerce, Bureau of the Census, Economic Research Report ER-1.

LEAA (Law Enforcement Assistance Administration, 1972). San Jose methods test of known crime victims. Statistics Division Technical Report # 1, Washington, D.C.

Lester, David (1972). Why People Kill Themselves. Springfield, Illinois: Charles Thomas.

McCain, Leslie J. and Richard McCleary (1979). The statistical analysis of the simple interrupted time series quasi-experiments. Pp. 233-293 in Quasi-experimentation: Design and Analysis Issues for Field Settings, by Thomas D. Cook and Donald T. Campbell. Chicago: Rand McNally.

McCleary, Richard, and Richard A. Hay, Jr. (1980).Applied Time Series Analysis for the Social Sciences. Beverly Hills:Sage Publications.

Miller, Louise S. and Carolyn Rebecca Block (1983). Illinois Murder Victim Data, 1973 to 1981: Guide to Quality, Availability and Interpretation. Illinois Criminal Justice Information Authority.

President's Commission on Law Enforcement and the Administration of Justice (1967). The Challenge of Crime in a Free Society. U.S. Government Printing Office.

Quetelet, Adolphe (1842). A Treatise on Man and the Development of his Faculties. English translation 1968. New York: Burt Franklin.

Sutherland, Edwin H. and Donald R. Cressey (1978). Criminology. Philadelphia: Lippincott.

Sylvester, Sawyer F. (1982). Adolphe Quetelet: At the beginning. Federal Probation 46 (December,4):14-19.

Wolfgang, Marvin E. (1966). Patterns in Criminal Homicide. New York: John Wiley & Sons.

Alternatives to Time Series Analysis in Population Projections

John P. O'Connell, *Office of Financial Management*

Because criminal justice operations remain largely a state and local responsibility, criminal justice research and analysis is very pertinent at that level. Even in the best of environments research can be a challenging and difficult endeavor. However, at the state level, criminal justice research takes on the added dimensions of urgency and policy. These added dimensions make the research process more complex. One result of this is the problem of selecting the appropriate type of analysis for the question being addressed.

Two recent Washington State research projects exemplify the problematic environment that often serves as the beginning point for state-level research. Both of these projects dealt with efforts to forecast future events. The two research questions discussed in this report are a crime forecast and the state inmate forecast. As will be shown, addressing the issue of crime forecasts logically precedes the problem of forecasting the state inmate population.

Forecasting Crime

In the late 1970's it became evident that the skyrocketing crime patterns that Washington State--and the nation--had experienced over the past decade were showing indications of slowing down. Because the volume of crime is very closely related to the number of arrests made, which are in turn very closely related to number of cases filed in court by prosecutors and so on through the criminal justice system, the leveling off or decline in the rate of increase of crime would be significant news for the operation of all criminal justice agencies. The problem was to determine whether the observed deceleration in the crime pattern after 1974 was a short term pause or whether the crime pattern had in fact reached a turning point.

Due to the work in the early 1970's of Wellford and Christenson, it was a widely accepted fact, by the time our crime forecast was undertaken, that the change in the volume of the criminally "at-risk" group--i.e., young males roughly between the ages of 15 and 29--clearly impacted the number of crimes in a state. All things held constant, the greater the number of at-risk males in your state the greater the number of crimes.

In addition to the number of at-risk males in a state, the second major variable affecting crime volumes is the probability for persons within the at-risk group to commit crimes. This is measured by the crime rate, in our case the number of reported crimes per 1,000 at-risk males.

Washington State has a very active demography program, so isolating changes in the size of the at-risk group was relatively simple. However, identifying the reasons for the growing crime rate was much more problematic. The review of the literature showed that there were three current methods that attempted to project or explain crime rates. These are represented by the following works:
(1) "Parametrizing Age, Period, and Cohort Effects," 1978, by Thomas W. Pullum. This study shows the importance of each of the listed variables in relationship with U.S. delinquency rates between 1964 and 1973. The study utilized log-linear analysis.
(2) "Forecasting Crime Data: An Econometric Analysis," 1978, by James A. Fox. This is a sophisticated effort using a nonrecursive regression model that predicts crime and clearance rates. It uses such independent variables as the number of police, the change in the number of police over-time, unemployment rates, and change in the size of the black population.
(3) "Property Crime Rates in the United States: A Macrodynamic Analysis, 1947-1977, with Ex Ante Forecast for the Mid-1980's," 1980, by Cohen, Felson, and Land. Quantifying their opportunity theory these authors use an ex ante forecast. A forecast is ex ante when it retains the variables in the model but uses projections of these variables as the basis for estimating crime rates rather than the known values themselves.

Although these forecasts were all state of the art theoretically and methodologically, they did not explain or predict the apparent downturn in the crime pattern that had piqued our interest. The question still remained: Is this apparent downturn in crime rate a lead indicator for the leveling or decline in the crime pattern or is it just a random dip?

An alternative method of explaining the variation in the crime rate was found in the social diffusion model. Diffusion models were adapted from the physical sciences for the social sciences by Dodd (1955) and Coleman (1959). In the most general sense, diffusion models are based on the concept that there are basic forces that contribute to the adoption of a social behavior among a population. When a new behavior appears in the population and social forces are activated, the behavior begins to be adopted throughout the population. The spread of the behavior throughout a population may be partial or total.

Applications of the diffusion-type models to crime forecasting were first observed in recent efforts by Snow and LaSante (1980). However, their application uses a physical science model, which obscures the social nature of the problem. MacCorquodale and Pullum's 1974 study provides an application of diffusion models that is more pertinent because it is applied to a social situation. These authors' study applies the diffusion model to the evaluation of women's acceptance of birth control practices. Using the MacCorquodale and Pullum social diffusion study as a guide, the application of this procedure was generalized to crime rates.

The first assumption posits an external force and the second assumption posits an internal force. Using the diffusion model for this purpose involves two underlying assumptions that represent the driving social forces. The first assumption is that forces outside of the adopting group operate at a constant level until all of those who are going to adopt a new behavior eventually do so. An example of such external force would be the adoption of a new product that is influenced solely by a marketing effort. In the case of crime rates, the outside force is possibly represented by the change in social stability. As a society becomes more unstable (e.g., the legitimacy of the system is questioned or there is great social change), the tendency of a population to commit crime increases.

The second assumption of the diffusion model used in this analysis is that the adoption of a behavior is proportional to the fraction of the population that has accepted it at time t. As the proportion of acceptors increases, a nonacceptor is increasingly likely to encounter acceptors in daily interaction. Put simply, a force that causes behaviors to be adopted in a given population is the interaction of individuals within that population. This force of diffusion is best called the interaction effect.

For many processes, including criminality, acceptance may come from either of the two assumptions. To represent this, the two above-stated models are added together to create a mixed diffusion model. By adding these two effects one assumes that they are mutually exclusive. The mixed model appears as follows:

$$Y = W_1(N-X) + W_2X(N-X)/N$$

Where:

Y = the number of people who adopt criminal behavior during any given time period

X = the cumulative number of people who adopt criminal behavior

W_1 = represents the external force that influences people to adopt criminal behavior

W_2 = represents the interaction among people that leads to the adoption of criminal behavior

N-X = represents the (potential) population that can be treated as potential acceptors. This captures the reality that potential acceptors are drawn from an ever-decreasing population

One of the important reasons for using the mixed diffusion model as a method for crime forecasts in Washington State is that the diffusion model follows the leveling off of the crime rate since 1976 more closely than do the other methods. Furthermore, the diffusion model provides a rationale for the leveling off of the crime rate after such a long period of increases (i.e., the growth curve is approaching its asymptote). Because of this new information the actual projection of crime volume is greatly simplified. Once it is determined what a reasonable projection of the crime rate might be (here it was predicted that it remains relatively constant for the immediate future), one merely needs to multiply the estimated at-risk group by the expected crime rates.

Combining the effect of W_1 and W_2 for the accumulation of the adoption of the behavior over time (X), it is obvious that the mixed diffusion model predicts that, once a force outside a population begins to exert its influence, there may be different periods of acceleration depending on the values of W_1 and W_2, but that eventually the adoption of a behavior will reach a limit. As it does so, the rate and the increase in the total number of adopters will slow down.

The basis of the mixed diffusion model is different from the other methods of predicting the crime rates discussed earlier. Where the methods discussed earlier focus on a specific set of prediction variables as a means of estimating a future phenomenon, the diffusion model focuses on the social process of two variables--outside influence and the influence of social interaction. In thus acceding to the diffusion model, the socially disorganizing events of the late 1960's and early 1970's will eventually run their course, culminating in a new equilibrium. Each period of social change, be it an aggravating change or an ameliorative change (i.e., the curves can either increase or decrease), can be followed by a period of relative stability or, because of new events, enter into a new epoch and a new period of diffusion of behavior.

Another thing to recognize about the diffusion of behavior is that the model does not require that all or even a majority of an eligible population adopt the behavior. A poorly diffused behavior would be represented by a relatively flat curve while a more successful diffusion would be represented by a much steeper curve.

Comparing the results of the mixed diffusion model to the previously discussed literature shows a major advantage of the mixed diffusion

model; it mathematically follows and theoretically predicts the leveling off of the crime rate in Washington State and, one could conjecture, the nation. Other methods of crime prediction failed to explain the leveling off of the crime rate. Therefore, even though these other models explain a remarkable percentage of variance, the mixed diffusion model appears to be a reasonable method of estimating the rate of crime.

As was noted earlier, once the asymptote of the diffusion curve is reached, it is then predicted that the rate of crime for the at-risk group will remain constant or decrease somewhat. For the state of Washington this appears to be the case. In the last couple of years the change in the rate of crime has stabilized. However, as the 1978 increase in the crime rate indicates, even if the crime rate has stabilized, we can still expect a wide degree of variation from year to year. The diffusion model would have us predict that on the average the crime rate will remain relatively stable. This points out a major shortcoming of the diffusion model; it is limited to predictions within one episode or period of change. Once the asymptote is reached one must look for emerging external forces of change that could lead to a new period of diffusion. As history has shown, a new epoch of change can be either in the positive or negative direction.

Graph 1 (which will follow at the end of the paper) shows the original forecast made for total reported crimes for Washington State using the mixed diffusion model. Although not extremely accurate, the results are fairly close. The crime rates have not increased at the rate they did in the 1970's, but they have still increased at least until 1982. However, since the late 1970's, crime, to a large extent, can be characterized as fluctuating widely without showing any true sign of direction. In general this is what the mixed diffusion model projected. Also in partial defense of the diffusion model the volume of the at-risk group did not increase nearly as greatly as was forecast in 1979. If actual population data were used, the mixed diffusion model would be somewhat closer.

Forecasting State Inmates

The need for improved prison population forecasts becomes critical as the need for correctional facilities and programs increases while available resources decrease. Increased competition for scarce resources requires a system that produces reliable forecasts of the size and composition of the prison population. To this end, this forecast takes into account the critical demographic and criminal justice system factors that produce changes in the prison population size. This forecast does not presume to provide an exact description of the future but, rather, makes a statement of what the future prison population will be if the crime, demographic, and criminal justice system factors follow their projected paths. The assumptions in this forecast are based upon the historical behavior of these critical factors and the expert consensus of key criminal justice decision makers.

This prison population forecast uses a computer simulation. A general flowchart of this system is presented in Figure A. Due mainly to data limitations, this forecast does not include all of the possible contributing factors that may explain changes in the prison population. However, we believe the most significant factors are included. Those factors which are included in the forecast are indicated on Figure A as solid lines and shaded areas. As can be seen on this flow chart, the prison population forecast includes key contributing factors such as demographic changes, superior-court felony convictions, the judicial decision to imprison, length of stay in prison, and the readmission of persons who fail once paroled.

The process by which the forecast was developed is unique. For the first time in this state a representative group of key criminal justice decision makers used a coordinated process for developing a prison population forecast. The catalyst for this involvement was the Governor's establishment of the Interagency Criminal Justice Work Group (GICJWG). One of the major charges of this group is to provide a coordinated interagency system for prison population forecasting. The involvement of the GICJWG went far beyond the normal managerial oversight that is the usual role when given such a technical task. The involvement of this group in the prison population forecast included review and evaluation of the methodology and data used, the establishment of the forecast operating assumptions, and close monitoring of the technical development.

The major outcome of this prison population forecast is a single indicator projection for FY 1982 to FY 1995. However, as a supplement to the single indicator projection, the forecast provides a wealth of detail in terms of the changing characteristics of the prison population over time. Therefore, it is possible to estimate not only the absolute change in the prison population, but also the changing composition of the prison population.

Equal in importance to the types of available detail in the forecast is the flexibility that is built into the computer model. Although the forecast produces a single-line estimate based on current operations and projected changes, it is also possible through alternative assumptions of the critical forecast factors to produce alternative forecasts. Changes in the system can be introduced that reflect various policy and system changes. The impact of these changes can be traced over time throughout the prison population. For example, the impact of the following questions could be evaluated:

- What if the violent crime rate continues to increase beyond the mid and later 1980's?

- What if the number of drug offenders being sent to prison stabilizes?
- What if the rate of parolees returning to prison drops significantly?
- What if the length of stay increases or decreases three years from now?
- What if the migration patterns change significantly?
- What if the probability of being convicted of a felony goes up 5%?
- What happens if all of the above happen at the same time?

Basically, the forecast operates according to the simplified formula:

Future		Present		New		Parole
Prison	=	Prison	+	Prison	+	Failures
Population		Population		Admissions		

- Prison Releases

Staff Document No. 72, "Programming Methodology for Calculating Washington State Inmate Population," provides a detailed review of the rationale for the operations of the computer model. For those interested in a more technical review there is the "Technical Programming Document."

The remaining two sections of this presentation show the most recent forecast and the monitoring and evaluation effort that is part of the forecasting process.

Monitoring and Evaluation

Once a forecast is published, the next step taken is to monitor and evaluate that forecast to determine if it and its assumptions are tracking correctly.

No forecasting effort can be presumed to provide an exact description of future events, and deviations between the forecast and actual events should be expected to occur. It is important to monitor and evaluate these deviations, because such deviations can provide valuable feedback regarding the reasons for current events. This new knowledge can then, in turn, be used to reassess the forecast model and assumptions.

The results of the monitoring and evaluation effort for the Fall 1981 forecast (that is, the forecast for FY 1982-1995) showed that for the first eight months of the forecast the maximum deviation of the forecast population from the actual population equaled 2.6 percent, an underestimation of 140 prisoners. However, within the next four months the deviation of the forecast from the actual more than doubled to 6.3 percent, representing an underestimation of 364 prisoners. This sudded and drastic change triggered a detailed analysis of the problem. From this it was concluded that the major portion of the sudden increase in prison population was due to:

o increased prosecutorial activity that led to a sharp increase in the number of new admissions to prison
o slower than forecasted rate of release of prisoners

Graph 2 and Table 1 (which will follow at the end of the paper) provide a summary overview of the monitoring effort for the Fall 1981 prison population forecast. A detailed review of this analysis can be obtained by reading OFM F&E Special Report No. 57.

The monitoring and evaluation of the Fall 1981 prison population forecast not only served as an early warning for the sudden and unanticipated change in the criminal justice system; it also proved to be a major source of information for updating the prison population forecast assumptions for the Fall 1982 prison population forecast.

