U.S. Department of Justice
Bureau of Justice Statistics

Indicators of crime and criminal justice: Quantitative studies

NCJ-62349, June 1980

Edited by

STEPHEN E. FIENBERG
Department of Statistics
and Department of Social Science
Carnegie-Mellon University

ALBERT J. REISS, JR.
Department of Sociology
Yale University

For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402

U.S. Department of Justice
Bureau of Justice Statistics

Harry A. Scarr, Ph.D.
Director

The Social Science Research Council is a private nonprofit organization formed for the purpose of advancing research in the social sciences. It emphasizes the planning, appraisal, and stimulation of research that offers promise of increasing knowledge in social science or of increasing its usefulness to society. It is also concerned with the development of better research methods, improvement of the quality and accessibility of materials for research by social scientists, and augmentation of resources and facilities for their research.

The Council's Center for Coordination of Research on Social Indicators was established in September 1972, under a grant from the Division of Social Sciences, National Science Foundation (grant number GS-34219; currently SOC-77-21686). The Center's purpose is to enhance the contribution of social science research to the development of a broad range of indicators of social change, in response to current and anticipated demands from both research and policy communities.

Library of Congress Cataloging in Publication Data
Main entry under title:

Indicators of crime and criminal justice.

Based on research initiated during the Workshop in Criminal Justice Statistics, held in the summer of 1975 by the Social Science Research Council.
"NCJ-62349."
Includes bibliographical references.
Supt of Docs. no.: J 29.2:In2
1. Criminal statistics—Addresses, essays, lectures.
I. Fienberg, Stephen E. II. Reiss, Albert J. III. Workshop in Criminal Justice Statistics, 1975. IV. Social Science Research Council.
HV6018.I52 364 80-607926

This project was supported by Grant No. 75-SS-99-6017, awarded to the Social Science Research Council by the Statistics Division, National Criminal Justice Information and Statistics Service, Law Enforcement Assistance Administration (now Bureau of Justice Statistics), U.S. Department of Justice, under the Omnibus Crime Control and Safe Streets Act of 1968, as amended. The project, entitled "Workshop and Fellowship in Criminal Justice Statistics," was directed for the Social Science Research Council by David Seidman and monitored by Paul D. White of BJS. Points of view or opinions stated in this document are those of the author(s) and do not necessarily represent the official position or policies of the U.S. Department of Justice.

Preface

The Social Science Research Council's Center for Coordination of Research on Social Indicators, organized in 1972, has had an ongoing interest in promoting quality research in the area of criminal justice statistics, especially at a national level. To this end, the Center's Advisory and Planning Committee created in 1973 a special Subcommittee on Criminal Justice Statistics, whose focus has been on understanding and measuring aspects of crime and the criminal justice system.

The present volume can be traced back to a meeting of the members of the Subcommittee with staff at the Law Enforcement Assistance Administration (LEAA) to discuss research projects of mutual interest. The two of us (Stephen E. Fienberg and Albert J. Reiss, Jr.) were convinced that an upgrading of the quality of research in the criminal justice area could be affected through the recruitment of statisticians and other quantitative social scientists to work on interesting statistical modelling and analysis projects in the area. We therefore proposed to hold a workshop focussing on the methodological and theoretical problems involved in the measurement of crime and its effects on society.

Thus it was that in the summer of 1975, the Social Science Research Council, with funding from LEAA* (now the Bureau of Justice Statistics), organized a month-long Workshop in Criminal Justice Statistics. The workshop brought together three groups: a faculty composed of seven experts in criminal justice; seven participants who were young academic statisticians, sociologists, and psychologists; and six government professionals.

The workshop's program of activities started with a series of lectures and seminars and concluded with the organization of the faculty and the participants into six working groups. Among the topics covered at the plenary meetings were conceptions of the criminal justice system, data sources, research approaches, methodological and substantive problems, and theoretical issues. The intensive interchange at these sessions sparked a high degree of interest in the six working groups, each of which then undertook a concentrated exploration of a problem area in criminal justice research. The topics on which they focused were the treatment of data in victimization surveys; measurement and experimental design in the study of police effectiveness; the sentencing process; decisionmaking by the U.S. Board of Parole; the deterrent effect of capital punishment; and the problem of designing a model of the criminal justice system.

The papers in the present volume are a direct outgrowth of research initiated during or as a result of the workshop activities. Earlier drafts of many of the papers were presented and discussed at a workshop reunion which took the form of a conference held in the fall of 1977, again sponsored by LEAA. The first paper in this volume reports on the highlights of that conference, and provides an overview to the frontiers of quantitative research in this area. The remaining sections of the volume are organized in large part around the research produced by the six working groups from the workshop.

The faculty of the workshop consisted of Albert D. Biderman, Bureau of Social Science Research (Washington, D.C.); Alfred Blumstein, Carnegie-Mellon University; John Clark, University of Minnesota; Michael J. Hindelang, State University of New York at Albany; Albert J. Reiss, Jr., Yale University; Richard Sparks, Rutgers University; and Leslie T. Wilkins, State University of New York at Albany.

The social scientist participants in the workshop were David W. Britt, Florida Atlantic University (currently at Nova University); William B. Fairley, Harvard University (currently a private consultant); Stephen E. Fienberg, University of Minnesota; Gary G. Koch, University of North Carolina; Kinley Larntz, University of Minnesota; Colin Loftin, Brown University (currently at the University of Michigan); and Howard Wainer, University of Chicago (currently at the Bureau of Social Science Research).

Government professionals participating in the workshop were Ken Brimmer, Mimi Cantwell, Linda Murphy, Tom Petersik, and J. Frederick Shenk, all of the U.S. Bureau of the Census; and Cynthia Turnure, Minnesota Governor's Commission on Crime Prevention and Control.

One of the most pleasing outcomes of the workshop (at least to us) has been the continuing work of the participants and faculty on methodological problems in criminal justice research. They have been intimately involved in such activities as: (a) the National Academy of Sciencies–National Research Council's Committee on Law Enforcement and Criminal Justice and that Committee's Panels on Deterrence and Incapacitation, and on Rehabilitation, (b) the NAS–NRC Committee on National Statistics and that Committee's Panel on the National Crime Survey, (c) the American Statistical Association's Committee on Law and Criminal Justice Statistics, and (d) SSRC's Advisory and Planning Committee on Social Indicators. They have also initiated new research projects on a wide variety of problems in the area. This continued activity is perhaps the best indicator of the success of the Workshop.

Almost as pleasing are the many professional collaborations and personal friendships that developed among those who took part in the Workshop. These go far beyond the work reported on in this volume.

We acknowledge with thanks the contributions of David Seidman of the Social Science Research Council, and of Elaine Javonovich, his assistant, who staffed the Workshop; and of the staff and director of the Council's Center for Social Indicators for their support and interest. Special thanks are due to Nancy McManus of the Social Science Research Council, who worked closely with us in structuring the volume, extracting manuscripts from authors, and copyediting the entire volume. Ms. McManus was responsible for seeing the manuscript through to completion, and we are indebted to her for her assistance.

STEPHEN E. FIENBERG
ALBERT J. REISS, JR.

*Grant No. 75SS-99-6017.

Abstract

The quantitative study of criminal justice is a rapidly expanding field, involving the application of relatively sophisticated statistical methods and probabilistic models. This volume brings together some of the products of a workshop in criminal justice statistics sponsored by the Social Science Research Council and the Law Enforcement Assistance Administration (now the Bureau of Justice Statistics) during the summer of 1975. An introductory essay describes the highlights of a followup conference held in October 1977, at which workshop participants described some of their subsequent research. This essay provides an overview of the volume and of the frontiers of quantitative research in this area. The remaining papers consist of detailed versions of conference presentations, as well as reports on other related research in such diverse areas as: macro models for criminal justice planning; the deterrent effects of punishment on crime; effects of plea-bargaining on case disposition; criminal victimization and models for its analysis; victim proneness; parole decision-making, and police patrol experiments.

Contents

Current research in criminal justice statistics: Report of a conference*

DAVID SEIDMAN**
Social Science Research Council
Washington, D.C.

The quantitative study of crime and criminal justice is a rapidly developing field of research. While perhaps most of this development can be attributed at least indirectly to the apparent rise in crime beginning in the 1960's, advances in particular areas can be traced to more proximate causes. The creation of major new data bases is perhaps the most obvious stimulus to quantitative research, and here the invention and institutionalization of victimization surveys have been particularly notable. There has been, within the criminal justice system itself, increased self-awareness, and this has stimulated attempts to rationalize the decisionmaking process within the system. The basis of rationalization, in a number of instances, has been quantitative study of the actual operations of the system, combined with attempts to develop prediction methods which would allow for new approaches to decision making. Finally, there has been increasing questioning of the effects of the criminal justice system on offenders, on the crime rate, and on society as a whole; this questioning has led to attempts to analyze quantitatively what those effects are, in ways which may provide guidance for criminal justice policy in the future.

These three areas of research were intensely studied at the Social Science Research Council's Workshop in Criminal Justice Statistics during the summer of 1975. Subsequent to the workshop, participants carried out research in these areas and others. The conference held in October 1977 provided an opportunity for presentation of research results and for discussion of research plans. The presentations and discussions do not represent the full range of research in criminal justice carried out by workshop participants since the summer of 1975, but focus on the three areas of research discussed above and emphasize research supported by the Law Enforcement Assistance Administration grant to the Social Science Research Council which made the workshop possible.

This report draws upon both the written conference papers and the discussions to present highlights of the conference in the the three main areas considered. It does not pretend to capture the full richness of the presentations, and the reader should turn to the conference papers for the complete picture. Together, the report and the papers provide an overview of current research on the frontiers of the quantitative study of crime and criminal justice and an assessment of current problems and directions for future research.

Victimization studies

MOST PUBLISHED reports of victimization surveys amount to little more than calculations of "victimization rates" for populations and subpopulations for particular time periods. While these calculations may be useful, they leave unresolved certain conceptual issues, contribute little to an understanding of crime and victimization, and fail to exploit the full potential of the surveys (in particular, the longitudinal features of the National Crime Panel Survey are underexploited). Several workshop participants have been studying victimization surveys in ways which probe more deeply into the structures and processes underlying the data. The primary focus has been on problems related to multiple and series victimization and the interactions of survey design features and resulting data. Considerable time at the conference was devoted to presentation and discussion of this work.

A useful introduction to problems of multiple victimization is a well-known finding concerning the distribution of individuals by number of reported victimizations within the reference period. Richard Sparks presented one version of this distribution, based on his London victimization survey. (See Sparks, Table 1.)[1] What is at issue here is the sense in which the overall victimization rate characterizes the risk of victimization for individuals in the group. If it is assumed that the overall rate represents the risk faced by individuals, then the distribution of individuals by number of victimizations ought to fit a Poisson distribution. The last column of Sparks's table shows the expected number of individuals by number of victimizations for a Poisson process. The standard finding emerges: "the numbers reporting no victimization at all, and the numbers reporting several incidents, are both greater than would be predicted by a Poisson distribution" (Sparks).

This result is a starting point for speculation and investigation. Sparks, for example, points out that slightly more complicated models fit the data better. In particular, models based on either "contagion" or heterogeneity of victim proneness in the population work reasonably well. Unfortunately, the two models cannot be differentiated in cross-sectional data, which points to the importance of exploiting the longitudinal features of NCS. Sparks finds the contagion notion an implausible explanation of the data, and this discussion will pursue the victim proneness direction.

If the explanation of the lack of fit of the Poisson model is that individuals differ in their victim proneness, a logical next step is to sort out individuals by their degrees of victim proneness, and see if the Poisson model fits within subgroups of the population. Sparks tried this with his London

*A report to the Bureau of Justice Statistics by the Social Science Research Council, 605 Third Avenue, New York, NY 10016, pursuant to Order Number 8-0073-J-LEAA. This report is based upon working papers presented at the conference. The quotations and materials discussed will therefore not necessarily coincide with the versions of the papers published in this volume.

**In September 1978 David Seidman left the Social Science Research Council for the Rand Corporation, Washington, D.C.

[1]Citations without dates refer to papers and other materials which served as the basis for conference presentations. Other citations are to materials included in the list of references at the end of this report.

data, but could not find any combination of attributes, such as age, sex, race, or social class, which could be used to categorize individuals to produce subgroups within which the Poisson model adequately fit the data. As he notes, it is not surprising that these simple demographic variables would be insufficient.

It is possible that the observed pattern actually results from measurement problems in the surveys. One such measurement problem has to do with series incidents. Typically the number of incidents in a series is estimated by the victims, and there appears to be a tendency to overestimate the number—a phenomenon common to a number of different areas of investigation. For a variety of reasons, incorporation of series victimizations into the distributions here considered is difficult, and their exclusion may substantially affect the shape of the distribution. Some indication of the effect can be obtained by examining such questions as whether series victims also report other victimizations, but this examination is difficult in a small cross-sectional survey. Another measurement problem is nonreporting. It may be, for example, that most nonreporting is by people who have been victims only once. Since the departures from the Poisson distribution are largely in the zero and one victimization groups, this would account for much of the lack of fit. Accounting for that pattern of nonreporting, however, is difficult.