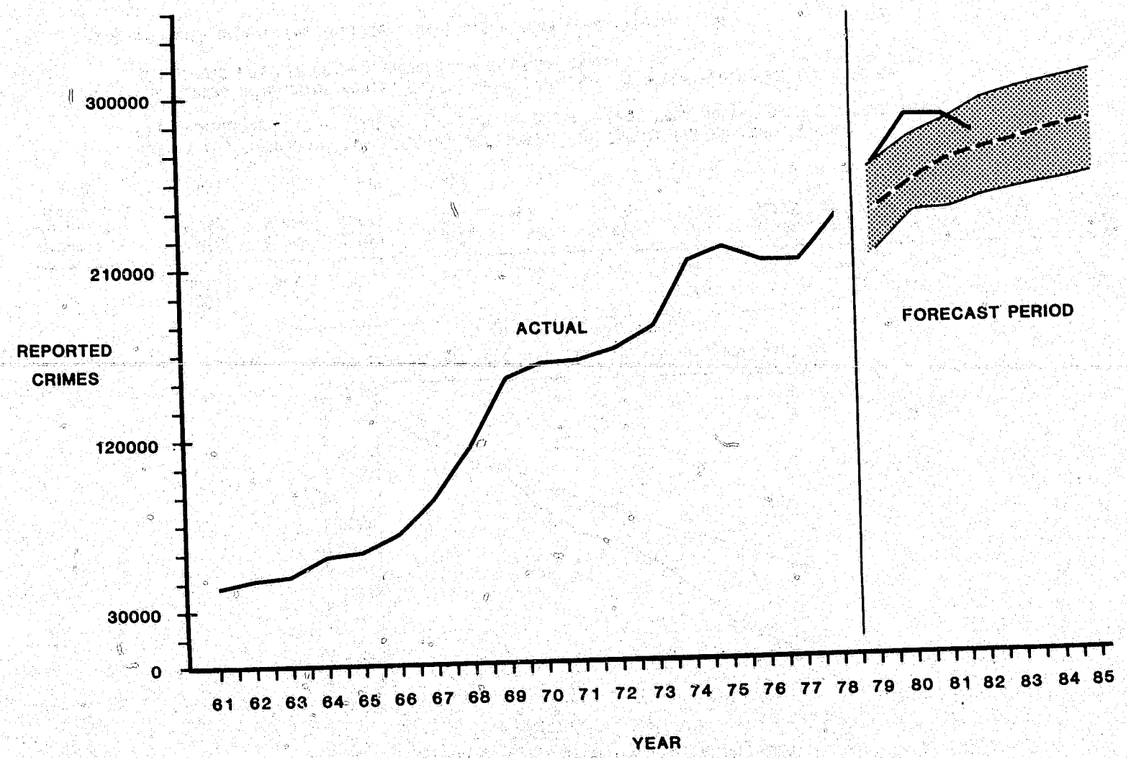
Findings

Findings for the FY 1983-FY 1996 prison population forecast, also known as the Fall 1982 Forecast, are presented in this section. See charts 3 and 4 and Tables 11 and 12 (which will follow at the end of the paper). The Annual Forecast reports on the expected number of prisoners as of the end of each fiscal year (i.e., June of each year) and the annual number of expected admissions and releases for fiscal years 1983-1996. This information is best suited for long-run issues such as capital planning and long-range criminal justice system planning. The Monthly Forecast, which is provided in the full forecast document, provides monthly admissions, releases, and population data for the fiscal years 1982, 1983, 1984, and 1985. The monthly information is most appropriate for shorter range efforts such as budget preparation, program planning, and forecast monitoring. Also provided in the full forecast document is a section showing the changes in the inmate population by crime type over the forecast period.

The Fall 1982 prison population forecast does not include the impact of two anticipated criminal justice system changes. First, the Fall 1982 prison population forecast does not include the impact of the recommendations of the Sentencing Guidelines Commission. However, the Fall 1982 prison population forecast, representing the current criminal justice system practices, is used by the Sentencing Guidelines Commission to provide a baseline of comparison for their recommendations. These recommendations are presented in a separate document prepared by the commission. Second, the Fall 1982 prison population forecast does not include the impact of the early release effort of the Board of Prison Terms and Paroles. At the time that the forecast was being developed, information regarding the magnitude of the early release effort was not available. The impact of the early release effort will be included in the monitoring phase of the forecast.

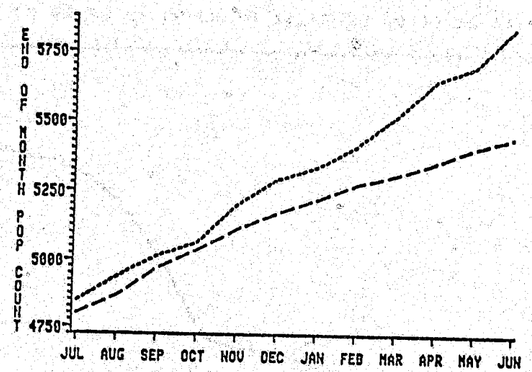
Graph 1

ACTUAL AND FORECAST REPORTED CRIMES
STATE OF WASHINGTON: 1961 to 1985



Graph 2

WASHINGTON PRISON POPULATION *
1981–1982



	1981 July	Aug	Sep	Oct	Nov	Dec	1982 Jan	Feb	Mar	Apr	May	1982 June
Actual:	4,848	4,939	5,018	5,068	5,306	5,294	5,342	5,418	5,525	5,654	5,706	5,845
Forecast	4,797	4,868	4,970	5,041	5,117	5,175	5,226	5,278	5,313	5,358	5,411	5,450
Difference F-A:	-51	-71	-48	-27	-89	-119	-116	-140	-212	-296	-295	-395

- THE PRISON POPULATION FORECAST UNDERESTIMATED THE ACTUAL PRISON POPULATION. THE UNDERESTIMATION BECAME SIGNIFICANT DURING THE LAST FOUR MONTHS OF THE FISCAL YEAR.
- THE STRONG UPSWING OF ADMISSIONS TO PRISON IS LARGELY EXPLAINED BY AN INCREASE OF ADMISSIONS FOR SEX CRIMES, ROBBERY, AND OTHER CRIMES.

*End of month population includes institutions and work release.

CONTINUED

1 OF 2

TABLE 1

BREAKDOWN OF THE DEVIATION BETWEEN FORECAST AND ACTUAL PRISON POPULATION BY MAJOR FORECASTING COMPONENTS FY '82

		Forecast	Actual	Difference	Percent Difference
New Admissions	Male	1,496	1,616	-120	-8.0%
	Female	104	103	+1	+0.9%
	Total	1,600	1,719	-119	-7.4%
Return Admissions*	Male	612	699	-87	-14.2%
	Female	26	23	+3	+11.5%
	Total	638	722	-84	-13.2%
Releases	Total	1,508	1,383	+125	+8.2%
Total (Admissions-Releases)		730	1,058	-328**	+44.9%

- THE TOTAL DIFFERENCE BETWEEN ACTUAL AND FORECAST FOR FY1982 EQUALS THE SUM OF THE TOTAL DIFFERENCES FOR EACH OF THE MAJOR FORECASTING COMPONENTS, NEW ADMISSIONS, RETURN ADMISSIONS, AND RELEASES. (119 + 84 + 125 = 328)
- THE DIFFERENCE BETWEEN THE FORECAST AND THE ACTUAL PRISON POPULATION FORECAST IS EXPLAINED BY SIGNIFICANT DEPARTURES FROM THE EXPECTED IN EACH OF THE MAJOR FORECASTING COMPONENTS
  - NEW ADMISSIONS ACCOUNT FOR 36% OF THE TOTAL DIFFERENCE
  - RETURN ADMISSIONS ACCOUNT FOR 26% OF THE TOTAL DIFFERENCE
  - RELEASE ADMISSIONS ACCOUNT FOR 38% OF THE TOTAL DIFFERENCE

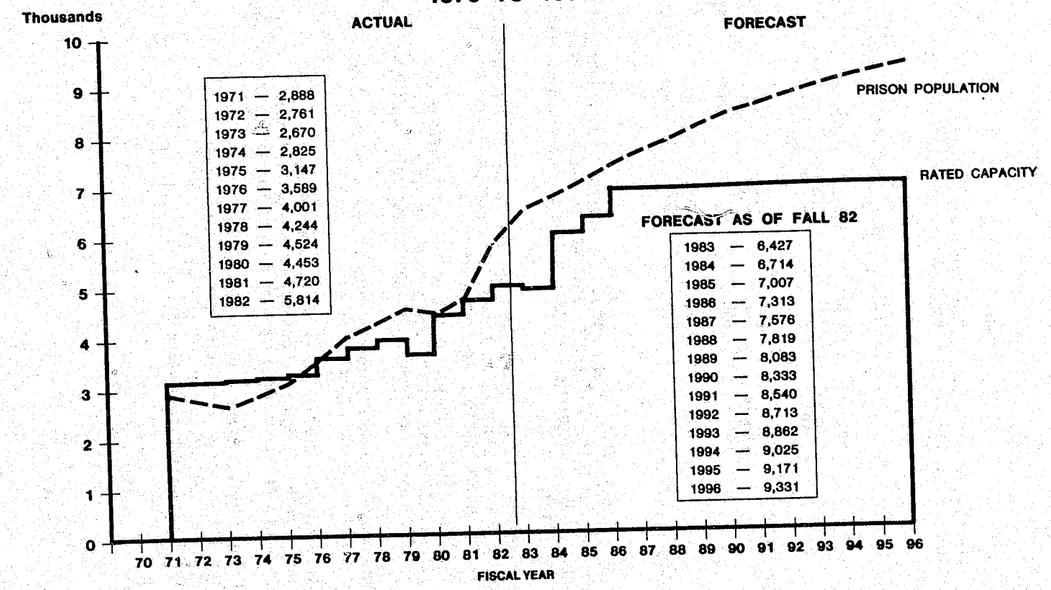
*Return admissions include parole returnees who are processed both through the courts and the Parole Board.

**Technical note: The difference of 67 inmates between the end of the comparisons (Graph 1) and the admissions versus releases analysis shown here can be accounted for by one or more of the following reasons.

(1) Relatively large changes in daily counts
(2) The forecast not accounting for non returning escapees or re-sentenced prisoners
(3) The forecast not accounting for special prisoners

CHART 3

TOTAL PRISON POPULATION AND RATED CAPACITY*
1970 TO 1996



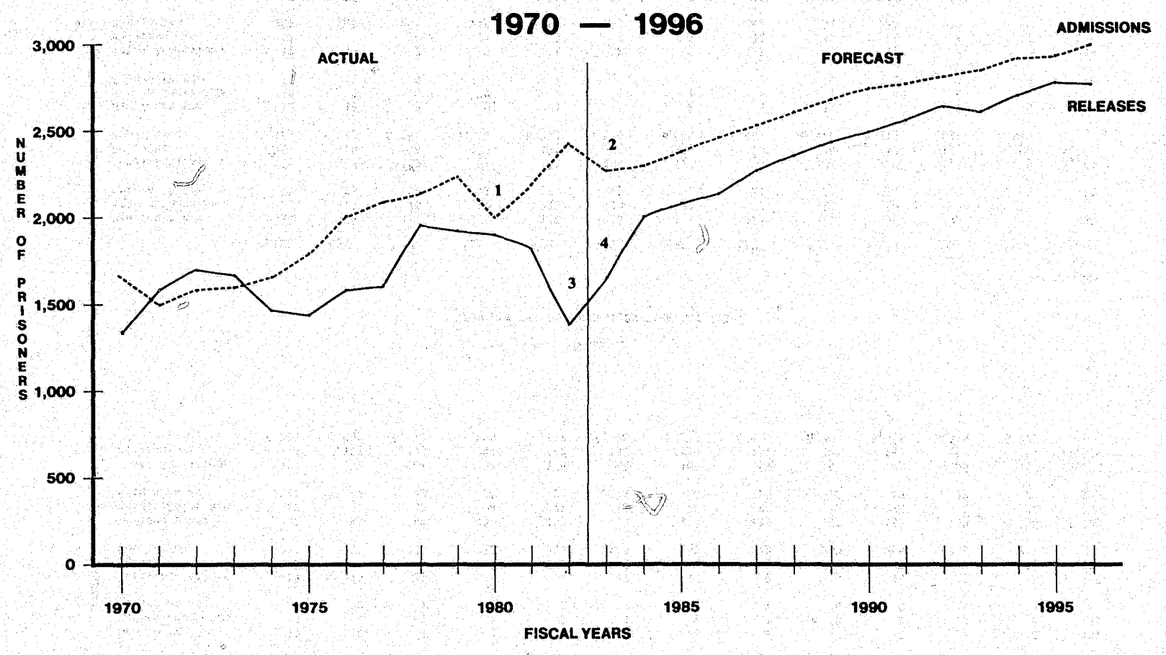
* PRISON POPULATION INCLUDES ALL STATE INMATES, BOTH IN PRISON AND ON WORK RELEASE AS OF JUNE 30TH.
RATED CAPACITY INCLUDES PRISON AND INMATE WORK RELEASE BEDS.

• PROJECTED PRISON CAPACITY INCLUDES THE FOLLOWING ANTICIPATED NEW BEDS.
.. DOUBLE CELLING AT THE CORRECTION CENTER. THIS WILL ADD 150 BEDS OCT. 1983, FEB. 1984, AUG. 1984, AND OCT. 1984 FOR A TOTAL OF 600 BEDS.
.. A 500 BED PRISON AT MONROE COMING ON LINE JULY 1984.
.. 200 BEDS AT THE McNEIL ISLAND FARM IN JULY 1984.
.. A 500 BED PRISON AT CLALLAM BAY IN NOVEMBER 1985.
.. A GRADUAL INCREASE OF APPROXIMATELY 40 WORK RELEASE BEDS AS PRISON RELEASES INCREASE DURING THE NEXT BIENNIUM.

OFM/FSE/
DATE:12/82 -16-535-

CHART 4

ANNUAL ADMISSIONS VERSUS RELEASES
1970 — 1996



1. ADMISSIONS DIP BECAUSE OF A 5% DROP IN THE JUDICIAL DECISION TO IMPRISON.
2. ADMISSIONS ARE EXPECTED TO DIP IN FY 1983 BECAUSE OF THE CURRENT DECREASE IN THE NUMBER OF REPORTED CRIMES.
3. RELEASES REACH A TWELVE YEAR LOW BECAUSE THE EARLY RELEASE PROGRAMS OF FY 1980 AND FY 1981 REDUCED THE POOL OF RELEASEES IN FY 1982 AND FY 1983.
4. EXPECTED RELEASES INCREASE RAPIDLY IN RESPONSE TO THE SURGE OF ADMISSIONS IN FY 1981 AND 1982.