If the deviation from the Poisson distribution appears to be mainly a matter of too many people reporting no victimizations, it seems reasonable to fit a Poisson distribution to the distribution of those reporting at least one victimization, and then project backwards to estimate the number "belonging" in the zero victimization group under Poisson assumptions. The "excess" actually found in that category might then be thought of as drawn from a population somehow lacking vulnerability to crime. The source of this immunity to crime—or at least to certain kinds of crime—is not immediately obvious. Presumably it has something to do with residential location or life style. Of course, while a model postulating an immune group and a group subject to Poisson process may fit the data, isolating the immune group for analysis may not be possible, since there is no obvious way to separate it from the group which, as a result of the random process, was not victimized in the reference period.

The plausibility of an absolutely immune group declines when a time perspective longer than the reference period of a single cross-sectional survey is adopted. This begins to suggest the importance of exploiting the over-time features of the National Crime Panel Survey. There are, of course, other reasons for relying upon longitudinal files of the NCS for analysis of multiple and series victimization. For one thing, the large number of incidents available in such files allows disaggregation to crime-specific victimization. For another, repeated interviews can be used to build long sequences of victimization experience, which allow, among other things, investigation of the stability of patterns. Finally, longitudinal files allow investigation of the time between victimizations, which may be important to an understanding of multiple victimization.

Albert J. Reiss, Jr., has constructed what we believe to be the only extant working longitudinal file of NCS data. He presented some results of analysis of this file to the conference. These results concerned the relatively limited question of whether in repeated victimization of a given unit there is a propensity for repeated victimization by the same type of crime, or whether repeated victimization is random as to crime type. While the question is limited, it appears central to an understanding of victim proneness and multiple victimization.

The analytical tool Reiss used is a crime switch matrix consisting of pairs of victimizations, where a preceding victimization by type of crime is compared with a following victimization by type of crime.

The construction of such a matrix presents a number of questions and problems. First, what is the unit at risk? Confining his analysis to households as units, Reiss notes that victimization of individuals within households may symbolically represent victimization of the household. (Similarly, individuals may not sharply distinguish household victimization from personal victimization when the focus is on personal victimizations. Because of the choice of focusing on households, Reiss's analysis is unaffected by this problem.) It follows that if the household is the unit of analysis, a researcher may either limit the analysis to household victimizations or include as well victimizations of individuals within the household. Reiss examined the data both ways.

Second, the matrix depends upon the time sequence of victimizations. NCS incident reports record the month of nonseries victimizations. Victimizations of a household within a month, however, lack the information required to establish the time sequence. The timing of series victimizations presents a more complicated problem, as the NCS records only the month of occurrence of the first incident in the series (and if the series began before the reference period, the date recorded is the first month of the reference period). Reiss therefore had to impose an arbitrary time order so as to establish a sequence for all reported victimizations. He did this by randomly assigning dates within months to events, with series treated as single events. The error introduced into the ordering by this random procedure is likely to be relatively small for nonseries incidents, as incidents occurring in different months are unaffected. The errors introduced by treating series victimizations as single incidents and then placing them (at a randomly selected date) in the month the series begins (or the first month of the reference period) are harder to assess because of the general paucity of information about series victimizations. However, Reiss has found that those who report series victimizations are highly unlikely to report other victimizations, so that the primary effect of his treatment of series victimizations is likely to be a reduction in the proportion of cases in a crime switch matrix which fall on the diagonal of the matrix (since events within a series are in theory all of the same kind). The significance of this will become clear as the results are discussed.

A preliminary examination of the data (Reiss's "Victim Proneness") enables one to compare the distributions of victimization incidents by type of crime where both incidents occurring in sequences and single victimizations are included with equivalent distributions not including single victimizations. It is apparent that the distributions are substantially similar, which leads Reiss to conclude that "multiple or repeat household victims are on the whole no more prone to victimization by certain types of crime than are all victims." If multiple victimization reflects victim proneness, it nevertheless seems not to reflect proneness to particular types of crime.

Reiss's crime switch matrix further exposes the structure of multiple victimization. As is perhaps clearer from a chi-squared analysis of the matrix, the major element of structure in the matrix is the heaping of incidents along the diagonal. The implication is that "the chances that a household or one or more of its members will be successively victimized by the same type of crime are greater than one would expect given the chances all households and their members have of being victimized by that type of crime." The analysis continues by examining whether the frequency of particular pairs depends upon which element of the pair occurs first. For example, is the sequence assault-personal larceny more or less frequent than the sequence personal larceny-assault? Reiss's Table 4 strongly suggests that the order within pairs makes no difference; that the structure is symmetrical with respect to order of occurrence.

Reiss's paper merely establishes these elements of structure. Conference participants offered some speculative explanations. Some possible explanations turn on methodological artifacts: reporting definitions, response bias, interviewing techniques, and the like. More substantive explanations focused on the characteristics of locations, patterns of communication, and the elusive concept of victim proneness.

It was noted that, because the NCS is a survey of household locations, there is in theory the opportunity to explore victim proneness and location proneness by examining over time (1) victimization experiences at locations where there is a change in individuals, and (2) victimization experiences of individuals who move out of a survey location. The importance of this kind of examination is suggested by the relationship between being victimized and moving. Reiss pointed out that both series victims and other multiple victims are highly likely to move. The questions which need to be answered are whether the high rate of victimization continues with a new household at the surveyed location and whether the high rate of victimization continues for the mover at a new location. While in theory the survey design might permit the first question to be addressed, in practice it does not, because the records do not carefully distinguish between new and old residents in continuing locations and because, in the case of new residents, no information is sought about victimization experience at the prior location. (That is, information is obtained which, depending upon the time of the move, may relate to the new resident's prior residence, but there is, in the recorded data, no way to distinguish between victimization at the old location and victimization at the new one.) The second question cannot be addressed at all, because no attempt is made to follow movers to their new residential locations. It might be possible to approach the problem relatively cheaply by adding to interviews with new residents a set of questions designed to elicit retrospective information about victimization experiences, though recall decay problems would be a substantial obstacle to obtaining reliable information.

Also important to questions of victim and location proneness is the character of the neighborhood of the interview location. Reiss has merged certain neighborhood characteristics onto his victimization tape, but what is perhaps the most important neighborhood characteristic for these purposes, the neighborhood crime rate, is unavailable.

Some questions about these patterns of multiple victimization relate to the length of time in panel for respondents. The Reiss data set contains a high proportion of respondents interviewed only a small number of times. It seems likely that the proportion of single victimizations would decline as the sample "ages." It is also possible that some of the concentration on the diagonal of the crime switch matrix would decline, the pattern of victimization followed by victimization of the same kind declining as time between victimizations increases. Reiss in other work is examining time between victimizations, but results of that work were not presented at the conference.

Stephen Fienberg has been developing models for analysis of victimization data in their full longitudinal structure, with particular attention to problems of multiple victimization and various time-related effects in the data, such as the effects of month of collection, time lag to reference month, and time in panel. A preliminary sketch of this approach is contained in section 5 of his "Victimization and the National Crime Survey: Problems of Design and Analysis." He expanded upon these ideas in a presentation at the conference.

In Fienberg's view, analysis of questions of multiple victimization depends upon models for the occurrence of victimizations over time. Without necessarily suggesting that it represents the true underlying structure, Fienberg proposes modelling victimization as a particular form of stochastic point process, a semi-Markov process. (The Poisson process, discussed above, is a special case of the semi-Markov process.) Such a process has two main features. First, the probability that one is next victimized by a crime of type j depends only upon the crime type of which one was last a victim. That is, the transition from victimization state to victimization state is governed by a matrix of probabilities resembling Reiss's crime switch matrix. Second, the interval between victimizations is random, though the distribution of time between victimizations depends upon the current victimization state. A convenient feature of semi-Markov processes is that, for many purposes, the transition matrix and the time interval distributions can be studied separately. The model is based on continuous time, while victimization data collection is one month at a time. Integrating these two different approaches presents modelling problems.

Fienberg also notes that the hierarchical structure of the victimization surveys (individuals within families, families within households, households within locations, locations within neighborhoods) complicates the modelling approach. These added complications will not be discussed here.

A fundamental question is whether the process is time homogeneous, or whether there are shifts in the process related to, say, time in panel or other causes. To answer this question and others relating to time between events, Fienberg proposes a research approach along the following lines. In order to obtain the required long records, he begins with a subfile of respondents with full 3½ year records. Assume we start by considering the household as the unit of interest. Then there are 42 months in which an incident of victimization may be reported. As a first step, consider only whether the household has or has not been victimized within a given month, ignoring differences among types of crime and multiple victimization within a month. (Note that the NCS treatment of series victimization causes difficulties here.) Then there are 242 possible time sequence patterns, a number far too large for analysis. In practice, the number of patterns found would probably be much smaller; Fienberg guesses that about 20 patterns would account for 95 percent of the records. It would then be possible to examine the time patterns for shifts associated with time in panel. As a next step, households could be scored each month according to whether there has been household victimization, personal victimization, or no victimization, and the same kind of analysis could be done. The analysis would then continue, examining finer grain structure at each step. The information on time sequencing of victimizations generated in this way feeds back into the models of generation of multiple victimization which are at the heart of the investigation. The discovery of time heterogeneity in long records would likely allow for various transformations inducing homogeneity. In addition to casting light on the underlying processes, such transformations would provide important information about the interpretation of aggregate victimization rates in cross-sectional analysis. The problem of the ill-fitting Poisson can be addressed, and records sorted in ways allowing examination of victim proneness.

Several of the conferees questioned whether it was wise to begin with a subfile of those with 3½ years of data, suggesting instead starting with an entering cohort and examining it over a six-month period. Reiss suggested that there will be sufficient multiple victimization within a six-month period to allow the kind of analysis Fienberg had proposed for the longer records. The suggested advantages of the entering cohort approach are that it avoids gaps in records (which Fienberg had noted as a complication facing his approach) and it significantly reduces problems associated with panel attrition. And, of course, the number of possible patterns is substantially smaller with 12 months than with 42 months.

Fienberg's answer was that time heterogeneity was more likely to be discoverable with the longer records. The conferees then suggested starting with entering cohorts and "aging" them. It is clear that if they are aged far enough, with movers deleted, it makes no difference whether one starts with entering cohorts and works forward or starts with completely aged cohorts: the same records are included. And starting in either place one might compare different cohorts at similar points in their "life cycles" in the panel to investigate stationarity of victimization apart from the question of time homogeneity Fienberg discussed. The question separating the two approaches, then, is how long a record is necessary to investigate time homogeneity. That, however, cannot be answered *a priori*. In the absence of adequate prior investigations, the only way to know whether it is necessary to use Fienberg's approach as compared with relying upon shorter records is to compare findings based on the longest possible records with those based on shorter ones.

The discussion of the two possible starting points for the investigation Fienberg proposed pointed once again to the problem of panel attrition. It has already been noted that those with high victimization are relatively likely to move from the household location and therefore be lost to the panel. There are at least three different reasons for this. First, high victimization and household propensity to move may independently be related to other causal variables, such as low income. Second, it seems plausible that some households may move because of high victimization. Presumably in such cases the household has decided there is location proneness. Third, the young are relatively likely to leave the household, going off to college or setting up their own households. Since the young have relatively high victimization rates, this process will reduce the victimization rates for the household considered as a collection of individuals. Either starting point for the Fienberg investigation will, provided it ultimately deals with complete long records, lead to an examination of a limited subset of the original panels, and the victimization experience of the movers is likely to differ systematically from the experience of those who remain. Furthermore, if, as is sometimes the case, it is impossible to tell from the data records whether the occupants of a household have changed from interview to interview, some of the "complete" long records will actually be spliced records of more than one household (in the nonlocational sense). Given the likelihood that the victimization experiences of the various households comprising the spliced record will systematically differ, these "completed" records will provide misleading information.

While the presentations and discussion at the conference focused on questions of multiple victimization, they pointed to several important problems in the design and analysis of victimization surveys which have bearing well beyond questions of multiple victimization. The overriding problem is that longitudinal analysis is required, while the attention of LEAA and the Bureau of the Census has been directed almost entirely to cross-sectional analysis. Related to this are several problems resulting from selective attrition of panels, largely caused by residential mobility and changes in family composition. Alteration in survey design is required to (1) provide information about the continuity of respondents within household locations, and (2) allow investigation of patterns of movers, pre- and post-move.

Conferees commented upon the Census Bureau's "Description of Activities," a proposed research agenda for fiscal years 1978 through 1982. This Description treats the matters mentioned in the previous paragraph, and it is worth quoting the entire treatment:

> *3. Questionnaire design research*
> Research on the most effective questionnaire design will focus on the following areas:
>
> ...
>
> d. A method of identifying respondents so they can be traced across time.
>
> ...
>
> *12. Conceptual issues regarding NCS*
>
> ...
>
> b. Mover/Nonmover Study (including a comparison of in-migrants with out-migrants).
>
> *13. Analytical issues*
>
> ...
>
> a. Longitudinal studies. An exploration of techniques and difficulties of exploiting NCS crime features to do longitudinal analyses.

It is noteworthy that mention of the longitudinal features of the design and their importance for analysis does not appear until the final page of the document. It was the clear sense of the conferees both that the placement of that brief discussion at the end indicated the priority attached to longitudinal analysis by those responsible for NCS, and that such low priority was totally inappropriate.