OFM/FAE
DATE: 11/82 -OFMC535-

TABLE 11

MONTHLY PRISON POPULATION FORECAST FY83

		July	Aug	Sept	Oct	Nov	Dec	Jan	Feb	Mar	April	May	June	Fiscal Year Totals
MALE	New Court Adm	135	133	132	133	132	133	130	129	138	136	132	130	1,593
	Return Court Adm	22	22	27	23	25	24	24	23	22	20	23	22	277
	Return P B Adm	22	22	27	23	25	24	24	23	20	19	22	22	273
	Total Admissions	179	177	186	179	182	181	178	175	180	175	177	174	2,143
	Releases	105	118	124	132	130	125	128	128	113	146	147	151	1,547
	Population	5,660	5,719	5,781	5,828	5,880	5,936	5,986	6,033	6,100	6,129	6,159	6,182	
FEMALE	New Court Adm	12	10	8	9	8	11	6	6	15	14	7	6	112
	Return Court Adm	0	0	5	1	2	1	2	1	0	0	1	0	13
	Return PB Adm	0	0	2	1	2	1	2	1	0	0	1	0	10
	Total Admissions	12	10	15	11	12	13	10	8	15	14	9	6	135
	Releases	8	11	9	15	12	5	7	9	8	7	10	16	117
	Population	231	230	236	232	232	240	243	242	249	256	255	245	
TOTAL	New Court Adm	147	143	140	142	140	144	136	135	153	150	139	136	1,705
	Return Court Adm	22	22	32	24	27	25	26	24	22	20	24	22	290
	Return PB Adm	22	22	29	24	27	25	26	24	20	19	23	22	283
	Total Admissions	191	187	201	190	194	194	188	183	195	189	186	180	2,278
	Releases	113	129	133	147	142	130	135	137	121	153	157	167	1,664
		5,891	5,949	6,017	6,060	6,112	6,176	6,229	6,275	6,349	6,385	6,414	6,427	

OFM/F&E
12/82

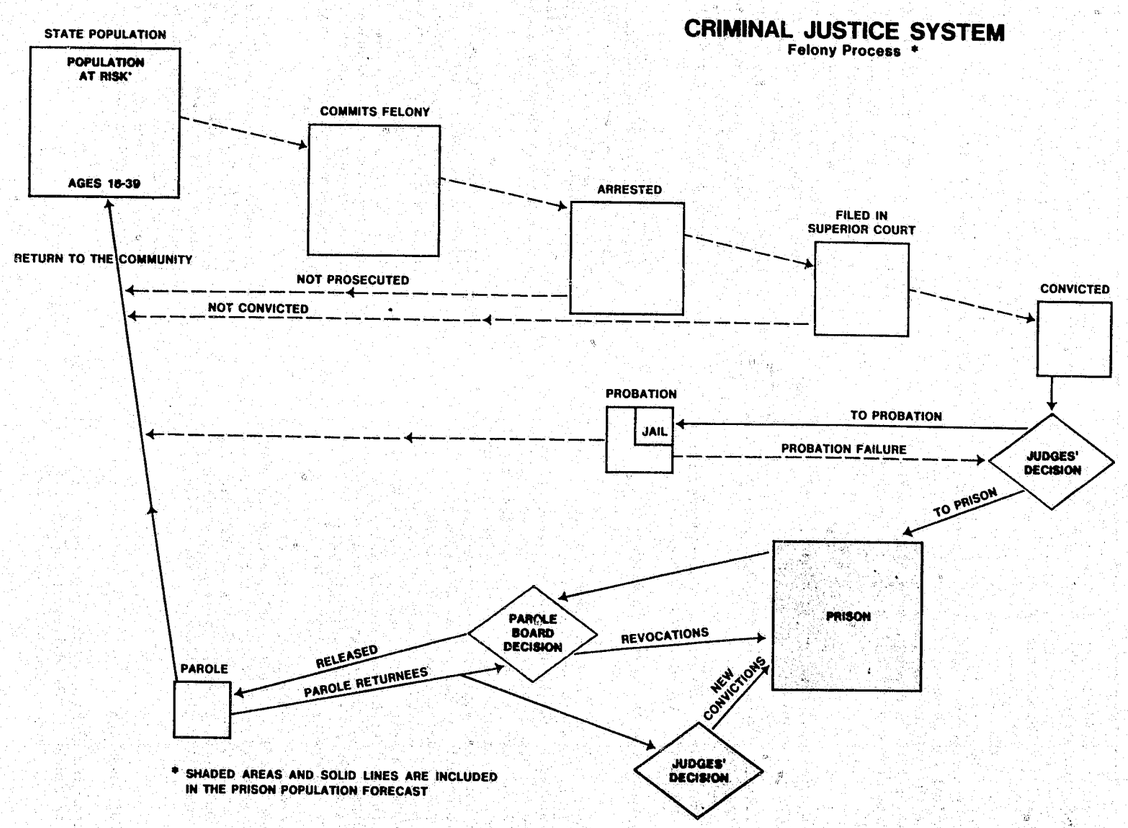
TABLE 12

MONTHLY PRISON POPULATION FORECAST FY84

		July	Aug	Sept	Oct	Nov	Dec	Jan	Feb	Mar	April	May	June	Fiscal Year Totals
MALE	New Court Adm	136	136	135	135	132	136	131	128	137	136	132	132	1,606
	Return Court Adm	23	21	28	22	25	24	23	22	23	23	22	21	277
	Return PB Adm	26	23	24	22	23	25	22	21	23	21	20	20	270
	Total Admissions	185	180	187	179	180	185	176	171	183	180	174	173	2,153
	Releases	184	126	143	140	151	155	155	164	164	187	159	159	1,887
	Population	6,183	6,237	6,281	6,320	6,349	6,379	6,400	6,407	6,426	6,419	6,434	6,448	
FEMALE	New Court Adm	12	10	9	9	8	11	6	6	15	15	8	6	115
	Return Court Adm	0	0	5	0	2	1	2	1	2	1	0	0	14
	Return PB Adm	1	0	3	0	2	0	1	0	2	1	0	0	10
	Total Admissions	13	10	17	9	12	12	9	7	19	17	8	6	139
	Releases	11	7	6	11	9	9	9	12	16	6	9	13	118
	Population	247	250	261	259	262	265	265	260	263	274	273	266	
TOTAL	New Court Adm	148	146	144	144	140	147	137	134	152	151	140	138	1,721
	Return Court Adm	23	21	33	22	27	25	25	23	25	24	22	21	291
	Return PB Adm	27	23	27	22	25	25	23	21	25	22	20	20	280
	Total Admissions	198	190	204	188	192	197	185	178	202	197	182	179	2,292
	Total Releases	195	133	149	151	160	164	164	176	180	193	168	172	2,005
	Total Population	6,430	6,487	6,542	6,579	6,611	6,644	6,665	6,667	6,689	6,693	6,707	6,714	

OFM/F&E
12/82

Figure A


CRIMINAL JUSTICE SYSTEM
Felony Process *
STATE POPULATION
POPULATION AT RISK*
AGES 18-39
COMMITS FELONY
ARRESTED
FILED IN SUPERIOR COURT
CONVICTED
RETURN TO THE COMMUNITY
NOT PROSECUTED
NOT CONVICTED
PROBATION
JAIL
TO PROBATION
PROBATION FAILURE
JUDGES' DECISION
TO PRISON
PRISON
PAROLE BOARD DECISION
REVOCATIONS
RELEASED
PAROLE RETURNEES
PAROLE
NEW CONVICTIONS
JUDGES' DECISION
* SHADED AREAS AND SOLID LINES ARE INCLUDED IN THE PRISON POPULATION FORECAST

Establishing Causation: The Role of Epidemiological Evidence

Shanna Helen Swan, *University of California, Berkeley*

The use of epidemiological evidence to establish a causal connection between an exposure and a disease outcome is relatively new in litigation. Until recently, it was often sufficient for a physician to testify to such a connection on the basis of his clinical experience. However, in recent years, large numbers of suits have been brought against drug and chemical manufacturers that rest, in large part, on the epidemiological evidence of an association between exposure to a drug, device, or chemical and a disease outcome. These cases include:

- Occupational exposure to asbestos and mesotheliqma,
- Prenatal exposure to diethylstilbestrol and vaginal cancer,
- Use of intrauterine contraceptive devices and pelvic inflammatory disease,
- Tampon use and toxic shock syndrome.

What these examples have in common is the availability of a number of relevant epidemiological studies in which an association has been analyzed, together with the absence of definitive information about causal mechanism. In such a situation, the epidemiologist, trained to synthesize and evaluate a wide range of epidemiological and medical evidence, may be uniquely qualified to assess the available data with regard to a causal connection between exposure and disease. This analysis must be done with considerable care; epidemiology is also subject to a wide range of methodological pitfalls which may render epidemiological evidence misleading or even useless.

Alternative models of causation will be discussed briefly. Subsequently, causation will be contrasted with association, and the factors considered in deciding whether an association is causal will be discussed. Two examples drawn from recent litigation illustrate these points.

Models of Causation

"Cigarette smoking is the major cause of lung cancer in the United States."
(Surgeon General's Report, 1982).

Perhaps the most widely accepted example of a causal connection which was initially established on the basis of epidemiological evidence is the relationship between cigarette smoking and lung cancer. It is useful in this context to examine what is meant by the above statement of causation. Several alternative models are considered. These are illustrated in Figure 1.

- One Cause-One Effect Model (Figure 1A)
  While this model incorporates the most intuitive notion of causation, it is probably the least applicable in this context. Under the one cause-one effect model every smoker would develop lung cancer and every lung cancer case would be attributable to smoking. In fact, only about 10% of smokers develop lung cancer and less that 85% of lung cancer cases can be attributed to cigarette smoking.[1]

- Multi-Effect Model (Figure 1B)
  This model correctly includes the fact that cigarette smoking has been shown to be causally related to a number of disease outcomes. However, it is not certain that every smoker will develop one or more of these diseases.

- Multi-Cause Model (Figure 1C)
  This model assumes that there are several facors, such as smoking and asbestos exposure that independently give rise to lung cancer. While it is the case that both smoking and asbestos exposure are, independently, causal factors for lung cancer[2], not every smoker or asbestos worker will develop lung cancer.

- Component Cause Model (Figure 1D)
  Under this model a number of component causes or factors must all, simultaneously, be present in order for smoking to lead to lung cancer. While this is theoretically appealing, no known collection of factors completely predicts which smokers will develop lung cancer.

- Indirect Causal Model (Figure 1E)
  In this model, a factor, for example genetics, leads to an outcome such as smoking, which, inturn, leads to lung cancer. This model may also postulate that the primary cause (genetic in this case) would also lead to the outcome (lung cancer) absent the intermediate cause. This model has been suggested by those researchers who are critical of the evidence for a direct causal connection between smoking and lung cancer.

- Probabilistic Model (Figure 1F)
  In this model, smoking increases the probability of lung cancer. This probabilistic model is in closest agreement with current thinking. The complete model should include several independent causal factors in addition to smoking and several largely unknown component causes. The probability of developing lung cancer for smokers can be estimated as a function of known risk factors such as age and occupation. The exposure, "smoking," can also be made more specific, and the probability of disease will then depend upon duration of exposure (number of years smoking), level of exposure (number of packs per day), and perhaps tar and nicotine content and so on. Both epidemiological and statistical methods are utilized in obtaining these estimates.

The statement "Smoking causes lung cancer" is, therefore, a probabilistic one that can be quantified using epidemiology and statistics. All the other, deterministic models discussed are inadequate to describe the nature of the connection between smoking and lung cancer. It is likely that a similar analysis of the causal connection between any exposure and a medical outcome will utilize a probabilistic model.

Association and Causation

Association between an exposure and a disease outcome is simply a statistical dependence that does not, in itself, imply causation. This association can be measured by a variety of parameters. One of these, relative risk, is most commonly used by epidemiologists to measure association. The relative risk of a disease associated with a particular exposure is defined as:

$$\frac{\text{Probability (disease among the exposed)}}{\text{Probability (disease among the unexposed)}}$$

This parameter can be estimated in several ways, depending upon the particular study design which is used. These estimates of relative risk can be obtained specific to a particular subpopulation or stratum and pooled across strata to provide a summary measure of association.

Once a measure of association has been obtained it can be tested for statistical significance under a model that assumes random selection of cases and controls. However, in practice, subjects are seldom selected at random. The usual study design is observational; subjects select their treatment or exposure. In this nonrandomized case, great care must be taken to insure the comparability of study groups with respect to all variables related both to outcome and exposure. Epidemiologists must identify such variables, or confounders, and take care to control for them in the design and analysis phases of the study. Further care must be taken to insure that study groups are representative of the population from which they are assumed to have been drawn, particularly with respect to the exposure and outcome under study. Failure to achieve this goal results in an important form of bias known as selection bias, which can severely limit the value of the epidemiological findings. Thus, the evidence must be critically reviewed to assess the presence of all potential biases and confounders that may have distorted the results.

Once a statistically significant association has been found and the problem of bias and confounding has been addressed, the question of causation vs. association must be considered. The principles for establishing causation have evolved gradually since Koch's postulates were originally developed for infectious diseases. The following factors are now generally considered by epidemiologists when evaluating a possible causal connection between any disease outcome and an exposure:

- Strength of the Association: What is the magnitude of the relative risk estimate and what is its statistical significance?
- Consistency of the Association: Has the result been replicated, preferably in studies of alternate designs?
- Biological Gradient: Is there a dose-response relationship between exposure and outcome?
- Biological Plausibility: Is the finding consistent with current biological thinking? Have reasonable mechanisms been postulated?
- Experimental Confirmation: Is there supporting evidence from laboratory studies, such as animal bioassays or challenge-rechallenge studies in humans?
- Collateral Evidence: Are there supporting results from other disciplines

or other epidemiological studies, such as time trends in disease that correlate with patterns of product sales?

- Elimination of Other Causes: In establishing causation in any particular instance, have other known, independent causes of the outcome been considered and eliminated?

- Chronology: Does the outcome follow the exposure?

- Quality of Epidemiologic Studies: Have all major sources of bias and confounding been considered and eliminated?

It is not necessary that each of these criteria be met. Rather, these are guidelines for evaluating the available evidence for causation. However, failure to meet the last two conditions would result in serious doubts about the causal nature of an association. The Surgeon General's reports have carefully discussed each of these factors in regards to smoking and lung cancer in arriving at a conclusion of causation.[3] In what follows, two examples drawn from recent litigation will illustrate the application of these criteria.

Two Case Histories

1. Diethylstilbestrol and Vaginal Adenocarcinoma

In 1947 the synthetic estrogen, diethylstilbestrol (DES), was first marketed for prevention of spontaneous abortion. It was widely used in the 50's and 60's, despite inadequate testing for safety or efficacy. It has been estimated that between two and four million offspring were exposed to this drug prenatally,[4] and in 1953 Dieckman et al. established the lack of efficacy of this drug for prevention of spontaneous abortion or premature delivery.[5] However, it continued to be marketed for use in pregnancy until 1971. At that time, Herbst et al. observed an excess of a rare vaginal cancer which was previously extremely rare in young women.[6] A case-control study found documented first-trimester exposure to DES in seven of the eight cases and none of the thirty-two matched controls. The drug was subsequently withdrawn for use in pregnancy. Extensive research followed and the International Agency for Research on Cancer has concluded that, "Diethylstilbestrol is causally associated with the occurrence of cancer in humans".[7] The principles for establishing causation will be illustrated using this example.

- Strength of the Association: The relative risk estimates in this case are very large. One estimate is provided using the augmented odds ratio on the unmatched data, which yields an estimate of 325. Alternatively, one can compute a lower confidence limit for the relative risk using the matched data. This method provides 7.5 as a lower 95% confidence limit for the relative risk. In either case the significance probability is less than $1/10^6$.

- Consistency of the Association: This finding was replicated by Greenwald et al. using five cases and eight matched controls. None of the controls had a history of synthetic estrogen exposure, while all five cases were exposed.[8]

- Biological Gradient: No dose response for cancer has been observed in humans although a strong dose response for impaired fertility has been observed in mice. A clear response relationship between gestational age at exposure and risk of cancer has been documented.[9]

- Biological Plausibility: DES has been known, since 1938, to cause cancer in animals. It has been shown in pregnant mice to concentrate in the fetal reproductive tract. It is believed that DES interferes with the normal replacement of vaginal columnar epithelium by squamous epithelium. Forsberg has demonstrated in mice that treatment with DES inhibits this transformation.[10]

- Experimental Confirmation: Experimental studies in rats and mice support the human epidemiological findings. In one study transplacental exposure of rats to DES resulted in genital malignancies in 20-40% of the exposed and in none of the unexposed controls.[11] In addition, DES has been shown to induce mutations in the mouse lymphoma syystem.[12]

- Collateral Evidence: Evidence of a wide range of vaginal/cervical and uterine anomalies occurring in exposed female offspring support the causal connection between DES and clear cell adenocarcinoma. These abnormalities have also been produced in test animals and, like the cancers, are inversely related to gestational age at exposure. Similar findings have been found in males, both in humans and test animals.

- Elimination of other Causes: To date no other cause for clear-cell adenocarcinoma in young women has been suggested. Some have argued that threatened spontaneous abortion, which was the indication for DES use, may itself have increased the risk of cancer. However, no other study has ever associated high-risk pregnancy with vaginal adenocarcinoma. Furthermore, the frequency of high-risk pregnancies does not correlate with the temporal pattern of clear-cell adenocarcinoma.

- Chronology:

  The appearance of vaginal adenocarcinoma occurs betwen seven and thirty years after the exposure to DES.

- Quality of Epidemiological Studies:
  Diagnostic Bias: Diagnosis of initial cases was blind with respect to exposure. Furthermore, diagnosis was done by cytology and was well defined.

  Ascertainment Bias: Initial cases were ascertained before DES had been associated with vaginal adenocarcinoma. The Registry that has been established to collect cases of adenocarcinoma does so without regard to exposure.

  Recall Bias: Ascertainment of exposure was not limited to recall but required written confirmation of exposures. All records were equally reviewed, regardless of outcome.