Decisionmaking

Considerable attention at the workshop had been given to aspects of decisionmaking in parole and sentencing. Several of the workshop participants continued research in this area and presented results to the conference.

The Salient Factor Score system used by the Federal Board of Parole approaches the parole prediction problem by assigning scores for each individual on nine categorical variables and adding the scores (equally weighted) to produce a single Salient Factor Score, which is used to predict parole outcome. Herbert Solomon (1976) reanalyzed the data earlier used to develop the Salient Factor Score system. He showed that, for a subset of four predictor variables, logit models, which weight the variables unequally, can achieve better prediction than the equal weighting of the Salient Factor Score system, in the sense of fitting the observed proportion of successes. (In part, no doubt, this better prediction occurs because Solomon's method produces a larger number of prediction categories than the Salient Factor system, given the same number of predictor variables.) Kinley Larntz felt that analysis could be pushed further (or at least beyond the published description of Solomon's work) and that it would be useful to approach a different data base with the same general questions. He described his work on this problem at the conference.

The State of Minnesota had undertaken a parole project similar to the federal effort, developing its own salient factor score system on the basis of an extensive body of data concerning released prisoners. Larntz obtained data for about 1000 released prisoners from the State of Minnesota, the data used in the development of the salient factor system, and reanalyzed it using multivariate log-linear methods. (Tables and graphs from Larntz's analysis are in his "Linear Logistic Models for the Parole Decisionmaking Problem.")

Larntz begins with a logit analysis using the eight categorical variables of the salient factor scores applied to 485 cases selected from the full file (this reduced sample is the "construction sample"). This analysis parallels that of Solomon, except for the larger number of variables. In addition, Larntz tested for two-factor interactions (which Solomon does not report having done). Only an age at first conviction-burglary interaction seemed to have much effect, with those first convicted at higher ages more likely to be parole failures if the instant offense is burglary, while whether the instant offense is burglary matters little for those first convicted at age 19 or younger. Larntz estimated models using the full eight variables, the eight variables with the one interaction, a subset of four variables (equivalent to the Solomon model), and four variables with the interaction, computing three kinds of chi-squared statistics for each. In addition, he computed the likelihood ratio chi-squared for an equally weighted model ("Burgess scale," the salient factor method). While models which include the interaction term do better than models which do not, differences in goodness of fit among the various models without the interaction term are exceedingly small. That is, it does not seem to matter much whether one uses four or eight variables (presumably because of high collinearity among the variables). Nor does it seem to matter much whether one uses the equal weights of the salient factor method or a more sophisticated weighting scheme.

Larntz then removed the restriction of the analysis to categorical variables. Several of the categorical variables of the first analysis were produced by categorizing measured scales, and Larntz therefore restored the original scales. Removing the restriction called for changing the analysis technique to logistic regression, and Larntz applied this technique to seven variables which are quite similar to the variables used in the earlier analysis. See Larntz, for partial results. Once again, the number of variables included in the model did not have a striking effect on goodness of fit. Indeed, a model containing only two variables, age at first conviction and number of previous parole failures, performed nearly as well as a seven-variable model. Larntz tested interactions, but did not include any in the reported models because they seemed to have little effect. It is striking that even the seven-variable model does not represent an enormous increase in goodness of fit over a model using no variables except a constant term. It is difficult to compare the logistic regressions with the logit models in terms of goodness of fit, since the variables used in the logit models were selected by the Minnesota parole project precisely because they produced a salient factor score which seemed to "work" well for the data Larntz was using. The variables used in the logistic regressions were somewhat less preselected.

The point of parole prediction models, of course, is not to fit observed probabilities of parole failure for categories of potential parolees. It is rather to provide predictions of success or failure which may serve as the basis for parole decisions. The important question, therefore, is how well the model is able to separate parole successes from parole failures. Larntz turned to the other half of the data set, the validation sample, for an approach to this question. The logic is roughly as follows. For each individual in the sample, the model will generate a prediction of success. The analyst then chooses a cutpoint on the probability scale. The group of individuals with predicted probabilities above the cutpoint will presumably include some parole failures and some parole successes. The group below the cutpoint will also include some successes and some failures. The difference between the proportion of all successes who fall above the cutpoint and the proportion of all failures who fall above the cutpoint is a measure of the ability of the model to discriminate properly. As the cutpoint is varied, this difference is likely to change, and the proportion of the whole group falling above the cutpoint will surely change. For each of the models Larntz varied the cutoff point and examined the key difference. These results are presented as a series of graphs. To facilitate comparison across models, Larntz graphed the differences not as functions of the cutoff point, but rather of the proportion of the total sample below the cutoff point.

It is to be expected that the logistic models will perform relatively well in this test, because the variables had initially been selected on the basis of their performance in the combined validation and construction samples. Performance of the logistic regressions near this level would, because less preselection of variables was involved, indicate that they are of a relatively high quality. The high peak associated with a logistic regression based on only two variables is therefore particularly striking.

In principle, analysis of this kind could be used to select an optimal cutpoint value. However, that selection depends upon criteria for optimality, and it is not entirely clear what those criteria should be. Maximum separation is obviously an attractive candidate. However, one might want to, say, sacrifice some degree of separation for a larger proportion of the total sample below the cutpoint.

Larntz next asked whether the effects of preselection of variables distorted the results, so that the estimated probabilities were in fact not very good estimates. He tested this possibility by taking the estimated probabilities from each model and using them to generate a binomial random variable (success or failure) for each individual. He then repeated the cutpoint and difference analysis for each model, using the generated outcome scores rather than the actual outcomes. Graphs for three models are presented in Larntz. The first two graphs are for logistic models. It is clear that in both cases the curve has shifted upwards, which suggests that the distribution of estimated probabilities is too spread out. Application of these models to a new data set should reveal considerable shrinkage. There is noticeably less upward shifting for the one logistic regression model included, indicating that the estimated probabilities are reasonably good. And it should be emphasized that the logistic regression involved here has only two predictor variables. The conclusion is that a very simple model, involving only two variables, performs quite satisfactorily—at least in comparison with the other models tested. It is obvious, however, that there remains considerable error in prediction.

Howard Wainer has also worked on the parole prediction problem, using data from the Federal parole system. Since his attempts to improve upon the predictive performance of the salient factor system by use of alternative estimation procedures had earlier been considered by most of the conferees, they were not discussed extensively at the conference. But it is worth mentioning that, by using psychometric methods applied to the variables used in the Federal salient factor scores, Wainer was able to improve on the salient factor score predictions. However, as with Larntz's various models, the improvement in predictive power was relatively small, and the resulting predictions are not very accurate. As Wainer and Perline have written, "our results are at least as good with this approach as with any other, but still leave a great deal to be desired."

The conferees were in general agreement as to why none of the approaches leads to very good prediction of parole success or failure. Leaving aside such common problems of prediction as, say, the influence of chance events on behavior, there is the problem that the parole prediction problem is considered only in the context of a highly homogeneous population representing a very small and extreme portion of the distribution of behavioral characteristics. The prison population is itself extremely restricted in its range of variation, and presumably an extreme portion of the prison population is seen as "obviously" unsuitable for parole, so the parole data base is even more restricted. As Leslie Wilkins put it, parole prediction is an attempt to predict far out in the tail of a distribution, which is a difficult thing to do. It is therefore not surprising that it appears difficult to improve upon Larntz's two-variable logistic regression model.

Wainer in fact told the conferees that he had given up hope of accurately predicting recidivism (or parole failure), and in his work on what appeared to be the parole prediction problem was actually addressing a different question. From the perspective of his analytic approach, events (or the variables involved in parole prediction) are viewed much like items in a test, and each event is characterized by its "rarity" ("difficulty" in the testing context). Recidivism, or parole failure, is viewed simply as an event with a certain amount of rarity. Individuals in the testing context are characterized by their level of ability, which, in the parole context, Wainer refers to as "probity." The statistical methods yield estimates both of the rarity of events and of the probity of individuals.

Wainer wants to be able to describe institutions, say in the criminal justice system, by the characteristics of the individuals entering them. Then, if one saw that recidivism among the products of some institution were declining, one could ask whether the decline resulted from a change in the nature of the institutional population or from something else, such as a new program which had been instituted. In other words, one would want to examine recidivism of the institutional population controlling for the probity of the population. The focus, then, is shifted from prediction of individual behavior to system characteristics.

Having introduced this perspective to the conferees, Wainer presented a problem for their consideration. In estimating rarity and probity, he has found occasional item-person interactions, or instances in which particular item-person combinations are badly fit by the model. The analogy in the testing context is to those occasions on which an individual with low ability nevertheless gets a difficult item right, or an individual with a high ability misses an easy item. Because of the probabilistic nature of the event-person combinations, some "errors" of this kind are bound to happen. However, they seem to happen more often than the probability model would suggest. Furthermore, even if they are expected to happen, they cause difficulty for the estimation of individual probity, particularly if the item is far from the individual's probity, because information about individual probity decreases with distance of the item from the individual probity while its influence on the estimate of probity increases. The problem, then, is how to reduce or eliminate the effects of these errors in estimating probity.

The statisticians at the conference discussed this estimation problem at some length, providing Wainer with some useful suggestions, but not reaching a clear solution.

While predictions of success are important in modern methods of parole decisionmaking, they are not the sole basis for the decision. The Federal probation guidelines combine salient factor scores (i.e., predictions) with offense characteristics ("seriousness") to produce "average total time served before release." Wainer applied two-way table decomposition methods to the matrix showing the midpoints of recommended sentence intervals for each combination of salient factor score and offense characteristic category in the salient factor system. His results indicate that "the existing parole policy corresponds closely to a multiplicative model in which seriousness of offense seems to count almost three times as much as offender characteristics among adults" (Perline and Wainer). The relatively small effect of the salient factor scores (that is, the measure of offender characteristics or of probability of recidivism) suggests either that the inaccuracy of the predictions of recidivism is taken into account in decisonmaking, or that characteristics of the offense are considered more important in determining time served. Wainer suggests that the two dimensions of the guideline scheme, salient factor scores and offense characteristics, correspond to two schools of thought on incarceration, rehabilitation, and retribution. It could be argued, then, that his finding concerning the relative importance of the two factors indicates the relative importance of the two views of the purpose of incarceration in parole decisionmaking.

Before parole boards can consider the decision to release a prisoner, there must be a decision to incarcerate the individual in the first place. David Britt and Kinley Larntz have investigated judicial sentencing decisions, using a sample of 200 cases from the Denver, Colorado, courts, originally collected under the auspices of the State

Courts Sentencing Project and provided by Leslie Wilkins. After elimination of cases because of inadequate data or because they involved "victimless crimes," 138 cases were analyzed to determine the major factors used by judges in reaching the decision of whether or not to incarcerate an individual convicted of a criminal offense. Logit models were used in the analysis.

In brief summary, the results indicate that dispositions of crimes against property appear related to variation in the seriousness of the crime and the offender's previous record in a straightforward way. "As the crime becomes more serious and the offender's previous record becomes poorer, the chances of the individual's being incarcerated increase." (Britt and Larntz). Furthermore, seriousness of offense does not seem to lead to incarceration for those with a good prior record. For crimes against persons, seriousness of offense did not play a large role in incarceration, presumably because nearly all the offenses involved were relatively serious ones. A two-variable model using the nature of the plea and whether the charge was reduced best fit the data. No characteristics of the offender affected the chances of being jailed once these two variables were taken into account.

The contrast between sentencing decisions, at least as they appear to be made in Denver, and parole decisions under modern guideline systems is worth noting. The parole decision system explicitly builds in a variety of characteristics of the individual in determining release. The sentencing decisions incorporate no individual characteristics (but for prior record in the case of property offenses) in determining when to incarcerate.

Several qualifications should be introduced. First, the sentencing study did not investigate length of sentence. Second, individual characteristics may be incorporated into prior events in the criminal justice system: the Denver data concern only convicted individuals. Individual characteristics may play a far larger role in the decisions to arrest, prosecute, and convict. A full investigation of the various factors determining outcomes in the criminal justice process would have to begin at a much earlier stage than the sentencing stage. It seems likely that the development of OBTS data will make such investigations easier to conduct in the future.

Deterrence

The problem of estimating the deterrent effect of criminal sanctions occupied considerable time at the workshop and considerable effort by workshop participants since then. The major work on deterrence since the workshop has been that of the National Research Council's Panel on Research on Deterrent and Incapacitative Effects (1978). The panel was chaired by workshop participant Alfred Blumstein and included two other workshop participants, Gary Koch and Albert J. Reiss, Jr. Daniel Nagin, staff to the panel, spent several weeks at the workshop, and Brian Forst, co-author of one of the panel's commissioned papers, reported on his research to the workshop. It is appropriate, therefore, that the conference discussion of deterrence began with a brief statement by Blumstein summarizing the panel's work.

Blumstein noted that the panel had focused on general deterrence, where the issue is the association between crime rates and sanctions. Virtually all the research on this issue has found, after controlling for other determinants of crime, a negative association between crime and sanctions. Whether there is a causal relationship is a more difficult question. The panel considered three principal aspects of this question:

1. Single equation models typically use, on the left-hand side of the equation, a fraction in which crime is the numerator, while the sanction risk on the right-hand side of the equation is a fraction with crime in the denominator. If there is error in the measurement of crime, it will generate a negative association between crime rate and sanction risk, even in the absence of a deterrent effect. Without good estimates of the magnitude of the error, therefore, it is difficult to know what to make of the association found in the data.

2. Imprisonment is typically the sanction used in deterrence studies. However, while the risk of imprisonment may have a deterrent effect, the fact of imprisonment may affect crime rates through incapacitation of active criminals. An estimated effect of imprisonment, therefore, may combine both deterrent and incapacitative effects in unknown proportions.

3. A negative association between crime and sanction may result from deterrence of crime by sanction, but it may also result from the inhibition of sanction by rising crime rates. This is particularly a problem in studies which, like most published studies of deterrence, rely on cross-sectional data at the state level. It is also a particular problem to the extent that studies rely on the sanction of imprisonment, because it is reasonable to suppose that imprisonment decisions are influenced by such factors as prison crowding. If imprisonment rates and crime rates are simultaneously determined, as this suggests, then simultaneous equations models may become appropriate. Estimation of simultaneous equations models requires appropriate identification restrictions, that is, variables which influence sanctions but not crimes, or variables which influence crime but not sanction risk. The choice of identifying restrictions is a difficult problem, which, in the panel's view, has not been adequately resolved.

In sum, a deterrent effect is not clearly demonstrated in existing studies. This brief presentation sparked considerable discussion among the conferees. A number of them argued that treatment of the simultaneity problem involves far more than choosing appropriate identifying restrictions and separating deterrence and incapacitation. The argument is that the relationship between imprisonment and crime rates is more complicated than the aggregate statistical models suggest. The ways in which adaptation to the pressures of changing crime rates affects prisons over time, and the related question of how changes in prisons affect crime rates, have not been studied to any substantial extent. The composition of prison populations may shift, the social environment of the prisons may change (arguably making prisons more—or less—attractive places to be), the symbolic meaning of imprisonment may change as prisons become politicized, and so forth. Further, it is reasonable to expect that the effects of imprisonment on crime rates depend upon prison release rates and the character of the prisoners released. At the very least, disentangling these factors requires disaggregation of the available data by offense—a recommendation with which the panel agrees and which is followed in a number of the studies. Whether modifications to account for the other problems and questions can be built into the kind of aggregate analysis common in the study of deterrence is another matter. In any event, the requisite data are not available to do so now even if in principle it can be done. The question of whether valid conclusions about the deterrent effect of sanctions can be reached through the use of aggregate models which do not include these complex problems was not resolved in conference discussion.

Discussion then turned to the problem of estimating deterrent effects in single equation models. Blumstein had noted that if crime is measured with error in the standard models for the problem, estimates of deterrent effects are biased. The standard model would take the following form:

$$C/N = A(I/C) + \ldots + e,$$

Where C = number of crimes
N = population
I = number of sanctions
e = random error term
A = parameter to be estimated

Estimates of A are biased if C is measured with error. Blumstein added, however, that if C is measured without error, then unbiased estimates of A are possible. The statement generated considerable controversy, less because conferees believed it to be false (it is not clear that any of them felt it was false, when properly qualified) than because its implications for the study of deterrence are problematic. In other words, presuming an unbiased estimate of A, is that estimate a measure of the deterrent effect of criminal sanctions? The question can be approached through two hypothetical experiments, the first proposed by Blumstein and the second by other conferees.

Blumstein's experiment is roughly this. Pick a value for A. Generate data for I and C, without error. (These may be real or hypothetical data.) Generate a random variable e. Then use the equation to generate values of C/N. Label I/C "imprisonment rate" and C/N (as generated by the model) "crime rate." Give these variables to a statistician and ask him to estimate A. The statistician should be able to produce an unbiased estimate of (the known value of) A. The structure of the model suggests that A is a measure of the effect of the imprisonment rate on the crime rate.

The second experiment uses a random number generator to produce values for C, I, and N. Then the two variables C/N and I/C are formed. The correlation between them is calculated. Given the construction of the variables, there should be a negative correlation, even though number of crimes, number of incarcerations, and population are random numbers. In other words, Blumstein's statistician, given "crime rate" and "incarceration rate" and asked to estimate (the unknown value of) A, will produce an estimated negative value. Since the data are purely random, it is difficult to see why A should be considered a measure of the deterrent effect of incarceration.

In objecting to Blumstein's hypothetical experiment, conferees noted that by fixing values of C and using them in the model to generate values of C/N, the experimenter would in effect be generating simply values of the reciprocal of N. This suggested to some of the conferees that the problem was incorrectly formulated. As one conferee said, "You cannot fix C and then talk about generating crime rates."

Another approach to the question formulates the basic equation in terms not of the quantities actually used in the standard approaches, but rather in conceptual terms. Thus "crime rate," or something like it, is seen as a function of "sanction risk," among other things. Call the coefficient of "sanction risk" B. Then the standard equation is seen as a particular realization of the conceptual model, using proxies for the conceptual variables. Then there are two separate questions which can be asked of the estimation. First, is the estimate of A unbiased? That question, which has apparently preoccupied the literature, appears to have the answer Blumstein suggested: if C is measured without error, an unbiased estimate of A is possible. The second question, however, concerns the relationship between A and B. And it would appear that the particular structure of the variables used in the estimation—that is, the presence in the denominator of the sanction risk variable of the numerator of the crime rate variable—means that A is not an unbiased estimate of B.

If the answer to the second question suggested above is correct, it immediately raises another problem. How does one estimate B? Colin Loftin, drawing upon his own deterrence research, suggests using sanctions rather than sanction risk as the explanatory variable. He noted that the deterrence research has generally found that measures of the severity of punishment have less of an effect on crime rates than do measures of the certainty of punishment. Severity measures, unlike certainty ("risk") measures, do not share a common element with crime rate measures. Loftin has used number of executions as a sanction variable in his capital punishment research. This removes the common term, but it leaves somewhat open the question of the underlying behavioral model.

A strong theme of much of the discussion of this problem was that the specification of deterrence models common in the literature may not allow the determination of deterrence effects. Considerable complexity must be added to the models. The "crime rate" needs to be decomposed. It might be viewed as a function of a distribution of individual rates of offending. As Reiss has written, "there are three general kinds of change that affect the incidence of crime in a population. They are changes in the *prevalence* of offenders in a population, changes in individual *incidence* of offending, and *interactions* between prevalence and incidence rates" ("Understanding Changes in Crime Rates"). In effect, the left-hand side of the deterrence equation must consider all of these. Both prevalence of offenders and individual rates of offending may be influenced by the criminal justice system in a number of ways, and these would need to be considered on the right-hand side. In addition to the general deterrent effect of imprisonment, one should consider the incapacitation effect, the effects of differential release rates, and other complications. In the opinion of some of the conferees, unless decompositions of this sort are essayed, it will be impossible to disentangle the various effects of sanctions and therefore impossible to estimate a deterrent effect. Analysis along these lines, however, is not a short-term project. The data simply do not now exist to implement it.

There remains the possibility that, short of implementing the more complex model described above, one can consider estimates derived from the standard deterrence model as measures of the aggregate effect of sanction risk, without worrying too much about what proportion, if any, of the aggregate effect is actually deterrence. Taking that approach raises the question of whether the findings of deterrence research can be accepted, subject to this rather substantial limitation.

Stephen Fienberg has taken a close look at some of the deterrence studies, and the results he reported to the conference suggest that estimates of the deterrent effects of sanctions should be treated with some caution. (Fienberg subsequently supplied a written version of his report, Brier and Fienberg, "Recent Econometric Modelling of Crime and Punishment: Support for the Deterrence Hypothesis?".)

Fienberg first examined Isaac Ehrlich's (1973) well-known analysis of 1960 data from 47 states and the reanalysis of these data done by Vandaele (1978) for the National Academy panel. Fienberg first notes the well-known shortcomings of Uniform Crime Reports data as measures of the variables required by Ehrlich's supply of offenses function and suggests that it would be appropriate to use statistical methods designed for the "errors in variables" problem and for multiple indicators of unobservable variables. Ehrlich's 1960 data were unavailable and published reports do not fully specify the simultaneous equations model actually used, so it was not possible to replicate the Ehrlich analysis. Vandaele had reanalyzed the Ehrlich data and concluded that inferences about deterrence are not sensitive to changes in the specification of the model. Using the data reported by Vandaele, Fienberg was able to duplicate the results. However, he questions the conclusion that inferences are not sensitive to changes in specification, noting that Vandaele's published results for murder, rape, and assault do seem sensitive to the choice of specification. This is no small problem, because the correct specification cannot be determined by analysis of the data themselves.

One of the key variables in the analysis is probability of incarceration, measured by the ratio of the number of commitments to the number of offenses. Vandaele had noted that for several states this ratio is greater than one, which, of course, is the upper limit for probabilities. This is not necessarily an indication of error in the data, because there are a number of plausible and reasonable explanations for ratios greater than one. It does, however, suggest that the ratio may not be a reasonable measure of the perceived probability of incarceration. Vandaele reestimated the equations deleting states with ratios greater than one and found little change in the results. Fienberg argues that to treat the data in this way is to assume that, while ratios greater than one may be bad data, there is nothing wrong with ratios less than one. However, Fienberg continues, the fact that the underlying process produces probability estimates greater than one suggests that there is something very wrong in the process as a whole, and therefore that there is no more reason to accept ratios less than one than to accept those greater than one. Using an analogy, he suggests that if a computer program computes 100 correlations, and 5 of them have values greater than 1.0, wisdom may dictate throwing out all 100 values, and not just the 5 impermissible ones. The impermissible values serve as evidence that none of the values are to be trusted. Fienberg's logic strongly suggests the conclusion that no deterrence studies relying upon such data can be accepted as providing reliable evidence of a deterrent effect.

Brian Forst (1976) attempted to replicate Ehrlich's study using 1970 data. There are differences in some of the variables used by Forst and Ehrlich, and Forst did not follow Ehrlich's specifications. In particular, Forst used more sociodemographic variables than Ehrlich. Further, Forst analyzed only the aggregate crime rate, while Ehrlich used rates for individual crimes. Forst's results differ strikingly from those of Ehrlich, as he found the coefficients of the deterrence variables insignificant. It is, of course, possible that these differences are to be explained by a change in patterns of criminal behavior over a 10-year period, but it is not clear what would explain such a change. The extra sociodemographic variables may also explain the difference.

Fienberg obtained most of the Forst data and reanalyzed it. He was able to duplicate Forst's results quite closely. However, one key variable, a measure of income dispersion, could not be obtained from Forst, and Fienberg therefore used several different measures of income dispersion. He discovered that the choice of a measure had substantial effects on most of the estimated coefficients. That is, the model is extremely sensitive to minor differences in the definition of a single variable. Since there are no strong grounds for preferring one definition to another, this suggests extreme caution in accepting the results of the analysis of deterrence effects.

One of the most widely discussed deterrence studies is Ehrlich's (1975) longitudinal study of the deterrent effect of capital punishment. Ehrlich's data could not be obtained, and Fienberg therefore reanalyzed the closely related data of Bowers and Pierce. While Fienberg discusses a number of methodological problems which raise questions about Ehrlich's specifications and conclusions, only two aspects of the reanalysis will be discussed here.

The Ehrlich data cover the years 1933-1969, and, because of the use of lags, analysis is limited to 1934-1969. Other analysts have noted that if the years 1963-1969 are deleted, the coefficient of the probability of execution variable loses statistical significance. To this Fienberg adds two points. First, by redesigning aspects of the analysis, so that the model was recursive, not simultaneous, he was able to analyze residuals, and these analyses show that the observations for 1963-1969 are discrepant. This reinforces the argument of other analysts that the relationship among the variables has changed over time, so that it is misleading to compute the deterrent effect based on the entire series. Second, analysis of residuals points to 1934 as an outlier in the data. Deleting that one observation has a substantial impact on the estimated coefficients. In the original scale of the variables, the coefficient of the execution variable is negative but not statistically significant. In a logarithmic specification, the estimated coefficient is positive. The data for 1934 show a small value for the murder rate and a high value for the execution variable (number of executions for murder in year $t+1$ divided by the number of convictions in year t). Fienberg suggests that the 1934 data are "suspect." Whatever the explanation, the fact that a single observation has such a substantial effect on the estimates reduces confidence in the findings generally.

Ehrlich's specification of the "murder supply" function is linear in logarithms, with one major exception. Time, used as a surrogate for improvements in medical technology, enters the transformed equation (that is, the equation after converting to linearity in the logarithms) untransformed. The reason for treating time differently from all the other variables is not clear. Fienberg reestimated the murder supply function using a fully logarithmic specification (that is, using logT instead of T). The effect was to change the sign of the coefficient of the execution variable. Fienberg then tried various other transformations of time and discovered that by this means he could change the results of the estimation almost at will. This might not be a great problem if there were some logical defense for the choice of any particular transformation, but none is apparent. Stated in very strong form, the appropriate conclusion would appear to be that longitudinal analysis of the Ehrlich capital punishment data can produce virtually any conclusion about the deterrent effect of capital punishment, depending upon the arbitrary choice of a scale for time.

The final paragraph of the Brier-Fienberg paper baldly presents their overall conclusions: "We can find no reliable empirical support in the existing literature either for or against the deterrence hypothesis. Moreover, we believe that little will come from further attempts to model the effects of punishment on crime using the type of data we have described in this paper." This reinforces the lessons suggested above: a reformulation of the problem must precede any advances in the study of deterrence.