  Confounding: Maternal history of pregnancy failure or bleeding was the only variable, other than DES exposure, related to the cancer outcome. Since these variables have had a uniform pattern of incidence both before and after the occurrence of this rare cancer, it is highly unlikely that they are causally related to the risk of cancer.

Therefore, it appears that exposure to DES during the first trimester is a significant causal factor in the subsequent development of clear-cell vaginal/cervical adenocarcinoma in young women.

2. Toxic Shock Syndrome and Tampon Use

The history of the toxic shock story and the methodological problems associated with it are well documented by Harvey, Horowitz, and Feinstein.[13] It is worth reviewing briefly:

In November 1978, Toxic Shock Syndrome (TSS) was described by Todd et al. in seven children ages 8 to 17, including three boys and five girls. The association with menstruation was not noted until one year later when Davis, of the Wisconsin Division of Health, noted three cases in menstruating women. Subsequently, the Center for Disease Control, Atlanta (CDC) was notified of large numbers of cases in young women and began carrying out surveillance for additional cases. In February 1980, Utah Health Department circulated a newsletter to physicians describing TSS and its association with menstruation. One month later, Schrock, of the Minnesota Health Department, published a letter in the Journal of the American Medical Association stating that TSS "appears to occur in women only, at an age of active menstruation, at or near the time of menstruation." In May 1980, CDC reported an association between tampon use and TSS, which was not related to brand or absorbency. National publicity followed, including speculation concerning a possible link to the new, highly absorbent tampon, Rely. Subsequently, reporting of TSS increased and self-reporting of cases was frequent. On September 19, 1980, CDC published the results of a second study that showed a statistically significant association between TSS and Rely brand tampons. On September 22, 1980, Rely was removed from the market and CDC stopped accepting direct case reports. On January 30, 1981, CDC published data showing a decrease in the number of reported cases, which they attributed largely to withdrawal of Rely and decreased tampon use.

An association between TSS and tampon use has been established and replicated. However, the epidemiological studies that demonstrated this association are subject to several sources of bias. The presence of these biases raises serious doubts about the causal nature of the association at this time.

Reporting Bias: In Minnesota, where surveillance for cases of TSS was active and ongoing, 83.8% of cases were tampon associated. The majority of the remaining cases reported to the CDC were the result of self-reports and passive surveillance; 92.4% of these cases were tampon related.[14]

Furthermore, while the overall pattern of cases collected by the CDC through passive surveillance shows a sharp peak in August and September of 1980 associated with the publicity concerning TSS and tampon use, the incidence of cases collected by Minnesota shows no such peak.

Recall Bias: Twenty-two cases of menstrually associated TSS were interviewed twice. Before the peak of the Rely publicity (7/11/80) only 32% reported using Rely. However, when interviewed after Rely was withdrawn from the market (10/1/80), 50% of the same women claimed to have been using Rely at the time of their ilness.[12]

Diagnostic Bias:
While physicians have previously diagnosed toxic shock in children, boys, and older women, publicity suggesting that TSS was a disease of menstruating women increased the likelihood that a diagnosis of TSS be made in a menstruating women.

It is possible that TSS is associated with tampon use only through its association with menstruation. If TSS is menstrually associated and if over 80% of menstruating women used tampons, irrespective of TSS, before the summer of 1980, the prevalence of tampon use observed in the Minnesota cases, which were those most likely to be free of bias, could well be explained. In this case, the excess observed in the CDC case series may well be the result of the biases mentioned. Therefore, at this time biases in the epidemiological data make it impossible to draw a conclusion of causation with respect to TSS and tampon use.

Epidemiological evidence can contribute uniquely to establishing causation between exposures and medical outcomes. In establishing the causal connection a number of conditions must be satisfied. In particular, great care must be taken to avoid biases and confounding, which may result in erroneous conclusions.

Notes

[1]Surgeon General's Report, Smoking and Health, 1982. U.S. Surgeon General's Office. The Health Consequences of Smoking: Cancer: A Report of the Surgeon General. Department of Health and Human Services, Public Health Services, Office on Smoking and Health, 1982.

[2]J. Saracci. An Analysis of the Epidemiological Evidence on the Asbestos-Smoking Interaction. Int. J. Cancer, 20:323, 1977.

[3]Surgeon General's Report, Smoking and Health, 1982. U.S. Surgeon General's Office.The Health Consequences of Smoking: Cancer: A Report of the Surgeon General. Department of Health and Human Services, Public Health Services, Office on Smoking and Health, 1982.

[4]O.P. Heinonen. Diethylstilbestrol in Pregnancy: Frequency of Exposure and Usage Patterns. Cancer, 31:573, 1973.

[5]W.J.Dieckman, M.E. Davis, L.M. Rynnkiewics and R. E. Bottinger. Does the Administration of Diethylstilbestrol during Pregnancy have Therapeutic Value? Am. J. Obstet. Gyn., 66 (5):1062-1081; 1953.

[6]A.L. Herbst and R.E. Scully. Adenocarcinoma of the Vagina in Adolescence: A report of Seven Cases Including Six Clear-Cell Carcinomas (so-called mesonephromas). Cancer, 25:745, 1970

[7]Int. Agency for Research on Cancer, IARC Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Humans, Vol 21:210, 1979.

[8]P. Greenwald, J.J. Barlow, P.C. Nosca and W.S. Burnett. Vaginal Cancer After Maternal Treatment with Synthetic Estrogens. N. Eng. J. Med., 285:7, 1971.

[9]A.L. Herbst, S.J. Robboy, R.E. Scully and D.C. Poskanzer. Clear-cell Adenocarcinoma of the Vagina and Cervix in Girls: Analysis of 170 Registry Cases. Am. J. Obstet. Gynecol., 119: 713, 1974.

[10]J. Forsberg. Late Effects in the Vaginal and Cervical Epithelia After Injections of Diethylstilbestrol in Neonatal Mice. Am. J. Obstet. Gynecol., 121: 108, 1975.

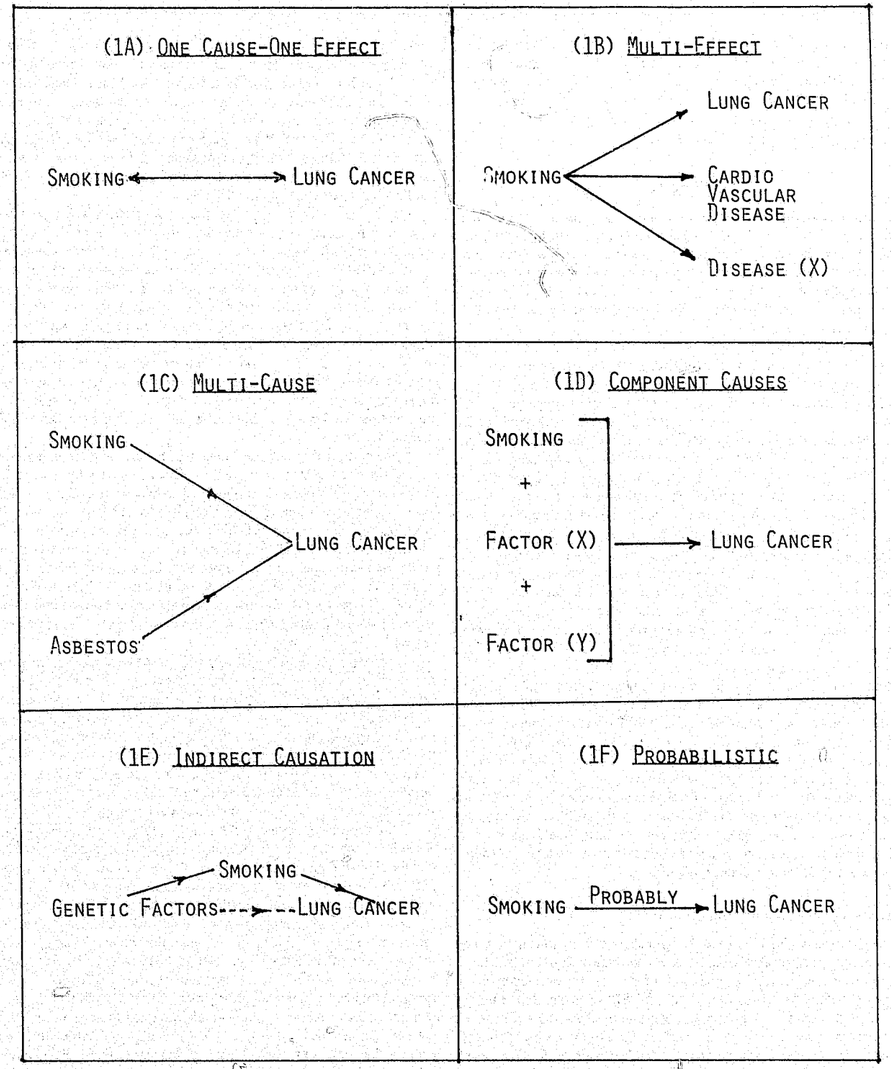
[11]H. Vorherr, R.H. Messer, U.F. Vorherr, S.W. Jordan and M. Kornfield. Teratogenesis and Carcinogenesis in Rat Offspring After Transplacental and Transmammary Exposure to Diethylstilbestrol. Biochem. Pharmacol., 28; 1865, (1979).

[12]D. Clive. A Linear Relationship between Tumorigenic Potency in vivo and Mutagenic Potency at the Heterozygous Thymidinekinase (TK +/-) Locus of L5178Y Mouse Lymphoma Cells Coupled with Mammalian Metabolism. In: D. Scott, B.A. Bridges and F.H. Sobels, eds, Progress in Genetic Toxicology, Amsterdam, Elsevier, North Holland, 1977.

[13]M. Harvey, R.E. Horwitz and A.R. Feinstein. Toxic Shock and Tampons; Evaluation of the Epidemiologic Evidence. JAMA, 248: 7, 1982.

[14]M.T.Osterholm and J.C. Forfang. Toxic Shock Syndrome in Minnesota: Results of an Active-Passive Surveillance System. J. Infect. Dis., 145: 458-464, 1982.

FIGURE 1: ALTERNATIVE MODELS OF CAUSATION



A Probabilistic Approach to Tracing Presumptions in the Law of Restitution

Michael O. Finkelstein*, *Columbia University*
Herbert Robbins†, *Columbia University*

The talk by Michael O. Finkelstein was subsequently published by Mr. Finkelstein and his co-author Herbert Robbins in Jurimetrics, Vol. 24, No. 1 at p. 65 (Fall, 1983). It appears here with the permission of the Section of Science and Technology of the American Bar Association.

ABSTRACT

Establishing an equitable lien or constructive trust over funds taken by fraud requires tracing. Where the misappropriated funds sought to be traced are mingled with other funds in an account and are subjected to withdrawals and deposits, the courts have used various presumptions to determine the extent of trust funds remaining in the account. It is suggested that probability theory yields rules and results that better accord with the legal theory underlying the remedies allowed in these cases.

It is well settled that a person who has been defrauded of money may assert an equitable lien or a constructive trust with respect to the funds in the possession of the wrongdoer. An equitable lien is a security interest in the funds that is limited to the amount of the claim. A constructive trust is an equitable ownership interest in the funds that is not so limited.[1] Important applications of this doctrine arise when the wrongdoer deposits the money in an account, either separately or mingled with other funds, and then becomes bankrupt. In that situation, if the claimant can "trace" the funds in the account he may recover on a priority basis, even though general unsecured creditors of the wrongdoer recover nothing because the estate available to them has been depleted.[2]

Prior to the Federal Deposit Insurance Corporation Act of 1933,[3] tracing claims frequently were asserted by defrauded depositors in failed banks who sought to recover their deposits on a preferential basis.[4] Although such cases are no longer common, tracing claims continue to arise in the wake of bank failures[5] and in a variety of other contexts, most of which involve claims against an insolvent debtor.[6] The current high rate of business bankruptcies promises to generate a new wave of such claims, perhaps at least equal to those of the 1930s.

In most cases involving tracing, the wrongdoer has commingled funds subject to the trust or lien with other funds. It is well settled that this does not destroy the claimant's equitable preference, but is likely to complicate tracing. We propose in this article to take a fresh look at two significant doctrines that determine the limits of tracing with respect to commingled funds. The first relates to the characterization of withdrawals from a commingled fund; the second to the determination of the lowest intermediate balance of a fund. In each case, highly artificial presumptions have been made to resolve uncertainty and to protect the claimant. We suggest a different approach: one that uses more plausible and neutral presumptions and then looks to mathematical probability based on the assumed facts to deduce the rules.

Withdrawals From a Commingled Fund

Once the right to an equitable remedy is established, the key to claimant's recovery is his ability to trace the funds. A rather elaborate body of law has developed on the requisites for tracing, and much of it favors the defrauded claimant.[7] A key presumption relates to the funds withdrawn from a commingled account. If funds withdrawn from the account are used for the wrongdoer's own purposes and dissipated, it is presumed that he selected his own funds for that purpose, so that claimant may trace the funds remaining in the account. Thus, as long as the balance in the account remains in excess of the amount of the claimant's funds there is a presumption against dissipation, and he may recover the full amount. If the balance in the account drops below the amount of the claimant's funds, the extent of depletion is determined by assuming that other funds were withdrawn first.

This rule derives from *Knatchbull* v. *Hallett*,[8] an English case in which a solicitor holding some bonds for a client sold them without authority and deposited the proceeds in an account with his bankers. He subsequently added some of his own funds to the account. Thereafter, he withdrew funds for his own purposes. When he died the balance in his account was (and apparently at all times had been) greater than the amount of his client's funds. The traceable assets would thus have been undissipated if the withdrawals had been deemed made from the solicitor's own funds, but substantially impaired if deemed made from the client's funds. Mr. Justice Frye held that he was bound by precedents (notably *Clayton's Case*[9]), to presume that the first funds deposited were the first withdrawn. Since the client's funds had been deposited first, he found that they had been substantially dissipated.

On appeal, the decision was reversed, two justices holding that when the solicitor made withdrawals for his own purposes, it must be presumed that he used his own funds. Jessel, Master of the Rolls, argued that if the monies were indistinguishable sovereigns in a bag, the fiduciary must be presumed to intend to take his own money when reaching into the bag. "His money was there, and he had a right to draw it out, and why should the natural act of simply drawing out the money be attributed to anything except to his ownership of money which was at his bankers."[10] Mr. Justice Baggalley agreed: "Can any reason be assigned why . . . I should not, as between myself and my *cestui que trustent* have the honest intention attributed to me of drawing against my own private funds and not against the trust funds, though it was the first paid in?"[11] This presumption, known as The Rule of Jessel's Bag, has been almost uniformly adopted by American courts.

In the particular context of *Knatchbull*, the presumption of honesty seems reasonable because the solicitor had not obtained money by wrongful means, nor had he clearly appropriated the funds to his own use. But most American courts, without considering the reasonableness of the presumption in other contexts, have simply adopted *Knatchbull* as superior to the first-in first-out rule in *Clayton's Case*.[12] This extension is most questionable when defendant has defrauded the plaintiff, or as a fiduciary has committed some act evidencing an intent to appropriate the money entrusted to him. In such circumstances it is incongruous to presume that the wrongdoer has become punctilious with respect to withdrawals, and in fact a few courts have refused to do so for this reason.[13]

A second problem relates to deposits: if the wrongdoer is presumed to be honest with respect to withdrawals, he should also be presumed to intend to restore the claimant's funds, if they have been depleted, by subsequent deposits of his own funds. This, however, is so implausible that the cases have split on the issue, with the majority and the *Restatement of Restitution* refusing to go that far, but rather requiring proof of an intent to make restoration.[14] The wrongdoer is thus presumed to protect the claimant's funds with respect to withdrawals, but not with respect to deposits, a patchwork of presumptions that is unlikely to reflect his actual intention in either respect.

A third problem is this: if the funds withdrawn are invested in traceable assets while the balance in the account is dissipated, *Knatchbull* would presume that the claimant's funds were dissipated. This seemingly unsatisfactory result led an English court in a subsequent case to reverse *Knatchbull* in that situation and to presume that traceable funds were withdrawn to make the investment.[15]

This reverse presumption arguably is not inconsistent with *Knatchbull*, but rather represents an extension of its logic: the honest wrongdoer is presumed to intend to use the claimant's funds to make the best investment, while dissipating his own funds. If this seems rather fanciful, the *Restatement of Restitution* embroiders the fancy into metaphysics by providing that if funds from a commingled account are withdrawn and invested in an asset, the claimant is entitled to a lien against the asset to the full amount of his claim, even if that exceeds the amount invested. Thus if a claimant has a $1,000 claim against a commingled account, and $500 is withdrawn and invested in securities, the *Restatement* would give the claimant a $1,000 claim against the securities.[16] The courts have not gone so far, but rather limit the preference to the amount invested in the case of an equitable lien, or to the proportionate part of the product in the case of a constructive trust.[17]

The artificiality of these presumptions makes it plain that the presumed intention of the wrongdoer—the original basis for the Rule of Jessel's Bag—is no longer a useful guide to what the rules are or ought to be. Nor is a desire to protect a recovery by the innocent defrauded claimant and to punish the tort-feasor or his estate a sufficient or even relevant basis on which to rest tracing doctrine. Since a constructive trust or equitable lien gives a defrauded creditor a significant advantage over general creditors only when the defendant is insolvent, the equities to be balanced in most cases are those of the defrauded claimant versus other general creditors; the equities of the tort-feasor are not involved.[18] A tracing rule that favors the defrauded claimant deprives the general creditors of their share of the traced assets and to that extent collides with the strong federal bankruptcy policy requiring equal treatment for all unsecured creditors. Thus it has been held that where trust funds are commingled with the debtor's property, a state statute that confers a lien on all the debtor's property without requiring tracing would create an invalid preference under the bankruptcy law.[19]

If tracing must stop somewhere in order to respect the equities of general creditors (and the bankruptcy law) where shall the line be drawn and on what theory? We suggest that the limits of tracing be determined by a reasonably coherent theory that is not convoluted in an effort to protect the defrauded claimant in the exceptional case in which he tilts with the tort-feasor.[20] In this approach the superior equities of the defrauded claimants are viewed as satisfied by their entitlement to a preference to the extent of tracing, but the limits of tracing are determined without reference to equitable considerations. In effect, this would mean replacing artificial presumptions designed to reward the claimant at the expense of the tort-feasor with a more plausible and neutral alternative.

An obvious neutral presumption is in fact a more realistic one: the wrongdoer did not care which money he used when he withdrew funds; in effect, he picked dollars at random without regard to their source. If that is assumed, probability theory applies to the results of the selection process.

Imagine that the claimant and the wrongdoer's funds are represented by labeled tokens in Jessel's Bag and the wrongdoer withdraws a given number of them at random without regard to their labels. The number of claimant's tokens removed (and hence the number remaining in the bag) will vary if this experiment is performed repeatedly (restoring the bag to its original state after each repetition), but the probabilities of various results can be calculated and the effect on the claimant's funds determined by probability theory.[21] To perform the calculations precisely, it would be necessary to decide in what units the withdrawals were made (i.e., the denomination of the tokens) and the choice is not obvious. However, the results that are important for our purposes are substantially independent of this choice.

First, the "expected" result (sometimes referred to as the "mean" result) is the one in which the proportion of the claimant's tokens among those withdrawn equals their proportion in the bag from which the withdrawal was made. A result as near as possible to the expected number (a result exactly equal to the expected number may not be possible because the expected number may include a fractional token) is among the most probable exact results, although this probability is still likely to be very small because the probability of *any* particular result is small.

Second, and of greater significance, the expected number is approximately equal to the median, i.e., there is approximately a 50 percent probability that the number of the claimant's tokens withdrawn would exceed the expected number, and approximately a 50 percent probability that it would be less than this. If a court were to presume that the proportion of claimant's tokens withdrawn equaled their proportion in the bag (or the next larger whole number of tokens if the expected number included a fractional token), there would be a probability of about 50 percent that the actual number of traceable tokens withdrawn was equal to or less than this and correspondingly a probability of about 50 percent that the number of claimant's tokens left in the bag was at least equal to the number obtained by subtracting the expected number withdrawn from the original number in the bag. A presumption to that effect would thus be supported by a preponderance of probability, which is the level frequently associated with a preponderance of evidence.[22]

This result is unchanged if there are multiple withdrawals, which may simply be aggregated and treated as a single withdrawal. If the trustee makes a deposit between withdrawals (represented by adding tokens to the bag), it will usually be of nontraceable funds. Whatever the characterization, if a subsequent withdrawal is at random from the bag as newly constituted, the same principle would apply: the expected number of claimant's tokens remaining in the bag after a series of deposits and withdrawals would be the number that would be obtained if at each stage the proportion of claimant's tokens among those withdrawn were equal to the proportion of claimant's tokens in the bag at the time the withdrawal was made.[23]

To illustrate this rule of proportionality, suppose the bag contains 100 dollar tokens, of which half are claimant's funds. If fifty dollars are withdrawn at random, the expected number of claimant's dollars remaining in the bag is 25. Assume the wrongdoer now deposits 25 additional dollars of his own funds, and then withdraws $25. The expected number of claimant's dollars then remaining would be two-thirds of 25, or 16.67, so that at the end of these transactions the expected balance of claimant's dollars in the account would be in excess of 16. If the court adopted the probabilistic approach suggested here, it would allow tracing to the extent of $16, as compared with $50 under the rule in *Knatchbull* or $0 under the rule in *Clayton's Case*.

Of course a proportionality rule would be less favorable to the claimant than the current rules. But it occupies a rationally defensible ground and one for which there is some case authority.[24] In addition, the *Restatement* proposes such a rule for the special case in which a constructive trust is sought against a conscious wrongdoer.[25] The *Restatement* theory appears to be that an equitable lien up to the amount of the claim attaches to all the funds in a commingled account (without regard to their identity), but that a constructive trust may only be asserted with respect to the claimant's property and its product. The *Restatement* thus recognizes the proportionality rule when tracing is required of the claimant's specific funds. Since most courts do require tracing of the claimant's funds and have not accepted the *Restatement*'s broader lien theory,[26] we may summarize the situation by observing that the proportionality rule marries two ideas: the judicial insistence on tracing as the basis for an equitable lien or constructive trust, and the *Restatement*'s rule of proportionality when it deems tracing to be required for the claimant's remedy.

The Lowest Intermediate Balance

Since a deposit does not usually restore traceable funds once depleted,[27] the sequence in which deposits and withdrawals are made may critically affect the amount of traceable funds remaining in an active account, even though the closing balance exceeds the amount of such funds. If, for example, sufficient withdrawals are made first, the traceable funds could be depleted, whereas they would have remained intact if sufficient deposits had been made first. This situation is reflected in the rule that traceable funds may not exceed the lowest intermediate balance of the account.

This rule creates a difficulty in active accounts because the sequence of deposits and withdrawals within a day generally is not known, since banks do not determine true intermediate balances but rather limit themselves to daily closing balances. For example, the Trustee of Franklin New York Corporation and Manufacturers Hanover Trust Company sought to trace and recover $30 million that allegedly had been obtained by fraud from Manufacturers and deposited in Franklin National Bank's account with the Federal Reserve Bank of New York.[28] Although this account was far in excess of $30 million at the close of each day from the day of deposit on April 3, 1974 to the day the bank closed on October 8, 1974, there were hundreds of deposits and withdrawals on each day, far exceeding in aggregate amount the opening balance in the account. However, the lowest intermediate balance could not be determined because the Federal Reserve Bank did not compute intraday balances.[29]

What is to be done when the order is unknown? One answer was given by the decision in *Republic Supply Co. of California v. Richfield Oil Co.*[30] In that case, the master who conducted the trial had three theories presented to him to deal with the intermediate balance problem. First, only daily closing balances should be considered. Second, deposits and withdrawals in the order posted by the bank should be accepted as the true sequence, even though it was shown that this sequence represented only "the arbitrary inclination of the posting clerk." Third, the lowest balance should be determined by deducting all withdrawals before crediting any deposits. The master chose the third theory, arguing that since plaintiff had the burden of proof, the nature of the sequence should be presumed against him. The district court adopted the master's position.

On appeal, the Ninth Circuit reversed. Since the choice was between the tort-feasor and the innocent claimant, the equities were not equal and this inequality justified putting the burden of proof on the defendant. Thus the court held that a prima facie case of tracing was proved by showing daily closing balances in excess of the amount of traceable funds; if in fact that account fell below the amount of the funds during the course of a day, the burden was on the defendants to prove that fact "with accuracy." If the court meant by this, proof of the actual sequence (and it would seem that the court did mean that), the presumption in effect gave this point irrebuttably to the claimant.

We have already criticized the court's balance of equities argument. The case illustrates the point previously made because the defendant was clearly

insolvent and hence its equities were not involved, but only those of the other creditors. In this situation, instead of presuming a sequence that favors one party or the other, it seems fairer to assume that each possible sequence of withdrawals and deposits was equally likely. On that assumption, the probability that the trust funds were dissipated to any given extent by a low intermediate balance can be calculated by a formula in certain simple cases and estimated by computer simulation generally.

For purposes of illustration, consider the simple case in which all deposits and withdrawals are of a fixed amount and are equal in number, so that the opening and closing balances are equal. Such a sequence may be modeled in familiar terms by assuming that a fair coin is tossed, say 200 times, with heads representing a $1 withdrawal and tails representing a $1 deposit, and that each face comes up 100 times. The probability of dissipation of $10 or more in such a case is equal to the number of different possible sequences of 100 heads and 100 tails in which, at some point, the number of heads leads the number of tails by at least ten, divided by the total number of possible sequences of 100 heads and 100 tails.

Most people do not have an intuitive feel for a probability of this type because it does not relate to any common experience (except perhaps gambling). Test yourself: in 200 tosses of a fair coin in which 100 heads and 100 tails appear, what is the probability that heads led tails by at least ten *at some point* in the process?[31] The presumption of nondissipation may seem reasonable if one believes that this sequence of tosses wound up in balance because by some self-correcting mechanism it stayed close to balance throughout the process. In fact, it is more correct to regard the process as having segments with possibly large leads in one direction that are counterbalanced by segments with leads in the opposite direction.[32]

When deposits and withdrawals are all of the same amount (although not necessarily equal in number) a simple formula (set forth in the margin) gives the probability of a dissipation equal to or greater than a given level.[33] To illustrate the formula, the table below sets forth the probabilities of various levels of dissipation in an account which is subjected to 100 one-dollar deposits and 100 one-dollar withdrawals, so that the closing balance is equal to the opening balance.

The table shows, for example, that the probability that $8 or more will have been dissipated is 0.522, while the probability that $9 or more will have been dissipated is 0.445. A court following the rule we have suggested would find that about $8.86 had been dissipated, because that is the expected amount.[34]

Computing probabilities in actual cases is more complicated than our simple formula would allow for two reasons. First, deposits and withdrawals are unlikely to be in identical amounts, so that probabilities determined by the formula will not represent the actual situation. It appears likely, however, that if

TABLE I

Probabilities of Dissipation Equal to or Greater Than Selected Amounts Assuming 100 $1 Deposits and 100 $1 Withdrawals

Amount Dissipated (k)	*Probability of Dissipating k or More*
$5	0.779
6	0.698
7	0.613
8	0.522
9	0.445
10	0.368
—	—
15	0.105
16	0.077
17	0.056
18	0.041
19	0.028
20	0.018

the formula is applied to an average figure for deposits and withdrawals, the probability of exceeding any given dissipation determined by the formula will be smaller than in the actual situation.[35]

Second, the formula gives the probability of dissipation if the rule in *Knatchbull* is followed and withdrawals are deemed to be made first from non-traceable funds; it does not apply if random selection is assumed. If there are $50 of claimant's funds in an account that has $100, and there are fifty $1 deposits and fifty $1 withdrawals, under the Rule of Jessel's Bag the expected dissipation would be $6.27, all of which would come from nontraceable funds since they are presumed to have been used first. But if random selection is assumed, our computer simulations indicate that the expected dissipation of claimant's funds would be about $30, so that only $20 of such funds would remain on the average.[36] Thus, as one would expect, the expected levels of dissipation of traceable funds are much greater than those indicated by the formula if the Rule of Jessel's Bag is rejected in favor of the presumption of random selection. Computer simulation is thus required if deposits and withdrawals are of differing amounts or if the presumption of random selection with respect to withdrawals is adopted.

Since the formula appears to understate the probability of dissipation, it may be used by general creditors to show that plaintiff's tracing claim must fail. This is likely to occur frequently because most bank account cases involve tracing over many days. Since the possible exhaustion of claimant's funds on each day is a separate event, the probability of nonexhaustion at the end of a number of days equals the product of the probabilities of nonexhaustion on each day. If there are appreciable probabilities of exhaustion on each day, the probability of exhaustion over a longer period will exceed 50 percent within a very few days. In that situation a more accurate calculation to show a higher probability of exhaustion would be unnecessary.

For example, in the table previously shown, there is a probability of 0.056 that $17 or more would have been dissipated. If the initial composition of the bag implied that such a dissipation would have exhausted the claimant's funds, the probability of nonexhaustion is 0.944 for a single day. But if the same situation were repeated for thirteen days, the probability that any traceable funds would remain at the end of that period would be less than 0.50.[37]

In the Franklin Bank case previously mentioned, there were many days in which the opening balance in the Franklin Bank's Federal Reserve account was in the range of $80 to $100 million; there were 400 to 500 deposits and withdrawals aggregating $2,500 to $3,500 million, and the closing balance was in the same range as the opening balance.[38] On such days, the probability that the account was completely exhausted at some point was about 30 percent on the basis of the formula, and the probability of nonexhaustion was about 70 percent. If only a single day were involved, we could not conclude that claimant's funds had been dissipated. But since there were more than twenty such days, the probability that any such funds remained becomes extremely small (less than one chance in a thousand).[39] It is thus unnecessary to inquire how much smaller the probability of nonexhaustion would be if more accurate methods were used.

Should We Bother?

One is entitled to ask whether refinement of these presumptions is worth the burden of introducing and understanding the technical methods of probability theory. We think that although the methods are technical, the burden is not very great and their use is justified. When the sequence and amounts of deposits and withdrawals are known, computing the expected undissipated amount using the pro rata approach is only slightly more difficult than computing the amount under the Rule of Jessel's Bag; a bookkeeper could do either calculation. Probability theory tells us only that it is reasonable to use a pro rata division of the withdrawals between trust and nontrust funds because that division generates a result that is consistent with a preponderance of probability approach. The outcome strikes a balance between general creditors and the defrauded claimant, which seems more in keeping with their balance of equities than an arbitrary resolution in favor of the claimant.

When the sequence of deposits and withdrawals is unknown, as in an active bank account, the application of a safe-bound formula will frequently indicate that it was more likely than not that the claimant's funds did not survive long enough to be traced, even under the favorable assumptions indicated by the formula. Of course, unless the parties agreed that the formula was correct, an expert would have to testify. In closer cases requiring more precise calculations, a computer simulation would be needed. However, the computer program is quite simple, and the computations could not reasonably be subject to much dispute. Indeed, the court might appoint a neutral expert to make them and report.

There is no doubt that the introduction of probability theory adds some complexity, but the burden does not seem unfair. The general creditors (or their representative, the trustee) should not object, since they would otherwise lose the point to a presumption. The court would not be greatly burdened, since the results of the computer calculation should not be disputed and may indeed be stipulated. The defrauded plaintiff may complain at losing the advantage of the presumption and having to work harder to sustain a tracing claim, but given the simplicity of the computer program, the extra work is not very great. Moreover, since the complexity arises because the claimant seeks to stretch the tracing doctrine to situations in which there was a substantial risk of dissipation, some extra effort to prove that no dissipation probably occurred does not seem inappropriate. Finally, as a matter of legal style, it seems preferable to deal with factual uncertainty by making neutral assumptions and calculating probabilities than by favoring one party with artificial presumptions when there is no strong policy basis for thus tilting the balance in civil litigation.

*Lecturer in Law, Columbia University. A.B., 1955, J.D., 1958, Harvard University, Mr. Finkelstein writes: "This article is dedicated to B. W. Nimkin, Esq., who for many years has graced the practice of law in New York City with both integrity and scholarship."