References

Ehrlich, Isaac (1975)
"The deterrent effect of capital punishment: A question of life and death." *American Economic Review* 65:397-417.

Ehrlich, Isaac (1973)
"Participation in illegitimate activities: A theoretical and empirical investigation." *Journal of Political Economy* 81:521-565.

Forst, Brian E. (1976)
"Participation in illegitimate activities: Further empirical findings." *Policy Analysis* 2:477-492.

National Research Council (1978)
Deterrence and incapacitation. (Panel on Research on Deterrent and Incapacitative Effects.) Edited by Alfred Blumstein, Jacqueline Cohen, and Daniel Nagin. Washington, D.C.: National Academy of Sciences.

Solomon, Herbert (1976)
"Parole outcome." *Journal of Research in Crime and Delinquency*, July, 107-126.

Vandaele, Walter (1978)
"Participation in illegitimate activities: Ehrlich revisited." In National Research Council, *Deterrence and incapacitation*, Washington, D.C.: National Academy of Sciences.

Conference participants

Alfred Blumstein
Urban Systems Institute
Carnegie-Mellon University

David W. Britt
Criminal Justice Program
Nova University

Stephen E. Fienberg
Department of Statistics
and Department of Social Science
Carnegie-Mellon University

Kinley Larntz
Department of Applied Statistics
University of Minnesota

Colin Loftin
Department of Sociology and
Center for Research on Social Organization
University of Michigan

Albert J. Reiss, Jr.
Department of Sociology
Yale University

Richard F. Sparks
School of Criminal Justice
Rutgers University

Howard Wainer
Bureau of Social Science Research
Washington, D.C.

Leslie T. Wilkins
School of Criminal Justice
SUNY at Albany

Guests

Charles R. Kindermann
Bureau of Justice Statistics
U.S. Department of Justice

Paul D. White
Bureau of Justice Statistics
U.S. Department of Justice

Staff

David Seidman
Social Science Research Council
Center for Coordination of Research on
Social Indicators

Conference papers and other materials

Brier, Stephen S. and Stephen E. Fienberg
Recent econometric modelling of crime and punishment: Support for the deterrence hypothesis?

Fienberg, Stephen E.
Victimization and the National Crime Survey: Problems of design and analysis

Larntz, Kinley and David W. Britt
The effects of plea bargaining on the disposition of person and property crimes: A research note

Larntz, Kinley
Linear logistic models for the parole decision-making problem

Loftin, Colin
Alternative estimates of the impact of certainty and severity of punishment on levels of homicide in American states

Reiss, Albert J., Jr.
Understanding changes in crime rates

Reiss, Albert J., Jr.
Victim proneness by type of crime victimization over time

Sparks, Richard F.
Criminal opportunities and crime rates

Wainer, Howard and Richard Perline
Quantitative approaches to the study of parole

Crime rates and victimization

Understanding changes in crime rates

ALBERT J. REISS, JR.
Department of Sociology
Yale University

Much of the empirical research in criminology is based on explaining changes in crime rates and much of the research evaluating law enforcement or criminal justice programs depends upon measuring changes in crime rates. The choice of measures of crime is much debated and their validity or reliability critically reported. Yet the relationship among measures of crime, e.g., between crime incident rates and aggregate rates of individual offending, is not well articulated. This stems in part from a lack of understanding of the nature of crime events, of the criteria for selecting measures based on such events, and of the sources of change in measures of crime. This paper attempts to clarify some of these problems in the measurement of crime using examples primarily from research on deterrence and incapacitation.

Understanding crime events

Any occurrence of crime is a complex event. Scanting most elements in the occurrence of crime events, the three critical ones for measuring crime are that an event may comprise one or more criminal acts, one or more offenders, and one or more victims. Of less critical significance are variation in duration of events in time and the number of settings involved in the criminal activity.

Now it is commonly understood that when a crime event is counted as a single crime act or incident, the number of victims and of offenders on the average exceed the number of crime incidents since some events involve more than one victim and/or more than one offender. But what often is overlooked is that when there is more than one offender for any incident, the offenders account for only a single crime event (even though the police understand that multiple arrests for a single incident clear only that incident, barring admissions by the offenders to other crime events).

Likewise, where there is more than one offender and more than one criminal act in an event, different offenders may commit different acts in that event. While these differences may be reflected in different charges or indictments for offenders involved in a common event, they will not be reflected in measures that assign each offender the *same* crime incident classification.

We shall not pursue here other implications of multiple offenses, offenders, and victims in crime events since we shall have occasion to draw the reader's attention to some of them in presenting a simple model for understanding changes in crime rates.

Understanding changes in crime incident rates

Whether a crime event is classified as a single incident or as multiple incidents, there are three kinds of change that affect the incidence of crime in a population. They are changes in the *prevalence* of offenders in a population, in individual *incidence* of offending, and in *interactions* between prevalence and incidence rates. Assuming no changes in individual incidence of offending (or interactions), for example, an aggregate crime incident rate may change solely due to changes in the prevalence of offenders in a population. Correlatively, if the prevalence rate remains constant, changes in the individual incidence of offending will affect the aggregate crime incidence rate. Because the prevalence rate and the individual incidence of offending may vary independently, their rates of change may be negatively as well as positively correlated.

Changes in the prevalence of offenders

Any change in the prevalence of offenders in a population comes about as a consequence of changes in replacement rates for a population of offenders. Changes in entry or exit from a population of offenders or of age at onset and desistance from a population affect replacement rates. We may identify several sources of change in replacement rates.

First, it has long been understood that changes in the *size of birth cohorts* or of *cohorts at risk in the population* affect the prevalence of offenders in a population if the probabilities of age at onset and desistance and individual rates of offending remain constant. Even when the prevalence rate for a cohort remains constant, any increase or decrease in the size of birth cohorts affects the prevalence of offenders in a total population after the cohort reaches the age of onset of offending.

Second, while it is understood that changes in length of offending career may have an effect on the crime rate, it is not generally appreciated that both *changes in age at onset and at desistance can affect the replacement rate of offenders* in a population and, therefore, the prevalence of offenders in a population.

Third, *changes in the number, size, degree of openness, composition, and territory of social networks may affect the prevalence of offenders in a population*, since these changes affect the potential for recruitment of offenders to offending groups. Both changes in market networks for the goods and services of crime as well as changes in formal and informal social networks may affect recruitment.

These changes are not necessarily independent one of another. Changes in social networks, for example, can affect changes in the age at onset or desistance from offending. It also is possible that beyond a certain point changes in the prevalence rate may increase both the number and size of social networks, a phenomenon sometimes observed in gang delinquency and organized crime. The relationship of processes of recruitment to the number and size of social networks is not well understood.

Changes in individual offending rates

Any change in individual incidence of offending likewise affects the crime incidence rate. We may think of three major types of change that affect individual rates of offending in a population.

First, *changes in opportunities to commit crime events can change individual rates of offending.* Assuming an offender, his or her probability of offending is some function of both opportunities and intervening opportunities to commit criminal acts. The opportunity to commit motor vehicle theft, for example, may increase both as the ratio of motor vehicles to the population increases and as the ease of a successful attempt increases.

Second, *changes in the size or structure of an offending group or network can alter the group and, therefore, an individual member's rate of offending.* There is plausible evidence that on the average an individual's rate of offending is greater as a member of an offending group than as a lone offender. Any reduction in the ratio of individual to group offending, therefore, should increase, on the average, individual rates of offending.

Third, *changes in the size and composition of crime networks should affect an individual's rate of offending.* Substantial changes in markets for illegal goods or services, for example, can affect the individual offender as an agent of supply, assuming elasticity of victims in the population.

Implications for research on deterrence and incapacitation

Little is known about the distribution of individual rates of offending or changes in individual rates of offending during an offender's career. In particular, almost nothing is known about the mix of individual and group offenses in an individual's rate of offending. The National Crime Survey of crime victims provides information on the number and selected characteristics of offenders in crime incidents reported by victims. For crimes against persons, the number of offenders is usually reported; for most crimes against property, offenders are not known. Table 1 provides information on the size of offending groups for selected types of crime incidents while Table 2 has information on the distribution of offenders by size of offending group for the same types of crime.

Although there is considerable variation in the individual or group contribution to offending in crime incidents reported by victims in Table 1, almost two of every three reported incidents involved only a single offender. For crimes against the person the proportion of incidents committed by a single offender ranged from only 30 percent of all serious assaults with a weapon followed by theft to 82 percent of all attempted rapes with no theft from the victim.

Unfortunately, we do not know how many *different* offenders are involved in these crime incidents reported during a 6-month period. Nonetheless, when we view these crime incidents in Table 2 from the perspective of an offending population, we see that only 30 percent of all offenders in these crime incidents were lone offenders. Again, there is considerable variation with only 10 percent of all offenders who commit a serious assault with a weapon followed by theft being lone offenders but 58 percent of all offenders who commit a rape with no theft being lone offenders. There is, on the average, 2.1 offenders for each crime incident against a person. There is considerable variation in this average with 2.9 offenders on the average for incidents of serious assault with a weapon followed by theft to 1.4 offenders for a rape with no theft.

Table 1.

Type of crime	Total incidents offenders known
Rape w/theft	30
Rape, no theft	112
Att. rape, w/theft	39
Att. rape, no theft	316
Robbery, w/weapon	752
Robbery, no weapon	596
Att. robbery, w/weapon	432
Att. robbery, no weapon	584
Ser. assault, theft, weapon	519
Ser. assault, theft, no weapon	71
Ser. assault, weapon, no theft	1,339
Ser. assault, no weapon, no theft	244
Minor assault, theft	485
Minor assault, no theft	1,947
Att. assault, w/weapon, no theft	3,289
Att. assault, no weapon, no theft	5,940
Purse snatch, no force	249
Att. purse snatch, no force	202
Pocket picking	441
Burglary, forced entry, something taken	95
Burglary, forced entry, 0 taken, damage	126
Burglary, forced entry, 0 taken, 0 damage	44
Burglary, unlawful entry, 0 force	588
Burglary, att. forced entry	374
Larceny, $250+	120
Larceny, $100-$249	208
Larceny, $50-$99	279
Larceny, $25-$49	302
Larceny, $10-$24	454
Larceny, under $10	749
Larceny, $ unknown	103
Att. larceny	842
Car theft	74
Att. car theft	120
Theft, other vehicle	14
Att. theft, other vehicle	26
Total	22,107
Percent	100.0

Given the limitations in these data that we do not know how many different offenders are involved in the reported incidents (we would assume that minor assaults without theft would involve fewer different offenders than rape, for example) nor the mix of individual and group offenses in an individual's rate of offending, we cannot estimate with any precision what on the average any offender contributes to crime incidents. We can say, however, that if an offender were involved in a group incident, it is not certain whether removing that offender from that population of offenders would have averted that incident. We can also say that where offenders are largely involved in group incidents, unless there is a deterrent effect on co-offenders in the event, it is unlikely that any crimes are averted by removing any single member of the offender group. We may also infer that unless offenses committed by lone offenders are committed in substantial proportion by high-rate lone offending persons or unless high-rate offending persons who commit group offenses have a substantial mix of individual offenses, a substantial number of offenders must be removed from a population to affect substantially the crime rate.

Research on the effects on the crime rate of incapacitating offenders generally makes a simple presumption that each offense in an individual's rate of offending is a single offender offense. Thus, it is assumed that the number of crimes averted by incapacitating an offender is equal to the individual's rate of offending. Clearly that is not generally the case when an individual commits an offense as a member of an offending group. Were we to assume that our estimate of 2.1 offenders per crime against the person holds for the person offenses in an individual's crime rate, then it would be more reasonable to assume that on the average incapacitation would avert only one-half a crime for each person crime in the individual's rate of offending. But this would be a reasonable estimate only if we also assumed that for group offenses the replacement rate for incapacitated members is 0 and that there are no deterrent effects on the other group offenders as a consequence of incapacitating one of its members.