†Higgins Professor of Mathematical Statistics, Columbia University. A.B., 1935, Ph.D. 1938, Harvard University.

[1] *See, e.g.*, RESTATEMENT OF RESTITUTION, §§ 160, 161, and 166. Which remedy is more advantageous for the claimant will depend on the success with which the funds have been invested.

[2] Most courts use the same tracing rules for equitable liens and constructive trusts. However, a few courts and the *Restatement of Restitution* have proposed a special rule for constructive trusts. This is discussed *infra* at 71.

[3] 48 Stat. 162.

[4] *See* Anno., *Following Trust Funds Deposited in Mixed Account of Trustee*, 102 A.L.R. 372.

[5] A case not in the usual mold involved Franklin New York Corporation, the holding company of the failed Franklin National Bank. The trustee in bankruptcy, Sol Neil Corbin, used a tracing theory to assert a constructive trust over the proceeds of a $30 million loan by Manufacturers Hanover Trust Company to the Franklin New York Corporation that was invested in Franklin National Bank shortly before its collapse. (Corbin v. Federal Deposit Insurance Corporation, 75 Civ. 2144 (filed E.D.N.Y. December 19, 1975)).

[6] *E.g.*, C. O. Funk & Sons, Inc. v. Sullivan Equipment, Inc., 89 Ill.2d 27 (1982) (tracing rules used to determine proceeds from the sale of collateral which were subject to the secured creditors lien under the Uniform Commercial Code); Utica Sheet Metal Corp. v. J. E. Schecter Corp., 53 Misc. 2d 284, 278 N.Y.S.2d 345 (S.T. Schenectady County 1967) (trust under New York lien law).

[7] *See* RESTATEMENT OF RESTITUTION §§ 202–215. For a review of the cases see Anno., *Following Trust Funds Deposited in Mixed Account of Trustee*, 102 A.L.R. 372. *See also* DOBBS, HANDBOOK ON THE LAW OF REMEDIES, 421–30 (1973); 1 PALMER, THE LAW OF RESTITUTION, §§ 2.16–2.18 at 193–217 (1978).

[8] 13 Ch. Div. 696 (1880).

[9] 1 Mer. 572.

[10] 13 Ch. Div. at 727–28.

[11] 13 Ch. Div. at 735.

[12] *See, e.g.*, I. & T. N. Bank v. Peters, *et al.*, 123 N.Y. 272, 25 N.E. 319 (1980).

[13] In *Mitchell v. Dunn*, 211 Cal. 129 (1930), the court, citing Crawford County Commissioners v. Strawn, 157 F. 49 (6th Cir. 1907), refused to adopt the presumption of honesty with respect to withdrawals because the trustee had demonstrated prior dishonesty.

[14] *See, e.g.*, Conqueror Trust Co. v. Fidelity & Deposit Co., 63 F.2d 833 (8th Cir. 1933); RESTATEMENT OF RESTITUTION, § 212 (b); Comment on clause (b).

[15] *In re* Oatway, Hertslet v. Oatway, 2 Ch. Div. 356 (1903).

[16] RESTATEMENT OF RESTITUTION, § 211, illus. 2.

[17] 1 PALMER, THE LAW OF RESTITUTION, § 217 at 209 (1978).

[18] A claimant may gain an advantage even when the wrongdoer is not insolvent by asserting a constructive trust and thus recovering the wrongdoer's profits, but this will generally not be a major factor with respect to funds in bank accounts.

[19] Elliott v. Bumb, 356 F.2d 749, 754–55 (9th Cir. 1966).

[20] Those cases may be left to special rules.

[21] The nmber of claimant's tokens included in the withdrawal is a random variable that has what is known as a hypergeometric distribution. For a discussion and derivation of formulas see, e.g., FREUND, MODERN ELEMENTARY STATISTICS 163 (5th ed. 1979).

[22] *But cf.*, FINKELSTEIN, QUANTITATIVE METHODS IN LAW, ch. 3 (1978). We note that if the numbers of claimant's and other tokens are sufficiently large, the probability that only the other tokens would be selected at random—the result presumed by Jessel as a result of intention—is vanishingly small.

[23] The expected number of claimant's tokens remaining in the bag after a series of deposits and withdrawals may be computed as the probability that a designated claimant token would not be removed in any of the withdrawals, times the number of claimant's tokens initially in the bag. The probability of nonremoval for each withdrawal is the ratio of the number of tokens not removed to the total number in the bag at that time; the probability of nonremoval after several withdrawals is equal to the product of these probabilities for each withdrawal. The value thus computed for the expected number of claimant's tokens remaining in the bag may be shown to be algebraically equivalent to that determined by assuming that the expected number was removed at each withdrawal.

[24] Orr v. Rose, 169 Okla. 387, 37 P.2d 300 (1934); Standish v. Babrock, 52 N.J. Eq. 628, 29A, 327 (1894). *In re Anjopa Paper & Board Mfg. Co.*, 269 F. Supp. 241, (S.D.N.Y. 1967) the court observed that "it might well be argued that the beneficiary claimant ought only to be entitled to a lien on the property in proportion to the amount of trust money in the account at the time of its purchase." 269 F. Supp. at 261 n. 29. But the court declined to do so on the ground of symmetry: if the *Knatchbull* presumption allowed full recovery of the balance of a fund not withdrawn, it should also allow full recovery of assets purchased by withdrawals from the fund. *Id.* The proportionality rule we suggest would be symmetrical by limiting recovery in both situations.

[25] RESTATEMENT OF RESTITUTION, § 211(2) and Comment on subsection (2). The rule is based on Professor Ames's article, *Following Misappropriated Property Into Its Product*, 19 HARV. L. REV. 511, 518–19 (1906) and Learned Hand's opinion in Primeau v. Granfield, 184 F. 480 (S.D.N.Y. 1911), *rev'd on other grounds*, 193 F. 911 (2d Cir. 1912), *cert. denied*, 225 U.S. 708 (1912). The rule is also consistent with the treatment of the case in which multiple parties have claims against a single fund. In that situation, each claimant's recovery is proportionately reduced by the amount of any withdrawals from the fund. SEE RESTATEMENT OF RESTITUTION § 213.

[26] *See, e.g.*, Republic Supply Co. v. Richfield Oil Co., 79 F.2d 375 (9th Cir. 1935); Brennan v. Tillinghast, 201 F. 609 (6th Cir. 1913); *In re* Anjopa Paper & Board Mfg. Co., 269 F. Supp. 241 (S.D.N.Y. 1967). In some cases the *Restatement* theory has been accepted. *See* Erie Trust Company's case (No. 1), 326 Pa. 198, 191 A. 613 (1937).

[27] *See, e.g.*, Maryland Casualty Co. v. City Nat. Bank, 29 F.2d 662 (6th Cir. 1928), *cert. denied*, 279 U.S. 847 (1929).

[28] *See* n.5, *supra*.

[29] Whether tracing should be allowed in this circumstance was not judicially resolved because the case was settled after the beginning of a bench trial.

[30] 79 F.2d 375 (9th Cir. 1935).

[31] *See* Table I *infra* for the answer.

[32] Many people think that the familiar newspaper chart of stock prices, which shows "waves" of price movement around the local average, displays trends from which future price levels may be projected. In fact the waves of the chart are similar to those that would be produced by coin tossings.

[33] Assume that n is the total number of deposits and withdrawals (not necessarily equal in number); w is the size of each single deposit or withdrawal; and aw is the amount by which the closing balance is less than the opening balance (a is a negative integer if the closing balance exceeds the opening balance). With this notation, the probability that at some point the opening balance was decreased by kw or more is:

$$P(\text{opening balance was decreased at some point by} \geq kw) = \frac{b(b-1)\ldots(b-k+1)}{(n-b+k)(n-b+k-1)\ldots(n-b+1)}$$

where $b = (n+a)/2$. This is approximately equal to e raised to the $-|2k(k-a)/n|$ power, where e is the base of natural logarithms 2.718.

[34] In the special case in which the opening and closing amounts are equal, the expected dissipation is equal to $0.627\sqrt{n}$ and the median dissipation is $0.589\sqrt{n}$.

[35] There are various ways of selecting such an average figure. For example, one may assume that the numbers of deposits and withdrawals are the actual numbers and then determine the figure that preserves the total aggregate deposits and withdrawals, or the figure that preserves the difference between them. Under either of these regimes it appears from computer simulation that the probability of dissipation will be larger in the actual case than that computed using the formula. However, a mathematical proof of this fact is not yet available.

[36] The authors are indebted to their colleague Prof. Bruce Levin for preparing the computer program and performing the simulation.

[37] $(0.944)^{13} = 0.47$.

[38] Stipulation of Facts, Joint Exh. A-1.

[39] $(.70)^{20} = .0008$.

Statistics in Employment Discrimination: The Use of Experts

James Beckett, *Criterion Incorporated*
Peter Lewin, *University of Texas at Dallas*

Introduction

Many people today have access to complex statistical techniques that they do not fully understand. This presents the statistics profession with a problem. An analogy with the modern science of microwave cooking may prove helpful. Just as in cooking, (just about) anyone can take a recipe and follow the step-by-step instructions and sometimes (hopefully) come out with a decent meal--but it is usually only a trained gourmet chef who can adapt or manipulate the recipe to meet changing or specific conditions. He (or she) has the understanding of the process of cooking. Not everyone wants to be a great chef (or a great statistician), but everyone should know where to find one for that special occasion (litigation). Gourmet food (complicated analysis) is not required for every occasion, but food that tastes good (and analysis that is appropriate) is always the objective. Wrong, ill-conceived statistical analysis (just like poisonous food) should not be tolerated by the profession since it is dangerous to one's health. To complete the analogy, the following correspondences are noted: (a) data = ingredients (bad ingredients imply garbage in, garbage out); (b) recipe = software, algorithm, or methodology (sometimes critical ingredients are not available or are inadvertently left out); and (c) oven or microwave = mode of analysis or computer. The computer (especially powerful micro-computers), just like the microwave oven, has become the great equalizer in permitting anyone to become an "instant statistician", regardless of training.

The Expert Witness in EEO Contexts

The difference between the statistician and the microwave button pusher is that the statistician knows the fundamental assumptions inherent in the methodology. This knowledge should and usually does prohibit the statistician from employing inappropriate techniques regardless of how good the result looks to the people who are paying him.

In using an expert, it is important to tailor the expert's role to his/her expertise in an optimal way. The lawyer and the expert bring different types of knowledge to any particular case. The lawyer brings a knowledge of law and of specific facts of the case. The expert brings a knowledge of his/her particular field of expertise and, perhaps, some experience of similar cases. If the expert is to be effective, he/she must not be made to go outside of his/her area of expertise.

Many aspects of EEO litigation involve matters of economics and the application of statistics to economic issues. Modern-day training in economics necessarily involves a fairly high level of statistical training. An economist with extra training in econometrics will generally be of greater value. This is because the economist can then testify on matters of statistical and economic relevance.

In general, the economist is likely to be strongest in areas involving multivariate analysis and weakest in those involving small-sample, non-parametric statistics. In many, perhaps most, class-action Title VII cases, however, some form of multivariate analysis is called for. In fact, in recent years, this area of law has provided an unprecedented and increasing number of statistical applications of the economic theory of <u>human capital</u>.

A person's human capital is the value of his/her productive skills, talents, and knowledge. This human capital is valuable to a firm in that it reflects a capacity to produce goods and services that can be sold for profit. A firm requires human capital stocks of certain types in order to achieve any production plan. Human capital differs from physical capital in that it is inseparable from its original owner; it is "embodied" in the individual worker. The relevance of this to the issue of employment discrimination is, thus, obvious. The question is: is the firm rewarding only the human capital characteristics of its employees or are other non-productivity related characteristics, such as race and sex, playing a role?

Typically then, the human capital context of a Title VII case will call for some knowledge of labor economics and multivariate statistics.

In addition to the statistician and economist, a variety of other possible experts may be of use in employment discrimination cases. Notably, experts in the area of industrial psychology may be used to evaluate the "validity" of employment related tests. Industrial psychologists and industrial relations experts are sometimes used to evaluate training programs and job-content profiles as they relate to earnings.

Expert witnesses are used in a wide variety of contexts besides EEO analyses. The role of the medical expert in establishing the cause and extent of injury is well established. Psychologists and psychiatrists can testify about states of mind. Rehabilitation experts can testify about job opportunities.

The economist and statistician also have opportunities to testify in areas other than Title VII. The statistician may be of use in a variety of criminal contexts (for example, fraud) in evaluating the likelihood of certain suspect outcomes. The economist is used routinely in personal injury and wrongful death cases to evaluate lost earnings and economic damage. The economist may also be used to evaluate lost profits or economic damage in any other context, notably antitrust litigation.

Every expert has his or her own style: loquacious or laconic, flamboyant or conservative, nervous or cool, authoritative or informative. However, the most prominent difference between witnesses or experts is usually based on the answer to the questions, "Which party has hired you to be a neutral expert in this matter?" or "Which side are you impartial toward--plaintiff or defendant?" The plaintiff expert is usually working with several distinct disadvantages: (1) he is usually less well funded, (2) he rarely has the defendant's "full" access to data and information relevant to the case, and (3) he does not get the last word. Of course, the plaintiff also has some advantages over the defendant, for example, getting to take the first shot in the duel.

Working with Attorneys and Clients--Anatomy of a Typical Case

The expert should not work in a vacuum; in order to design and carry out a decent analysis of the issues involved, typically meetings or initial consulting sessions are held in which the basics of the case are presented by those in attendance. These early meetings are attended by three (sometimes only two for plaintiffs) types of people. The triumvirate of involvement in a case is slightly different for the two opposing parties: the plaintiff has expert/attorney/named plaintiff, whereas the defendant has expert/attorney/personnel expert. Notwithstanding the above advantages to the defendant's statistician, sometimes the most valuable resource of the defendant's statistical expert is the company's representative (legal liasion, personnel expert, data expert, computer expert) who can give clear, personal, and detailed answers to virtually any question about the employment practices and data in question.

Most cases are received by referrals and often by word of mouth in the network of attorneys, bar associations, etc. Our experience has been that typically the labor attorneys that do (practice) this type of work do a lot of this type of work. Of course, news of bad cases travels faster and farther than news of good cases. A bad case (for an expert) is not necessarily a lost case, but is one in which the expert did not come across well or was trapped in an inconsistency.

It is, therefore, important that the expert and the attorney have a good understanding. This requires that there be careful and extensive preparation and full and frank exchange between them. The attorney should be aware of what the expert (and the opposing expert) can and cannot say. It should never be forgotten that, while a qualified expert may have the right to be <u>heard</u> in a court of law, he/she does not have the right to be <u>believed</u>. Minimizing the extent of "surprises" to attorney and expert will serve to enhance the expert's credibility.

Careful preparation starts with the discovery process. The role of the expert in discovery is limited but vitally important. This can best be explained by noting that any case is distinguished from every other case by four main factors:

1. The issues: There are many possible issues in Title VII litigation: hiring, placement, promotion, transfer, training, pay, conditions, and termination. Only rarely will a single case involve them all, and frequently only one or two will be involved.

2. The data: The data relevant and available will differ from case to case.

3. The process: Each case will involve standards and practices that are more or less unique to the firm or organization in question. These may be thought of as characterizing aspects of the employment process.

4. The scope: Each case will have a particular scope of application, a relevant time period and a relevant class or subclass of employees.

Having obtained information on these four factors, the expert will then be in a position to help determine the extent and type of discovery necessary for the conduct of the case. The type of relevant decision will include what types of study would be appropriate, what type of data is necessary to do the studies, the availability of data in the required form, what would be entailed to convert it if it were not in the required form, and so on.

An example of the way in which woefully inadequate discovery can prejudice the whole case is presented by the following situation. A large national insurance company was sued for hiring discrimination. Incumbent statistics were relevant, but an expert was engaged only at a fairly late stage and the statistics were not obtained.

The plaintiff's task was merely to establish an adverse statistical impact in the hiring of sales persons by the insurance company in 1975. The insurance industry had been accused by some of having been slow to incorporate females into non-clerical positions. Nonetheless, plaintiffs in this case failed to establish a <u>prima facie</u> case of adverse treatment of females.

The statistical issue presented was the simplest one--a straightforward comparison of snapshot-utilization to availability. The first question--utilization--should have been

determined by the lawyer during discovery. The lawyer's failure to pursue discovery resulted in the expert's being provided a "universe" as follows (Table 1):

	Male	Female	Total
1975			
Job A	6	2	8
Job B	6	0	6
1976			
Job A	2	0	2
Job B	11	1	12

Table 1

The second issue--availability--should have been addressed jointly by the lawyer (who should have known what the job requirements were) and the expert (who should have been able to locate and defend the best availability statistics). In fact, the lawyer failed to discover the true qualifications for the job, and the expert attempted to defend general population, rather than labor force or insurance industry, as the proper measure of availability.

The defendant's expert was able to accept direct instructions concerning the data to be analyzed (those were the numbers, and the jobs had different qualifications), locate availability statistics that best reflected the job-content of the jobs, and demonstrate that zero utilization of females was, in fact, not adverse, given the data discovered.

Adequate and appropriate discovery lays the foundation for the analysis. A statistical model must be chosen. The choice of a statistical model is a crucial aspect of any analysis. In a Title VII context, it will most likely involve both economic and statistical theory. Regarding the former, the general problem of misspecification is relevant. This was recently stated quite penetratingly for a regression context as follows:

> . . . in economics . . . the theory used to derive tests ordinarily does not generate a complete specification of which variables are to be held constant when statistical tests are performed on the relation between the dependent variable and the independent variables of primary interest. Accordingly, in such cases there will be a set of often very different candidate regression-based tests, each of which has equal status with the others since each is based on a different projection of the same underlying model. Except in the unlikely event that the explanatory variables are mutually orthogonal, the conditional regression coefficients, which generally form the basis for the test statistic, will depend on the conditioning set. We conclude from this that, if a theory which does not generate a complete specification of the regression test is nonetheless to have testable implications, these must be robust over the permissible alternative specifications. If the restrictions indicated by the theory are satisfied in some projections, but not in others that have an equal claim to represent implications of the theory, one cannot conclude that the theory has been confirmed.
>
> The fact that the observable implications of valid theories must obtain over a broad (but usually incompletely specified) set of regressions rather than for a single regression introduces a large and unavoidable element of imprecision into hypothesis testing . . . .

Although the above quotation was taken from an article dealing with the demand for money in a macroeconomic context, it applies with equal force to the case of earnings discrimination. Human capital theory suggests a set of earnings functions that is potentially very large. A subset is the set of linear or log-linear equations which have become traditional. It can be shown that this type of analysis can result in a finding of discrimination when, in fact, no such discrimination exists. The problem arises because, although the linear regression model fits the data adequately, it is not able to adequately represent the details of the salary administration process. In particular, where the work force is composed of women and minorities in decreasing proportions in the higher educational and skill levels where women and minorities with high educational and skill levels are typically relatively recent entrants into the work force and where education and skill information is limited, the chances of a spurious finding of discrimination increase.

It is important, therefore, for the expert to attempt to learn as much as possible of the relevant aspects of the firm's employment process and to attempt to "model the process" as accurately as possible. Failure to do so may result in a successful <u>modeling of the data</u> (good fit, high R^2), but a failure to <u>model the process</u>.

In attempting to get a more accurate picture of the employment process as a whole, the expert may want to use a series of different statistical models, one appropriate to each relevant issue. This is often a substitute for or a complement to a conglomerate regression-type analysis which tries to represent the entire employment process. For example, initial placement and salary progression could be analyzed separately. The analysis of salary progression, which involves promotion, merit and general increases, transfers, and possibly training, can itself involve separate analyses, each addressing a specific issue. This in itself offers an advantage over the traditional regression analysis, which does not identify those employment practices that contribute to, or possibly are the sole sources of, discriminatory employment activity. And, in this respect, the expert may be led to the use

of various non-parametric, more basic types of models and tests.

If the plaintiff prevails on any of the issues, damages are likely to be awarded. The expert as economist-statistician has an important role to play in this second phase of the case. A model must be chosen that will yield a specific dollar amount.

It is important, in this regard, to remember that the point of reference is that amount which will make the plaintiffs whole, i.e., restore to them the value of the opportunities or earnings that they would have obtained <u>but for</u> certain discriminatory actions. Thus, for example, the entire productivity-adjusted earnings gap against women should not be figured as a loss, since, in the absence of discrimination, men would have gotten <u>less</u> and women more. The correct estimate of loss is the difference between a sex-neutral wage and the wage actually received. This type of consideration pervades all aspects of phase-two analysis.

Many cases go on for a long time. All cases essentially are either settled (sometimes everybody wins) or decided (sometimes everybody loses). Many decisions and most settlements are complex, where a compromise of sorts is reached by the parties themselves or through the court. Of course, appeals are not uncommon and can certainly greatly prolong the eventual "final" decision.

Settlement is one of the ways and, indeed, a common way that a case may end. The statistician or economist can certainly be of assistance to attorneys in the settlement "discussions" that frequently take place. These discussions sometimes even result in an actual settlement. Many settlements occur on the steps of trial for a variety of reasons that neither side will ever admit to later--as both sides wish to call the settlement a "victory" . . . that was mercifully forced upon the other side.

The decision on a case (the trial judge's opinion) can take up to a year and a half or more after the trial, depending on the length of the trial. It is interesting to read opinions where judges summarize one's testimony and comment on one's credibility. One must remember that the judge is faced usually with two eminently qualified opposing experts each contradicting the other and that the judge has the task of deciding, in effect, which expert is more credible; if both are equally credible, implicitly the tie goes to the expert on the prevailing side.

Real Case Examples Using Anonymous Data

Case Example 1: Misleading data or misleading interpretations of data can arise in a regression context. With many independent variables being measured on a large work force, the ideal of having complete data on every single employee is admirable. However, sometimes there are variables that are important predictors that were simply not maintained over the whole time period or were not maintained for some types of employees. Important missing data can be deleted or salvaged. Salvaging the missing data can involve estimation or control, either of which is usually preferable to the wholesale throwing away of large numbers of cases (employees). Table 2 shows why one must be suspicious of the reasons for which data (certain variables) are missing for some employees. In a preliminary regression on starting salaries for a firm, high school graduates (the default value of a set of education dummy-variables) appeared to be worth more than Bachelors or Masters due (in part) to the many higher level employees who had attached a resume (but had not filled out the education portion of the application) that detailed their educational attainment. The "Second Glance" at the table shows the effect of separating out the Missing from the default group. Resumes were not originally asked for in this case. Thus, the final step shown in Table 2 was based on the addition of the resumes which were later requested and received.

Case Example 2: Underspecified models arise in many situations, some of which cannot be helped (data not available). Table 3 shows a situation where underspecification may be argued from the residual analysis performed using the opposing expert's regression on pay as a function of a few human capital variables. The lower level individuals all seemed to be underpaid, whereas the higher level individuals almost uniformly appeared to be overpaid, in other words, not much different from what one would come up with using a regression with only a constant term.

Case Example 3: Since the observations correspond to people with names outliers can be traced, and sometimes can be legitimately thrown out of the analysis. They can be very influential with small samples and surprisingly so with large samples. Table 4 shows some of the "anomalies" and outliers that have been encountered in actual cases.

Case Example 4: Violated assumptions can involve or imply statistical assumptions or process assumptions; both types can cause incorrect conclusions. The most commonly violated statistical assumption is the assumption of independence. Tests performed year after year on the same employees are not strictly independent, but what are they? Are they positively serially correlated or are they negatively serially correlated or neither? Violations of process assumptions can cause obvious problems, as in the situation shown in Table 5, which shows the problems that are caused when terms are not defined carefully.

Case Example 5: The relative adequacy or inadequacy of a model often depends on where the burden of proof lies. Consider the results for an analysis of promotion presented in Table 6. The following cross-examination of the expert proved effective.

Q. ". . . the professional degree variable has a negative and a significant impact. Is that plausible?. . ."

A. ". . .The answer is no."

Q. "Then doesn't that result . . . cause you to question the model?"

A. "Not the result in and of itself."

Q. "If that sort of result repeated itself . . . would you have questions raised?"

A. "Again, let me stress that I'm not particularly concerned with the explanatory power of the model. The purpose of the model was not for predictive purposes. . . ."

Q. "Did you then limit your investigation here essentially to whether or not the [race/sex] variable appeared significant?"

A. "That's the primary focus of the analysis."

Q. "Isn't it possible that that particular variable . . . could be spuriously insignificant if the model itself is suspect?"

A. "Statistically, there is a chance that could occur."

Q. "And isn't a way to determine whether the model itself is suspect, to examine the results obtained on all other variables as to stability, plausibility and explanatory power?"

A. "That is one way of evaluating the model."

- - -

Q. "Is it true in that case [an equation with very low explanatory power], for example, that you could draw no conclusions at all from any of the independent variables?"

A. "Oh, I draw the conclusion that none of them are significant."

Q. "Is that a fair conclusion, or can you just draw the conclusion that none of them explain what occurred? In other words, can you draw the conclusion specifically, for example, that the [race/ sex] variable was not significant, as opposed to the conclusion that your . . . equation here simply failed to explain that occurred?"

A. "Well, the fact that none of the variables, including the [race/ sex] variable, was not significant, it is of some import."

Q. "Is it of import only granting the assumption that your model in general and this equation in particular have some meaningful explanatory power; must you assume that before you can reach the conclusion of no significance on the [race/sex] variable?"

A. "As I explained before, all of the other independent variables can be insignificant and it's still possible to obtain a significant estimate on the [race/sex] variable."

Q. "Does that follow with regard to each and every of the independent variables, can each of the independent variables be insignificant and you draw some conclusion from that?"

A. "I would conclude for the model as estimated . . . that they have no significant effect in terms of how the model specified on the likelihood of being promoted."

Q. "Would another conclusion that you could draw be that the model failed to isolate whatever are the significant effects on the likelihood of being promoted?"

A. "One could conclude that as well."

Conclusion

There are no simple rules when it comes to the analysis and presentation of statistics by an expert. The above examples illustrate some of the oft-encountered problems and potential solutions. However, every situation has some new ingredient that calls for careful examination and judgment in the application of statistical techniques.

First Glance		Black	White
	HS (default)	40	60
	Some College	30	10
	College Degree	20	20
	Masters	10	20
Second Glance		**Black**	**White**
	Missing	10	40
	HS	30	20
	Some College	30	10
	College Degree	20	20
	Masters	10	10
True Picture		**Black**	**White**
	Missing	5	5
	HS	30	20
	Some College	30	15
	College Degree	25	40
	Masters	10	20

Table 2

	1/1/75	1/1/76	1/1/77	1/1/78
GS-7	-5997.	-5896.	-6889.	-11436.
GS-9	-4274.	-5098.	-5598.	-6924.
GS-11	-2591.	-3314.	-3915.	-4642.
GS-12	-476.	-1170.	-1716.	-2182.
GS-13	1095.	852.	697.	611.
GS-14	3160.	2748.	2845.	3125.
GS-15	6177.	6058.	6758.	7580.

Table 3. Table of Average Residuals

Observed Salary	Expected Salary	Residual	Explanation or Comment?
19,000	20,000	-1,000	
20,000	20,000	0	
20,000	19,000	1,000	
20,000	15,000	5,000	Night Collector
26,000	20,000	6,000	Red Circled
18,000	23,000	-5,000	Refused Promotions
50,000	25,000	25,000	Related to Owner
30,000	20,000	10,000	Outstanding Performance
15,000	20,000	-5,000	Poor Attendance
15,000	20,000	-5,000	Interrupted Tenure

Table 4. Extract of Residual Analysis with Potential Outliers

For a large employer, level 5 has 30% blacks and is the level immediately below level 6. There have been 50 openings at level 6 during the past year which were filled from within. Only 10 of these selections went to blacks--a shortfall of 5.

The "catch" in this particular case (and there isn't always a catch) was that the expert counting these "job fills" or "opportunities for advancement" assumed that all the openings were filled by promotions from level 5 when, in fact, many were demotions during this year of retrenchment or belt-tightening. In fact, the data give an entirely different look when that fact is taken into account.

	Black	White	Total
Promotions	9 (30%)	21	30
Demotions	1 (5%)	19	20
Total Job Fills	10	40	50

Table 5

(PLEASE NOTE THAT TABLE 6 IS AT THE END OF THIS ARTICLE.)

*Cooley, Thomas F. and Stephen F. LeRoy, "Identification and Estimation of Money Demand," The American Economic Review, December 1981, p. 825.

VARIABLE	Grade 5	Grade 7	Grade 9	Grade 11	Grade 12	Grade 13	Grade 14	Grade 15
INTERCEPT	-3.5301*	-0.5547	0.2991	-1.7748*	-2.0330*	-1.3217	0.3710	-4.4784
RACE/SEX	-0.5828	-0.0476	0.4227	0.1737	0.3621	-0.0543	-0.3451	-0.0949
PTIGLTYR	5.0598*	1.3817*	1.0294*	2.7354*	2.6069*	0.0180	0.2469	1.6684
PTIGGTYR	(a)	-0.0363	-0.9880*	0.5322	0.5730	1.6221*	0.3091	2.6490
PTIGGT2	(a)	0.4407	0.1587	-0.1494	-0.0371	-0.4045*	-0.0150	-0.2636
PRMD	(a)	0.6678	0.3417	-0.0839	0.8140*	0.1130	-0.3419	0.6045
VETD	-0.0380	0.1289	0.1328	0.3380	0.1144	0.1207	0.3864	-0.6313
LOCD	4.6709*	0.0939	0.0564	0.4837*	0.2346	0.3616	-0.2412	-0.0477
LAWD	(a)	(b)	(a)	2.1625*	2.2429*	-0.2414	0.3044	1.9154*
RELTD	5.4454*	1.7561*	1.0283	-0.2662	0.2119	-0.9623*	-0.1695	0.7358
LBD	-0.0841	-0.3723	-0.1401	-0.2915	-0.0639	-0.1276	-0.0735	(a)
PROFD	(a)	0.1859	0.0730	-0.0874	0.0836	0.5717	-0.4355	-2.4055*
ADVDEGD	-0.6585	-0.0311	0.0706	0.3680	-0.2071	1.1527*	-0.2206	-0.5850
YEAR2D	-2.4264*	-1.6929*	0.2940	0.4574	1.2237*	1.6762*	2.3811*	(a)
YEAR3D	-0.6240	-1.4481*	0.1914	-0.9928	0.0777	-0.1166	1.0694	(a)
YEAR4D	-0.2413	-1.7567*	0.4623	-1.7413*	0.2842	-0.1299	1.2681	(a)
YEAR5D	-4.2735*	-2.2113	1.4186*	0.1983	1.2625	1.6965*	2.1444*	(a)
YEAR6D	-2.8920*	-3.1324*	-1.6385*	-1.4568*	-2.5113*	-0.0461	0.4295	(a)
EXREL	-0.0570	1.3224*	-1.0811*	-1.1721*	-2.4685*	0.3034	-1.1769	-1.5602
EXREL1	(a)	-0.5324*	0.2399*	0.1998*	0.4153*	-0.0920	0.2137	(a)
EXGOV	-0.0504	0.0168	0.0152	0.0116	-0.0581	0.0027	-0.1163	-0.0494
EXGOV2	(a)	-0.0026	-0.0007	-0.0015	0.0012	0.0018	0.0047	(a)
AGE	0.0017	0.0022	-0.0275*	-0.0118	-0.0213	-0.0676*	-0.0607*	0.0368
Model Chi-Square (Degrees of Freedom)	117.68* (15)	209.31* (22)	161.20* (21)	281.95* (22)	284.20* (22)	155.33* (22)	39.01* (22)	17.55 (14)

*Significant at a 95 percent confidence level.

Notes: (a) This variable was excluded from the model in order to ensure a non-singular matrix.