If we are reasonably correct in the foregoing argument, some more general inferences can be drawn. In investigating the effects of incapacitation on the crime rate, it is assumed that incapacitation effects can be separated from, and investigated apart from, deterrent effects. Clearly that is not the case where group offending is involved. Whenever an individual is involved in group offenses, unless incapacitation of that member has a *deterrent* effect on the

Percent distribution of incidents by size of offending group for selected types of crime incidents: All National Crime Survey incidents reporting offender information, July 1, 1972, to December 31, 1975

Percent of incidents by size of offending group									Total percent
1	2	3	4	5	6-10	11-14	15-19	20+	
70.0	13.3	10.0		6.7	-	-	-	-	100.0
79.5	11.6	4.5	1.8	1.8	0.9	-	-	-	100.1
64.1	17.9	10.2	2.6	2.6	2.6	-	-	-	100.0
81.9	5.7	3.8	2.9	1.6	3.5	0.6	-	-	100.0
36.8	35.2	16.2	3.9	2.8	4.1	0.3	0.5	0.1	99.9
44.8	23.7	15.3	6.0	4.2	5.4	0.3	0.2	0.2	100.1
54.6	24.1	11.1	4.9	1.4	3.2	0.7	-	-	100.0
49.7	22.8	12.7	5.6	3.4	4.1	0.9	0.4	0.5	100.1
30.0	28.9	18.9	10.2	5.6	4.4	1.0	0.5	0.5	100.0
39.4	24.0	19.7	11.3	1.4	2.8	-	-	1.4	100.0
59.7	12.6	9.1	6.1	3.1	5.8	1.4	0.7	1.5	100.0
74.2	10.7	5.3	2.9	0.8	4.5	0.4	-	1.2	100.0
44.3	23.5	18.8	5.6	1.6	5.4	-	0.4	0.4	100.0
70.8	10.1	7.3	4.0	2.4	3.9	1.1	0.4	0.2	100.2
67.2	12.4	7.6	4.1	2.7	4.2	0.7	0.4	0.8	100.1
70.6	11.0	6.7	4.0	2.6	3.9	0.5	0.3	0.4	100.0
61.9	25.3	0.9	2.8	0.8	0.4	-	-	-	100.1
61.4	24.3	8.9	3.5	1.5	0.5	-	-	-	100.1
65.1	21.8	8.9	2.7	1.6	-	-	-	-	100.1
52.6	30.5	10.5	4.2	2.1	-	-	-	-	99.9
68.3	22.2	5.6	1.6	1.6	0.8	-	-	-	100.1
61.4	25.0	11.4	0.2	-	-	-	-	-	100.0
72.8	17.4	6.6	1.7	0.5	1.0	-	-	-	100.0
66.0	18.7	9.6	2.7	1.1	1.3	0.5	-	-	99.9
60.0	25.0	9.2	2.5	0.8	1.7	0.8	-	-	100.0
70.7	19.2	6.7	1.9	-	1.5	-	-	-	100.0
66.7	19.0	5.7	4.7	2.1	1.4	-	0.4	-	100.0
75.2	14.9	7.0	1.0	1.7	0.3	-	-	-	100.1
71.1	16.3	6.6	3.8	0.9	1.1	0.2	-	-	100.0
73.3	16.4	6.4	2.0	1.1	0.7	0.1	-	-	100.0
68.9	19.4	5.8	4.9	-	1.0	-	-	-	100.0
56.5	23.8	11.1	5.8	1.2	1.4	0.2	-	-	100.0
71.6	17.6	6.8	4.1	-	-	-	-	-	100.1
45.0	30.0	17.5	2.5	3.3	1.7	-	-	-	100.0
78.6	7.1	7.1	-	-	7.1	-	-	-	99.9
53.9	30.8	15.3	-	-	-	-	-	-	100.0
14,210	3,505	1,932	925	517	749	120	61	88	
64.3	15.9	8.7	4.2	2.3	3.4	0.5	0.3	0.4	

other offenders involved with him, the effect of his incapacitation could easily be zero crimes averted, as the co-offenders could continue to account for the same number of crime incidents. If, however, the incapacitation of an offender has a deterrent effect on co-offenders, incapacitation may contribute substantially to reducing the crime rate. Any model of the effects of incapacitation on the crime rate should reflect deterrent effects on the network of offenders as well.

Models of the effect of incapacitation on the crime rate assume that decreasing the *prevalence* rate of offenders in a population has a corresponding effect on the crime rate. There are several reasons why this may be a weak assumption.

The models generally ignore replacement rates of incapacitated offenders. The size of an incapacitation effect depends upon the size of the group, the fluidity of its boundaries, and the rate of replacement. To measure the rate of replacement is difficult, since much depends upon the age of the offender, the mix of offenses, the group contribution to an individual's rate of offending, the structure of social networks, and any other factors that affect changes in replacement rates. If we assume that individuals are inducted into criminal careers largely in co-offending and that offending networks are reasonably large, we would expect high replacement rates. Indeed, since within any offending group, or within a network of any reasonably large size, say ten or more, any individual offender is quite likely to be linked with different offenders in different offenses, it is possible that this condition of relatively free substitution of group members in committing any given group offense may work against replacement if any member is selectively withdrawn, i.e., the group is simply reduced in size. But it also is possible that ease of substitution in offending may generate recruitment and actually increase group size.

Just how replacement takes place is not well understood. It seems unlikely that any high offending group recruits directly from the nonoffending population and thereby increases the prevalence rate of offenders in a population. Rather, if one thinks of groups with fairly open boundaries in a larger social network, then replacement follows somewhat the principles of replacement in a vacancy chain (White 1970). A vacated position in a high offending group is more likely to be filled by marginal offenders than by new offenders. Recruitment of *new* entrants to an offending population thus is unlikely to occur by direct replacement in high offending groups but indirectly in a chain of replacement; such direct effects on replacement will depend very much on the structure and extent of social networks. One might expect that networks are more extensive and open in high population density areas and areas where crime rates are high. Thus one might expect higher replacement rates in large cities than in rural areas. Similarly, replacement may be higher in young than older age groups since the boundaries of the former are more permeable and fluid.

The effect of incapacitation on replacement rates is undoubtedly complex not only for the reasons already mentioned but because replacement occurs over a period of time. Replacement is not always an immediate response to vacancy. All other things being equal, the longer a vacancy exists, the more likely any vacancy is to be filled by processes of internal shifts into vacancies with a final closure by recruitment from outside. Unless there are deterrent effects or substantial discontinuities in social networks that affect recruitment levels, one would expect replacement to be fairly high.

Nonetheless, one way that any organization may close off a vacancy is to vacate the position. Thus, it is possible that the effect

Table 1a:

Type of crime	Not known
Rape, w/theft	-
Rape, no theft	1.8
Att. rape, w/theft	-
Att. rape, no theft	0.9
Robbery, w/weapon	2.2
Robbery, no weapon	4.6
Att. robbery, w/weapon	1.4
Att. robbery, no weapon	2.5
Ser. assault, theft, weapon	4.2
Ser. assault, theft, no weapon	12.3
Ser. assault, weapon, no theft	4.6
Ser. assault, no weapon, no theft	2.4
Minor assault, theft	2.6
Minor assault, no theft	1.3
Att. assault, w/weapon, no theft	5.2
Att. assault, no weapon, no theft	2.4
Purse snatch, no force	20.7
Att. purse snatch, no force	3.3
Pocket picking	56.5
Burglary, forced entry, something taken	98.0
Burglary, forced entry, 0 taken, damage	88.3
Burglary, forced entry, 0 taken, 0 damage	86.8
Burglary, unlawful entry, 0 force	93.4
Burglary, att. forced entry	91.5
Larceny, $250+	95.7
Larceny, $100-249	96.7
Larceny, $50-99	96.5
Larceny, $25-49	96.9
Larceny, $10-24	96.7
Larceny, under $10	96.6
Larceny, $ unknown	95.1
Att. larceny	81.8
Car theft	96.5
Att. car theft	90.8
Theft, other vehicle	95.9
Att. theft, other vehicle	79.0
Total number	90,484
Percent	80.4

0.0 = less than half of one percent

of incapacitation on a group or network may be to reduce its size. But it is well to bear in mind that unless a group is at maximum productivity, the same amount of work may be accomplished by a smaller number of members in a smaller number of positions. Reductions in prevalence of offenders in a population can always be offset by an increase in individual rates of offending. From a sociometric perspective, if the probability that any member of a group of a given size will be selected into a subset of offenders for a given offense is one in "n" incidents, then the elimination of any member may only increase his probability of being selected, which in turn increases somewhat his individual rate of offending. It is also possible that replacement is related to the length of incapacitation of one of its members since the likelihood of return declines with the length of expected absence. Long-term incapacitation may lead to quicker replacement than short-term incapacitation.

We have noted that current models of the effect of incapacitation on the crime rate assume that reducing the prevalence of offenders in a population reduces the crime rate. We have proposed a more robust model, one where incapacitation can increase as well as decrease the crime rate. It also should be apparent that incapacitation can reduce the crime rate only under the conditions of less than total replacement in offending groups and no increase in the individual rates of offending of those who remain in the population.

The proposed model assumes then that incapacitation can *increase* both the *prevalence rate*, through a complex process of recruitment to offending groups which actually increases the offending proportion of a population, and *individual rates of offending*. Indeed, if the effect of incapacitating a member of an offending group increases the recruitment effort of all remaining members, the replacement rate might be greater than one, and the offending group might increase in size. Assuming that the supply of crime victims for a population of offenders is elastic, an increase in group size can increase the group offending rate. The problem is undoubtedly more complex since it depends upon both aggregate effects of recruitment in a population of offenders and reallocation of offending groups whose rates of aggregate offending vary.

Percent distribution of number of reported offenders by type of crime for all National Crime Survey incidents: July 1, 1972, to December 31, 1975

Number of offenders									Offense total		Percent reporting offenders
1	2	3	4	5	6-10	11-15	16-20	21+	Percent	Number	
70.0	13.3	10.0	-	6.7	-	-	-	-	100.0	30	100.0
78.1	11.4	4.4	1.7	1.7	0.9	-	-	-	100.0	114	98.2
64.1	18.0	10.3	2.6	2.6	2.6	-	-	-	100.1	39	100.0
81.2	5.6	3.8	2.8	1.6	3.5	0.6	-	-	100.0	319	99.1
36.0	34.5	15.9	3.8	2.7	4.0	0.3	0.5	0.1	100.0	769	97.8
42.7	22.6	14.6	5.8	4.0	5.1	0.3	0.2	0.2	100.1	625	95.4
53.9	23.7	11.0	4.8	1.4	3.2	0.7	-	-	100.1	438	98.6
48.4	22.2	12.3	5.5	3.3	4.0	0.8	0.3	0.5	100.1	599	97.5
28.5	27.6	18.0	9.7	5.3	4.2	0.9	0.6	0.6	100.0	544	95.8
34.6	21.0	17.3	9.9	1.2	2.5	-	-	1.2	100.0	81	87.7
57.0	12.1	8.7	5.8	3.0	5.5	1.3	0.6	1.4	100.0	1,403	95.4
72.4	10.4	5.2	2.8	0.8	4.4	0.4	-	1.2	100.0	250	97.6
43.2	22.9	18.3	5.4	1.6	5.2	-	0.4	0.4	100.0	498	97.4
69.9	9.9	7.2	4.0	2.3	3.8	1.1	0.4	0.1	100.0	1,972	98.7
63.8	11.7	7.2	3.9	2.6	4.0	0.6	0.4	0.7	100.1	3,468	94.8
68.9	10.7	6.5	3.9	2.6	3.8	0.5	0.3	0.4	100.0	6,087	97.6
49.0	20.1	7.0	2.2	0.6	0.3	-	-	-	99.9	314	79.3
59.3	23.4	8.6	3.5	1.4	0.5	-	-	-	100.0	209	96.7
28.3	9.5	3.9	1.2	0.7	-	-	-	-	100.1	1,014	43.5
1.0	0.6	0.2	0.1	0.0	-	-	-	-	99.9	4,952	2.0
8.0	2.6	0.7	0.2	0.2	0.1	-	-	-	100.1	1,077	11.7
8.1	3.3	1.5	0.3	-	-	-	-	-	100.0	333	13.2
4.7	1.1	0.4	0.1	0.0	0.1	-	-	0.0	99.8	9,073	6.6
5.5	1.6	0.8	0.2	0.1	0.1	0.0	-	-	99.8	4,457	8.5
2.6	1.1	0.4	0.1	0.0	0.1	0.0	-	-	100.0	2,796	4.3
2.3	0.6	0.2	0.1	-	0.0	-	-	-	99.9	6,374	3.3
2.3	0.7	0.2	0.2	0.1	0.1	-	0.0	-	100.1	8,057	3.5
2.3	0.5	0.2	0.0	0.1	0.0	-	-	-	100.0	9,700	3.1
2.3	0.5	0.2	0.1	0.0	0.0	0.0	-	-	99.8	13,806	3.3
2.4	0.6	0.2	0.1	0.0	0.0	0.0	-	-	99.9	22,368	3.4
3.3	0.9	0.3	0.2	-	0.0	-	-	-	99.8	2,129	4.9
10.2	4.3	2.0	1.0	0.2	0.3	0.0	-	-	99.8	4,667	18.2
2.4	0.6	0.3	0.1	-	0.0	-	-	-	99.9	2,192	3.5
4.0	2.7	1.6	0.2	0.3	0.2	-	-	-	99.8	1,344	9.2
3.0	0.3	0.3	-	-	0.3	-	-	-	99.8	369	4.1
11.3	6.5	3.2	-	-	-	-	-	-	100.0	124	21.0
14,210	3,505	1,932	925	517	749	120	61	88	-	112,591	-
12.6	3.1	1.7	0.8	0.5	0.7	0.1	0.1	0.1	100.1	-	19.6

There is also a complex relationship between the mix of individual and group offenses in an individual's rates of offending and membership in offending groups. It does not seem reasonable to assume that if an individual commits an offense at some times alone and other times as a member of an offending group that the lone offense is entirely independent of the group. There is some reason to believe that in the early stages of a person's criminal career, most of the offenses any offender commits are group offenses. Later criminal careers, by comparison, appear to have a larger proportion of individual offenses. Just how groups affect the propensity of persons to offend alone is not known. What is known is that among young offenders, the continuous lone offender is uncommon, but just how common he is at later ages is not known.