(b) All employees who possessed this characteristic were promoted; thus, this variable perfectly predicts promotions. The coefficient is automatically set to 500 by the LOGIST procedure and the values of the remaining coefficients are adjusted in order to fit the likelihood function.

Table 6. Estimated Logistics Models: All Employees Estimated Coefficients

Appendix A: Program

Second Workshop on Law and Justice Statistics

Sponsored by ASA's Continuing Education Department
and
The Bureau of Justice Statistics*

Sheraton Centre, Toronto, Canada

Saturday-Sunday, August 13-14, 1983

This workshop was developed by the ASA Committee on Law & Justice Statistics.

S. James Press, Chair
Alan E. Gelfand, Vice-Chair & Workshop Coordinator

Saturday, August 13

Morning

Session I — Chair: Katherine K. Wallman, Council of Professional Associations on Federal Statistics

9:00- 9:10 Introductory Remarks:
S. James Press, University of California, Riverside
Alan E. Gelfand, University of Connecticut

9:10- 9:30 Bureau of Justice Statistics: Present and Future Plans
Ralph Rossum, Bureau of Justice Statistics

9:30-10:00 Providing Statistical Support to the Policy Process
Sue Ann Lindgren, Bureau of Justice Statistics

10:00-10:30 Issues in Using Statistics in the Policy Process
Marion R. Metcalf, Office of Policy & Management Analysis

10:30-10:50 M O R N I N G B R E A K

10:50-11:20 Structure and Process of Collecting and Analyzing Justice Statistics in Canada
Gaylen A. Duncan, Canadian Centre for Justice Statistics

11:20-12:00 The Organization of the Ministry of the Solicitor General of Canada and the Role of the Statistics Policy Advisor
Penny Reedie, Ministry of Solicitor General, Canada

12:00- 2:00 L U N C H B R E A K

Afternoon

Session II — Statistical Methodology in Law & Justice Statistics (Part One)

Chair: George G. Woodworth, University of Iowa

2:00- 2:50 Analysis of a Y-Stratified Sample: The Georgia Charging and Sentencing Study
George G. Woodworth, University of Iowa

2:50- 3:30 Some Consequences of Convenience Samples in Criminal Justice Research
Richard A. Berk, University of California, Santa Barbara

3:30- 3:50 A F T E R N O O N B R E A K

Statistical Analysis in the Courts

Chair: Albert J. Reiss, Jr., Yale University

3:50- 4:25 The Civil Litigation Research Project: Lessons for Studying the Civil Justice System
Herbert M. Kritzer, University of Wisconsin, Madison

4:25- 5:00 An Examination of U.S. Appellate Court Opinions
Alan E. Gelfand, University of Connecticut

*With special acknowledgment to the Criminal Justice Statistics Association for assistance in planning the workshop.

Sunday, August 14

Morning

Session III — Chair: Thomas A. Henderson, Criminal Justice Statistics Association

9:00- 9:40 Data Analysis Problems at the State Level: Data Problems and Data Collection
Michael H. Rabasca, Statistical Analysis Center, Delaware

9:40-10:30 Policy Relevance in Criminal Justice Research
Paul Stageberg and Daryl Fischer, Statistical Analysis Center, Iowa

10:30-10:50 MORNING BREAK

10:50-11:25 Is Crime Seasonal?
Carolyn Rebecca Block, Statistical Analysis Center, Illinois

11:25-12:00 Alternatives to Time-Series Analysis in Population Projections
John P. O'Connell, Statistical Analysis Center, Washington

12:00- 2:00 LUNCH BREAK

Afternoon

Session IV — Statistical Methodology in Law & Justice Statistics (Part Two)

Chair: Alan E. Gelfand, University of Connecticut

2:00- 2:30 Establishing Causation: The Role of Epidemiological Evidence
Shanna Helen Swan, Department of Health Services, California

2:30-3:00 Two Statistical Methods for Examining Claims of Employment Discrimination
Joseph Gastwirth, George Washington University

3:00- 3:30 A Probabilistic Approach to Tracing Presumptions in the Law of Restitution
Michael O. Finkelstein, Barrett, Smith, Shapiro, Simon and Armstrong, New York

3:30- 3:50 AFTERNOON BREAK

The Statistician as Expert Witness

3:50- 5:00 Statistics in Employment Discrimination: The Use of Expert Witnesses
James Beckett, Criterion Incorporated
Peter Lewin, University of Texas, Dallas

Appendix B: Addresses of Speakers

James Beckett
Criterion Incorporated
13140 Coit Road, Suite 318
Dallas, TX 75240

Richard A. Berk
Department of Sociology
University of California, Santa Barbara
Santa Barbara, CA 93106

Carolyn Rebecca Block
Statistical Analysis Center
Illinois Criminal Justice Information Authority
120 South Riverside Plaza
Chicago, IL 60606

Gaylen A. Duncan
Canadian Centre for Justice Statistics
19th Floor, R.H. Coats Building
Tunney's Pasture
Ottawa, Ontario, K1A 0T6
Canada

Michael O. Finkelstein
Barrett, Smith, Shapiro, Simon and Armstrong
26 Broadway
New York, NY 10004

Daryl Fischer
Office for Criminal & Juvenile Justice Planning
523 East 12th Street
Des Moines, IA 50319

Joseph Gastwirth
Department of Statistics
George Washington University
Washington, D.C. 20052

Alan E. Gelfand
Department of Statistics
University of Connecticut
Storrs, CT 06268

Thomas A. Henderson
Criminal Justice Statistics Association, Inc.
444 North Capitol Street, Suite 122
Washington, D.C. 20001

Herbert M. Kritzer
Department of Political Science
University of Wisconsin, Madison
Madison, WI 53706

Peter Lewin
Criterion Incorporated
13140 Coit Road, Suite 318
Dallas, TX 75240

Sue A. Lindgren
U.S. Bureau of Justice Statistics
633 Indiana Avenue, N.W.
Washington, D.C. 20531

Marion R. Metcalf
Office of Policy and Management Analysis
Criminal Division
U.S. Department of Justice
Room 2214
Washington, D.C. 20530

John P. O'Connell
Statistical Analysis Center
Forecasting and Estimation Division
400 East Union, ER-13
Olympia, WA 98504

S. James Press
Department of Statistics
University of California, Riverside
Riverside, CA 92521

Michael H. Rabasca
Statistical Analysis Center
Suite 207, Treadway Towers
Dover, DE 19901

Penny Reedie
Ministry of Solicitor General
Sir Wilfred Laurier Building
340 Laurier Avenue, West
Ottawa, Ontario K1A 0P8
Canada

Albert J. Reiss, Jr.
Department of Sociology
Yale University
Box 1965, Yale Station
New Haven, CT 06520

Ralph Rossum
U.S. Bureau of Justice Statistics
633 Indiana Avenue, N.W.
Washington, D.C. 20531

Paul Stageberg
Office for Criminal & Juvenile Justice Planning
523 East 12th Street
Des Moines, IA 50319

Shanna H. Swan
Department of Health Services
2151 Berkeley Way
Berkeley, CA 94704

Katherine Wallman
Council of Professional Associations on Federal Statistics
806 15th Street, N.W., Suite 440
Washington, D.C. 20005

George G. Woodworth
Department of Statistics
University of Iowa
Iowa City, IA 52240

☆ U.S. GOVERNMENT PRINTING OFFICE: 1984-421-934:11212

NCJ-93310, 6/84
Proceedings of 2nd Conference
on Law & Justice Statistics

NCJRS REGISTRATION

The National Criminal Justice Reference Service (NCJRS) abstracts documents published in the criminal justice field. Persons who are registered with the Reference Service receive announcements of documents in their stated fields of interest and order forms for free copies of Bureau of Justice Statistics publications. If you are not registered with the Reference Service, and wish to be, please provide your name and mailing address below and check the appropriate box.

Name **Telephone** ( )	☐ **Please send me a NCJRS registration form.**
Number and street	☐ **Please send me the reports listed below.**
City **State** **ZIP Code**	

(Fold here)

U.S. DEPARTMENT OF JUSTICE
Bureau of Justice Statistics
Washington, D.C. 20531

PLACE STAMP HERE

User Services Department 2
National Criminal Justice Reference Service
Bureau of Justice Statistics
U.S. Department of Justice
Box 6000
Rockville, Maryland 20850

(Fold here)

If you wish to receive copies of any of the Bureau of Justice Statistics Reports listed on the reverse side, please list them below.

_____

_____

_____

_____

_____

Bureau of Justice Statistics reports

(revised June 1984)

Single copies are available free from the National Criminal Justice Reference Service, Box 6000, Rockville, Md. 20850 (use NCJ number to order). Postage and handling are charged for multiple copies: up to 10 titles free; 11-40, $10; more than 40, $20; libraries call for special rates (301/251-5500).

Public-use tapes of BJS data sets and other criminal justice data are available from the Criminal Justice Archive and Information Network, P.O. Box 1248, Ann Arbor, Mich. 48106, (313/764-5199).

National Crime Survey

Criminal victimization in the U.S.:
- **1973-82** trends, NCJ-90541, 9/83
- **1981** (final report), NCJ-90208
- **1980** (final report), NCJ-84015, 4/83
- **1979** (final report), NCJ-76710, 12/81

BJS special reports:
- **The economic cost of crime to victims,** NCJ-93450, 4/84
- **Family violence,** NCJ-93449, 4/84

BJS bulletins:
- **Criminal victimization 1983,** NCJ-93869, 6/84
- **Households touched by crime,** 1983, NCJ-93658, 5/84
- **Violent crime by strangers,** NCJ-80829, 4/82
- **Crime and the elderly,** NCJ-79614, 1/82
- **Measuring crime,** NCJ-75710, 2/81

The National Crime Survey: Working papers, vol. I: Current and historical perspectives, NCJ-75374, 8/82

Crime against the elderly in 26 cities, NCJ-76706, 1/82

The Hispanic victim, NCJ-69261, 11/81

Issues in the measurement of crime, NCJ-74682, 10/81

Criminal victimization of California residents, 1974-77, NCJ-70944, 6/81

Restitution to victims of personal and household crimes, NCJ-72770, 5/81

Criminal victimization of New York State residents, 1974-77, NCJ-66481, 9/80

The cost of negligence: Losses from preventable household burglaries, NCJ-53527, 12/79

Rape victimization in 26 American cities, NCJ-55878, 8/79

Criminal victimization in urban schools, NCJ-56396, 8/79

Crime against persons in urban, suburban, and rural areas, NCJ-53551, 7/79

An introduction to the National Crime Survey, NCJ-43732, 4/78

Local victim surveys: A review of the issues, NCJ-39973, 8/77

National Prisoner Statistics

BJS bulletins:
- **Capital punishment 1983,** NCJ-93925, 7/84
- **Prisoners in 1983,** NCJ-92948, 4/84
- **Prisoners 1925-81,** NCJ-85861, 12/82

Prisoners in State and Federal institutions on December 31, 1981 (final report), NCJ-86485, 7/83

Capital punishment 1981 (final report), NCJ-86484, 5/83

1979 survey of inmates of State correctional facilities and 1979 census of State correctional facilities:

BJS special reports:
- **The prevalence of imprisonment,** NCJ-93657, 6/84
- **Career patterns in crime,** NCJ-88672, 6/83

BJS bulletins:

Prisoners and drugs, NCJ-87575, 3/83

Prisoners and alcohol, NCJ-86223, 1/83

Prisons and prisoners, NCJ-80697, 2/82

Veterans in prison, NCJ-79632, 11/81

Census of jails and survey of jail inmates:

Jail inmates 1982 (BJS bulletin), NCJ-87161, 2/83

Census of jails, 1978: Data for individual jails, vols. I-IV, Northeast, North Central, South, West, NCJ-72279-72282, 12/81

Profile of jail inmates, 1978, NCJ-65412, 2/81

Census of jails and survey of jail inmates, 1978, preliminary report, NCJ-55172, 5/79

Parole and probation

BJS bulletins:
- **Probation and parole 1982,** NCJ-89874 9/83
- **Setting prison terms,** NCJ-76218, 8/83

Characteristics of persons entering parole during 1978 and 1979, NCJ-87243, 5/83

Characteristics of the parole population, 1978, NCJ-66479, 4/81

Parole in the U.S., 1979, NCJ-69562, 3/81

Courts

The prosecution of felony arrests, 1979, NCJ-86482, 5/84

Habeas corpus (BJS special report), NCJ-92949, 3/84

State court caseload statistics: 1977 and 1981 (BJS special report), NCJ-87587, 2/83

State court organization 1980, NCJ-76711, 7/82

State court model statistical dictionary, NCJ-62320, 9/80

A cross-city comparison of felony case processing, NCJ-55171, 7/79

Federal criminal sentencing: Perspectives of analysis and a design for research, NCJ-33683, 10/78

Variations in Federal criminal sentences, NCJ-33684, 10/78

Federal sentencing patterns: A study of geographical variations, NCJ-33685, 10/78

Predicting sentences in Federal courts: The feasibility of a national sentencing policy, NCJ-33686, 10/78

State and local prosecution and civil attorney systems, NCJ-41334, 7/78

Expenditure and employment

Justice expenditure and employment in the U.S., 1979 (final report), NCJ-87242, 12/83

Justice expenditure and employment in the U.S., 1971-79, NCJ-92596 (forthcoming)

Privacy and security

Computer crime:
- **Electronic fund transfer and crime,** NCJ-92650, 2/84
- **Computer security techniques,** NCJ-84049, 9/82
- **Electronic fund transfer systems and crime,** NCJ-83736, 9/82
- **Legislative resource manual,** NCJ-78890, 9/81
- **Expert witness manual,** NCJ-77927, 9/81
- **Criminal justice resource manual,** NCJ-61550, 12/79

Privacy and security of criminal history information:
- **A guide to research and statistical use,** NCJ-69790, 5/81
- **A guide to dissemination,** NCJ-40000, 1/79
- **Compendium of State legislation:** NCJ-48981, 7/78
- **1981 supplement,** NCJ-79652, 3/82

Criminal justice information policy:
- **Research access to criminal justice data,** NCJ-84154, 2/83
- **Privacy and juvenile justice records,** NCJ-84152, 1/83
- **Survey of State laws** (BJS bulletin), NCJ-80836, 6/82
- **Privacy and the private employer,** NCJ-79651, 11/81

General

BJS bulletins:
- **Federal drug law violators,** NCJ-92692 2/84
- **The severity of crime,** NCJ-92326, 1/84
- **The American response to crime: An overview of criminal justice systems,** NCJ-91936, 12/83
- **Tracking offenders,** NCJ-91572, 11/83
- **Victim and witness assistance: New State laws and the system's response,** NCJ-87934, 5/83
- **Federal justice statistics,** NCJ-80814, 3/82

Sourcebook of Criminal Justice Statistics, 1983, NCJ-91534, forthcoming 10/84

Report to the nation on crime and justice: The data, NCJ-87068, 10/83

BJS five-year program plan, FY 1982-86, 7/82

Violent crime in the U.S. (White House briefing book), NCJ-79741, 6/82

Dictionary of criminal justice data terminology: Terms and definitions proposed for interstate and national data collection and exchange, 2nd ed., NCJ-76939, 2/82

Technical standards for machine-readable data supplied to BJS, NCJ-75318, 6/81

Justice agencies in the U.S., 1980, NCJ-65560, 1/81

Indicators of crime and criminal justice: Quantitative studies, NCJ-62349, 1/81

A style manual for machine-readable data, NCJ-62766, 9/80

Please put me on the mailing list(s) for:

- ☐ **All BJS reports** — 30 to 40 reports a year
- ☐ **BJS Bulletin** — timely reports of the most current justice data
- ☐ **Courts reports** — State court caseload surveys, model annual State court reports, State court organization surveys
- ☐ **Corrections reports** — results of sample surveys and censuses of jails, prisons, parole, probation, and other corrections data
- ☐ **National Crime Survey** — the Nation's only regular national survey of crime victims
- ☐ **Sourcebook of Criminal Justice Statistics** — a broad spectrum of data from 153 sources in an easy-to-use, comprehensive format (433 tables, 103 figures, index)
- ☐ Please send me current report(s) checked on the list above.

END