Models for measuring the effect of incapacitation on the crime rate usually assume that individual rates of offending are a constant—a tenuous assumption. Where variation is assumed, it is postulated that the distribution of individual rates of offending is such that a relatively small proportion of individuals make a disproportionate contribution to the crime rate. If, however, individuals with high rates of offending account disproportionately for group offenses, then their contribution to the overall crime rate is substantially reduced. Correlatively, if high offending individuals are disproportionately lone offenders, or if offenses with a single offender contribute disproportionately to higher offender rates, then their contribution to the overall crime rate is substantially greater. The effect of selective incapacitation of offenders with high rates of offending then may be particularly sensitive to the mix of individual and group offenses in individual and group offending rates.

Similarly, from the perspective of a group offending rate, if certain offenders are disproportionately involved in most group offenses, their incapacitation may have a substantial effect on the group offending rate. This would be the case particularly if such high-rate offending persons are less easily replaced, and their incapacitation would have a greater deterrent effect on the propensity of less frequently involved offenders to commit crime than would the incapacitation of less frequently involved offenders on the propensity of the most involved offenders.

It is conceivable that the selective incapacitation of offenders who are disproportionately involved in the offenses of a group might *not* have a substantial effect on the remaining members. Several conditions might offset the expected effect of incapacitation. One is that the remaining members recruit additional members (Hoekema 1973). A second is that the group reallocates members of offending positions within the offending group. Short and Strodtbeck (1965) note that vacancies in the gang leadership position commonly are filled by reallocation. Another possible result is fragmentation of the original group and its members' selective recruitment to other groups where their individual offending rates might increase; or the remaining members might become two or more nuclei around which new offending groups form. In either case, there is a potential for increasing both the prevalence rate and individual rates of offending. Whether these alternative possibilities occur and with what frequency, or whether the deterrent effect on remaining members is the most likely result, requires investigation.

It is well to bear in mind that replacement is a continuous process in many different kinds of groups, including offending groups. The American population is fairly

Table 2.

Type of crime	Total number offenders
Rape, w/theft	48
Rape, no theft	156
Att. rape, w/theft	68
Att. rape, no theft	513
Robbery, w/weapon	1,780
Robbery, no weapon	1,429
Att. robbery, w/weapon	861
Att. robbery, no weapon	1,382
Ser. assault, theft, weapon	1,488
Ser. assault, theft, no weapon	179
Ser. assault, weapon, no theft	3,537
Ser. assault, no weapon, no theft	480
Minor assault, theft	1,164
Minor assault, no theft	3,860
Att. assault, w/weapon, no theft	6,996
Att. assault, no weapon, no theft	11,676
Purse snatch, no force	392
Att. purse snatch, no force	327
Pocket picking	679
Burglary, forced entry, something taken	164
Burglary, forced entry, 0 taken, damage	189
Burglary, forced entry, 0 taken, 0 damage	68
Burglary, unlawful entry, 0 force	876
Burglary, att. forced entry	625
Larceny, $250+	212
Larceny, 100-249	310
Larceny, 50-99	474
Larceny, 25-49	425
Larceny, 10-24	705
Larceny, under 10	1,095
Larceny, $ unknown	157
Att. larceny	1,530
Car theft	114
Att. car theft	238
Theft, other vehicle	24
Att. theft, other vehicle	42
Total	44,263
Percent, incidents against persons only	
Percent, all incidents	

mobile and transient. Turnover rates in the membership of schools and communities are very high where crime rates are high and offenders are disproportionately concentrated. This suggests that peer group structures are far from static. There is continual recruitment and replacement of members. Just how substantial replacement of members is for offending groups because of residential mobility is not known, but it is possible that transiency itself has an effect on the prevalence rate of offenders in a population. The effects of recruitment and replacement on offending may be greater for young than for older persons and greater for closed than for open social networks or groups.

Codicil

Readers will be quick to detect weaknesses in the proposed model for understanding changes in the crime rate and its implication for conclusions about the effects of deterrence and incapacitation on the crime rate. Some will be quick to note that the structure of social networks and groups is less formal than the argument supposes. Others will argue that human behavior is less rational than the argument assumes. Still others will suggest that the burden of evidence hardly warrants a model which assumes human intervention to reduce crime rates may actually increase them. And so it may be. Beneath the argument, however, lies a presumption that offending groups may not be all that different from other kinds of groups in a society. And the proposed model provides a basis for examining whether incapacitation may increase as well as decrease crime rates.

References

Hoekema, A. J. (1973)
Rechtsnormen en Sociale Feiten: Theorie en Empirie rond de Kleine Haven diefstal. Rotterdam.

Short, James F. and Fred L. Strodtbeck (1965)
Group process and gang delinquency. Chicago: University of Chicago Press.

White, Harrison (1970)
Chains of opportunity. Cambridge: Harvard University Press.

Percent distribution of estimated number of offenders by type of crime and mean number of offenders per incident for all National Crime Survey incidents where number of offenders was reported.
July 1, 1972, to December 31, 1975

Percent of offenders for each size of offending group									Total percent	Number of incidents	Number offenders per incident
1	2	3	4	5	6-10	11-15	16-20	21+			
43.7	16.7	18.8		20.8					100.0	30	1.6
57.1	16.7	9.6	5.1	6.4	5.1				100.0	112	1.4
36.8	20.6	17.5	5.9	7.4	11.8				100.1	39	1.7
50.5	7.0	7.0	7.0	4.9	18.3	5.3			100.0	316	1.6
15.6	29.8	20.6	6.5	5.9	14.8	1.5	4.2	1.2	100.1	758	2.3
18.7	19.7	19.1	10.1	8.7	19.0	1.9	1.3	1.5	100.0	603	2.4
27.4	24.2	16.7	9.8	3.5	13.8	4.6			100.0	434	2.0
21.0	19.2	16.1	9.6	7.2	14.8	4.9	2.7	4.6	100.1	591	2.3
10.4	20.2	19.8	14.3	9.8	13.2	4.6	3.7	4.2	100.2	519	2.9
15.7	19.0	23.5	17.9	2.8	9.5			11.7	100.1	71	2.3
22.6	9.6	10.4	9.2	5.9	18.5	7.3	4.7	11.9	100.1	1,339	2.6
37.7	10.8	8.1	5.8	2.1	19.6	2.7		13.1	99.9	244	2.0
18.5	19.6	23.5	9.3	3.4	19.0		3.2	3.6	100.1	485	2.4
35.7	10.2	11.0	8.1	6.0	16.7	7.4	3.3	1.6	100.0	1,958	2.0
31.6	11.6	10.7	7.7	6.4	16.8	4.2	3.4	7.5	99.9	3,319	2.1
35.9	11.2	10.2	8.1	6.7	16.8	3.6	3.0	4.5	100.0	5,991	1.9
39.3	32.1	16.8	7.1	2.6	2.0				99.9	250	1.6
37.9	30.0	16.5	8.6	4.6	2.4				100.0	202	1.6
42.3	28.2	17.2	7.1	5.2					100.0	443	1.5
30.5	35.4	18.3	9.8	6.1					100.1	95	1.7
45.5	29.6	11.1	4.2	5.3	4.2				99.9	126	1.5
39.7	32.4	22.1	5.9						100.1	45	1.5
48.9	23.3	13.4	4.6	1.7	5.8			2.4	100.1	595	1.5
39.5	22.4	17.3	6.4	3.2	6.9	4.3			100.0	379	1.6
34.0	28.3	15.6	5.7	2.4	8.0	6.1			100.1	120	1.8
47.4	25.8	13.5	5.2		8.1				100.0	210	1.5
39.2	22.4	10.1	11.0	6.3	17.2				100.0	281	1.7
53.4	21.2	14.8	2.8	5.9	1.9		3.8		100.0	303	1.4
45.8	21.0	12.8	9.6	2.8	6.1	1.8			99.9	460	1.5
50.1	22.5	13.1	5.5	3.7	3.9	1.2			100.0	753	1.5
45.2	25.5	11.5	12.7		5.1				100.0	104	1.5
31.1	26.1	18.2	12.8	3.3	6.7	1.8			100.0	851	1.8
46.5	22.8	13.2	10.5		7.0				100.0	76	1.5
22.7	30.3	26.5	5.1	8.4	7.1				100.1	124	1.9
45.8	8.3	12.5			33.3				99.9	15	1.6
33.3	38.1	28.6							100.0	26	1.6
14,210	7,010	5,796	3,700	2,585	6,366	1,620	1,128	1,848		22,303	
30.2	14.2	12.7	8.5	6.3	16.1	4.1	3.0	4.9	100.0		2.1
32.1	15.8	13.1	8.4	5.8	14.4	3.7	2.5	4.2	100.0		2.0

Criminal opportunities and crime rates

RICHARD F. SPARKS
School of Criminal Justice
Rutgers University

A statistic very commonly used by criminologists is the *crime rate*—e.g., the number of crimes known to the police per 100,000 population. In recent years, with the development of victimization surveying, similar use has been made of the *victimization rate*—e.g., the number of victimizations reported per 1,000 persons or households. It is argued here that such rates need very careful interpretation and that for many purposes they may be extremely misleading. It is also argued that in the calculation of such rates it is generally desirable to take into account *opportunities* for committing illegal acts as well as the population of potential offenders at risk.

Purposes for which crime rates are calculated

The purpose of this paper is to consider some conceptual and statistical properties of crime and/or victimization rates. Before doing that, however, it is necessary to review briefly the reasons why such rates are used as measures of criminal behavior or its consequences and, more generally, why it has been thought important to measure crime or victimization in particular times or places. There appear to be four main reasons:

1. Historically, the first purpose for collecting statistics on crime and criminals appears to have been the measurement of the "moral health" of nations, cities, etc. The names given by the earliest demographers and statisticians to their measures of crime—*Moralstatistik, statistique morale*—give a sufficient indication of this objective; if the numbers of crimes or criminals increased, then in some sense the moral "health" of the nation was growing worse. Something of this concern appears to linger on in popular interpretations of crime statistics: a rising crime rate is not infrequently seen as an indication of increased depravity or decreased probity, or as a sign of a usually ill-defined "social pathology." Somewhat similarly, Taylor, Walton, and Young (1975:42) have argued that from a radical perspective crime statistics can be used as an "examination of the extent of compliance in industrial society (in quite the same way ... as it is possible to use statistics on strikes as an index of dissensus in direct class relations at the work-place)."

2. A second purpose for measuring crime has been the evaluation of the effectiveness of the machinery of social control. As is well-known, Bentham (1778) was one of the first to urge that accurate measurement of crime was a necessary adjunct for the legislator; he urged the collection of statistics on convictions and prisoners as "a kind of *political barometer*, by which the effects of every legislative operation relative to the subject may be indicated and made palpable."

3. A third reason for measuring crime is the estimation of the risk of becoming a victim. This concern is present, though often implicit, in contemporary efforts to develop "social indicators" (cf. Bauer 1964, esp. chap. 2). As victimization surveying has developed over the past decade, the assessment of risk has become increasingly prominent; indeed, it appears to have been one of the main objectives of the National Crime Surveys (NCS) now being conducted by the Census Bureau for the Law Enforcement Assistance Administration (see Penick and Owens 1976:143-45).

4. Finally, the measurement of crime has been a necessary preliminary to the development and testing of criminological theories. Typically the testing of such theories has involved comparisons of crime rates in different places or types of place (for example, cities versus suburbs), or over time, or attempts at correlating changes in candidate independent or explanatory variables with changes in crime rates.

It seems intuitively obvious that each of these four objectives requires the calculation of crime rates—that is, the number of crimes committed in a given place and time period must be standardized for the population in that place at that time period. It seems *prima facie* absurd to compare the number of homicides in the United Kingdom with the number of homicides in the United States, given that the United States has about four times as large a resident population as the United Kingdom; similarly, not much can be inferred from a comparison of the number of thefts in the United States in 1933 with the number of thefts in the United States in 1977, given that the U.S. resident population rose by over 70 percent in that time period. Whether our concern is with social morality, social control, the assessment of risk or the testing of theory, it is necessary to control for variations in crime which are due to differences in the numbers of persons able to commit crimes; criminological theories, for example, are theories about the causes of crime, and not about population growth. Similarly, we would not conclude that people were becoming more dishonest, social control less effective, or life more dangerous, without standardizing for population size.

The calculation of crime and victimization rates

A simple rate, like a crime or victimization rate, is a function of only two elements: (1) a number of acts, events, situations, etc., which occur in a given place and time period; and (2) a number of persons or other elements present in the same place and time period. Thus a crime rate R_c is typically defined by

$$R_c = \frac{kC}{P} \quad (1)$$

where C = number of crimes committed,
P = number of persons available to commit crimes,
k = a number chosen either to give a convenient rate or a convenient base (e.g., 100,000 persons).

Thus a rate R_c has a natural verbal translation: "For every k persons, R_c crimes were committed."

A victimization rate R_v is typically defined in a similar fashion, but with two differences. First, the numerator of the right-hand side contains the number of victimizations rather than the number of crimes; depending on the definitions of these two things, and the "counting rules" used for each, they need not be identical.[1] Second, the denominator is typically the number of persons, organizations, etc. *capable of being victims*. Thus the current NCS surveys, for example, compute commercial victimization rates to a base of (nongovernmental) recognizable businesses; no account is taken of the population of persons able to rob or burgle those businesses. (See, e.g., National Criminal Justice Information and Statistics Service 1976:123.)

Rates of this kind measure the *incidence* of crime or victimization, since their numerators contain (essentially) numbers of events. But analogous rates can be constructed which measure the *prevalence* of crime-committing or being a victim; for such rates the numerator is usually the number of persons (organizations, etc.) who had committed one or more crimes, or been a victim on one or more occasions, in a given time period. Since a single offender may commit more than one offense in a given time period, or a person or organization may be a victim on more than one occasion, incidence and prevalence rates are not necessarily identical. (Compare death rates, where the number of deaths is necessarily identical with the number of persons who die.)[2]

[1]For a discussion of the definitions and "counting rules" used in the *Uniform Crime Reports*, and their relation to those used in the National Crime Surveys, see Hindelang (1976:89-97).

[2]In the case of phenomena which have some temporal duration (e.g., diseases), a further distinction is sometimes drawn between "point-prevalence" rates and "period-prevalence" rates. Crime and victimization prevalence rates are of the latter type, i.e., they give the percent of the population at risk who were criminals or victims within a time period such as a year.

Crime and victimization rates raise a number of well-known problems of measurement. Typically we are interested in the numbers of crimes which are actually committed; but statistical series like the *Uniform Crime Reports* of course give only the numbers of crimes "known to the police." Similarly, victimization surveys aim to measure the numbers of victimizations which actually occur; what they get instead is the number of victimizations correctly recalled by survey respondents, reported to interviewers, etc. Each rate thus has, in its numerator, a "dark figure" of incidents which are not counted. Similar problems can also occur with the denominators of these rates, through underenumeration in a census; typically, however, these are much less serious. Having noted these problems of measurement, I shall from now on ignore them; my interest is in the interpretation of crime and/or victimization rates, and not with the accuracy of the counts of incidents or persons which they may involve.

The first of these problems of interpretation is a purely statistical one. It is obvious that a rate defined as in (1) is a kind of average; it is in fact a function of the arithmetic mean number of crimes committed per person. But such an average, taken over the whole of a population, clearly need not represent the experience of any individual or subgroup within that population. A death rate for the whole of a population—sometimes called a "crude" death rate—may conceal considerable variations in the incidence of death in various subgroups of that population. For this reason it is customary to calculate separate rates for subgroups whose experience is known to be different; e.g., age-specific or race-sex-specific rates, or rates associated with different causes of death as well as with different populations. (Cf. Reiss 1967:22-23.) Such rates make possible between-group comparisons; for instance, of the risk of dying of heart disease at age 15 compared with the risk at age 75. Moreover, if the subgroups used to calculate such "specific" rates are reasonably homogeneous with respect to the phenomenon being measured, the resulting rates will not be very misleading as within-group descriptions of experience or risk. For example, if every white male aged 21 on his last birthday had an approximately equal chance of contracting smallpox by his 22nd birthday, an age-specific rate of infection of smallpox would give an accurate measure of risk to each individual, though of course it would not describe any individual's actual experience (since either he catches smallpox or he does not). The same thing is true for phenomena like crime or victimization, which can involve the same individual more than once in any noninfinitesimal time period. If every member of a given subgroup were to commit or suffer (say) exactly two crimes per year, then the resulting rate would necessarily reflect each individual's experiencce in that year. Though this is of course unlikely to happen, a crime rate would still not be too misleading, provided that the within-group variance were small, relative to the subgroup mean (i.e., in proportion as the coefficient of variation approached zero). Finally, even if the within-group variance were considerable, the rate might not be too misleading, provided the distribution were approximately normal (or more generally were symmetrical about its mean).

It seems clear that this is generally not the case, however, either for crimes committed or victimizations experienced. Data from a number of studies to date strongly suggest that the frequency distributions of crime-related events are typically extemely skewed, with the majority of the population having no crimes or victimizations in a given time period, and at the other extreme a small proportion of the population having a great many. It follows that a crime or victimization incidence rate will be an extremely misleading descriptor of the group's experience, or of the risk of crime or victimization.

This point can be illustrated with data taken from a victimization survey which I conducted in three Inner London areas in 1973 (see Sparks, Genn, and Dodd 1977, chap. 4). Table 1 gives the numbers of respondents reporting 0, 1, 2, ... incidents of victimization of various types as having happened within the survey reference period (approximately the calendar year 1972), together with sample victimization rates per person. It will be seen that for two of the three areas, and for the sample as a whole, the total victimization rate is in excess of 1.0 per person. A naive interpretation of these rates might suggest that

Table 1. Distribution of victimization incidents in three Inner London areas in 1972, and expected numbers based on Poisson distribution (all three areas)

Number of incidents	Brixton		Hackney		Kensington		Total		Expected numbers ($\lambda = 1.07$)
	Number	Percent	Number	Percent	Number	Percent	Number	Percent	
None	101	56	104	58	93	51	298	55	187
1	40	22	40	22	40	22	120	22	200
2	13	7	11	6	32	17	56	10	107
3	14	8	18	10	8	4	40	7	38
4	6	3	2	1	3	2	11	2	10
5	2	1	3	2	1	1	6	1	2
6+	6	3	1	1	7	4	14	3	1
Totals	182	100	179	100	184	100	545	100	(545)
Total numbers of incidents:	208		151		223		582		—
Mean Numbers of incidents:	1.14		.84		1.21		1.07		—

everyone in the sample was a victim at least once in the year; or, alternatively, that the risk of victimization in those areas was about 100 percent, i.e., virtual certainty; yet, as the table shows, over half of the respondents in each area reported no incidents at all.

Similar findings have emerged from a number of other victimization surveys done in recent years (see, e.g., Aromaa 1971, 1974; Wolf and Hauge 1975; Reynolds 1973). The same general picture appears to be emerging from the NCS surveys. In the commercial victimization survey conducted in Houston, Texas, for example, the aggregate rate for robbery and burglary was 1.278 incidents per establishment; yet nearly 60 percent of the businesses surveyed reported no incidents at all as having occurred within the 1-year reference period (see Penick and Owens 1976:127-30; Hindelang 1976:22). In the case of the NCS surveys this problem is especially serious, since in LEAA's published reports to date on these surveys, the victimization rate is virtually the only statistic used.

Given a skewed distribution of the kind disclosed by Table 1, the victimization rate might still have some readily interpretable meaning in terms of victimization experience and/or risk, if the occurrence of multiple victimization were approximately random, i.e., if it more or less conformed to a Poisson distribution with the overall mean as a transition rate. As the right-hand column of Table 1 shows, however, this is not the case: the numbers reporting no victimization at all, and the numbers reporting several incidents, are both greater than would be predicted by a Poisson distribution. Compound distributions—based on assumptions of "contagion" or increasing probability of victimization (Coleman 1964:299-305) or of heterogeneity of "proneness" to victimization (Greenwood and Yule 1920) give a somewhat better fit to data like those of Table 1; so does the skew distribution first described by Yule (1924) and discussed by Simon (1957:145-64), which is based on slightly different assumptions. Using the London survey data an attempt was made to identify empirically (following a suggestion in Coleman 1964:378-80) subgroups of the sample with different mean rates of victimization, for whom the distributions of incidents could be adequately described by separate simple Poisson processes. Unfortunately, this attempt was unsuccessful. No set of criteria—based on attributes such as age, sex, race, or social class, either singly or in combination—was found by which the sample could be subdivided into groups in which the frequency distribution of multiple victimization was no greater than would be expected by chance. (See Sparks, Genn, and Dodd 1964:166-67; for a similar analysis with no better result, see Aromaa 1973.)

In summary, in the present state of our knowledge, even specific subgroup rates are apt to be extremely misleading as descriptors of the experience of, or the risk of, victimization. A prevalence measure—such as the percentage who are victimized on one or more occasions in a given time period—is somewhat less misleading. But such a measure completely masks the extreme cases of multiple victimization which occur; if this is to be avoided, then the full frequency distribution must be presented.[3]

Opportunities and rates

The concept of opportunity is familiar in criminology, chiefly through the work of Merton (1938) and Cloward and Ohlin (1960).[4] Though seldom explicitly referred to, the concept has also played a part in many less elaborate attempts to explain variations in crime rates over time or place. Thus, for example, it has often been noted that there are well-marked seasonal variations in observed patterns of crime, with crimes of violence typically being more common in the summer months and crimes such as burglary being more common in the winter: a common explanation for such findings is that social interaction is greater in the summer, thus providing greater opportunity for interpersonal violence,

[3]Though the evidence is much less complete, it appears that the committing of crimes is distributed in a similar fashion. Carr-Hill (1971) found that convictions for crimes of violence among adult males in England and Wales displayed a distribution not unlike that of Table 1: most of the population had no convictions, while a small proportion had many. As with victimization, a crime rate based on such a distribution would be extremely misleading: it would greatly overstate the involvement in crime of the majority, while, of course, understating the activity of the "crime-prone" minority.

[4]Though it is interesting to note that, in general, both Merton and Cloward and Ohlin tended to regard legitimate and illegitimate opportunities as alternatives, and thus mutually exclusive: either one obtained a legitimate job or he joined the rackets. But—at least in Western industrial societies—the major opportunities for illegitimate gain open to most people involve theft of some kind from their places of employment: thus legitimate and illegitimate opportunity structures are intimately connected. For a study of blue-collar theft illustrating this point, see Horning (1964).

whereas longer hours of darkness in the winter months provide greater opportunity for undetected entry into others' property.

More recently, a few researchers have explicitly considered the relations between crime and opportunities for it. Before considering these approaches, however, we need to examine the relations between an opportunity for committing a crime, and the commission of crime itself. It is clear that, as a matter of ordinary language, the existence of an opportunity for a crime to be committed is a logically necessary condition of that crime's occurring. That is, if we are prepared to assert that an opportunity to commit, e.g., a theft at a particular time and place did *not* exist, then we should normally be compelled to say that no theft did in fact take place. Thus it is a necessary truth, and not merely a very well-confirmed hypothesis, that no motor cars were stolen in the United States (or anywhere else) in the year 1850; that no room air-conditioners were stolen in the year 1900; that no color television sets were stolen in 1930; and that no credit-card frauds were committed in 1940. The opportunities for those crimes simply did not exist in those years.

The proposition that *changes* in opportunities to commit crime will lead to *changes* in the numbers of crimes actually committed appears to be a hypothesis—to involve a contingent matter of fact, and not a truth of logic. But the matter is more complicated than that. Certainly it is not necessarily true that if the amount of stealable property increases, the number of thefts will, *ceteris paribus*, increase. However, a *decrease* in the quantum of stealable property, social interaction, etc., may of necessity lead to a decrease in thefts, assaults, etc., if it results in some individuals who formerly had opportunities to commit these acts no longer having them. Thus, suppose that in a time period t_1 every member of a population of N persons has some opportunities to steal; a further increase in opportunities in that population may lead to more thefts, or it may not. Suppose at t_2 the number of opportunities for theft is reduced, so that k individuals are completely without opportunities (so that thefts can only be committed by $N - k$ members of the population); all other things being equal, the number of the thefts will *necessarily* decrease. Evidently if we are comparing numbers of thefts committed at t_1 and t_2 we must take account of changes in opportunities between those two periods; and this is so whether t_1 precedes t_2 or follows it.

It follows that, for any of the four objectives of measurement mentioned earlier, opportunities for crime need to be taken into account in calculating crime rates for comparisons across time or place. Thus, if the crime rate is to be used as an indicator of social morality, probity, violence-proneness, collective wickedness, etc., it is in effect being interpreted as an average tendency in the population to behave in certain illegal ways; but a person's actually behaving in those ways presupposes that he has the opportunity to do so. Suppose that we associate with each person in the population a tendency to steal, assault others, etc.; borrowing a bit of economists' jargon, we might speak of a *propensity* to steal, assault, defraud, etc. Such a propensity can be defined as the conditional probability that an individual will steal, assault, etc., *given that* he has the opportunity to do so. The evidence discussed earlier suggests that the distribution of this propensity in the population will be skewed rather than normal. The unconditional probability that an individual will steal, $p(T)$, would then be given by the product of this conditional probability or propensity, and the probability $p(O)$ that he has the necessary opportunity:

$$p(T) = p(T|O) \cdot p(O) \qquad (2)$$

But $p(O) = 0$ by definition, when no opportunities exist; thus $p(T)$ will also necessarily be zero under those conditions. Thus if the average (or marginal) propensity to steal remains constant in a population between t_1 and t_2, but opportunities for thefts increase, the probability of theft (and probably the numbers of thefts actually committed) will increase; but that is not an indication that the population is becoming any more dishonest. In somewhat old-fashioned language, we might describe such a situation by saying that the number of temptations had increased (so that some who formerly had no such temptations available now had them); not that people were becoming any more susceptible to such temptations as they had.

The unconditional probability $p(T)$ can obviously serve as a transition rate in a Poisson process leading to actual thefts. The numbers of thefts occurring would in this case be a random variable depending in part on "theft-proneness" and in part on chance factors; models such as those of Greenwood and Yule (1920) would fit this situation. Such models assume that crime-proneness is not uniform in the population; and while the evidence for this proposition is admittedly thin, the proposition itself is intuitively reasonable. This point has been neglected in some recent attempts at modelling criminal behavior (see, for example, Avi-Itzhak and Shinnar 1973; Shinnar 1975, 1977). Shinnar does refer to the average number of crimes committed (which he designates by λ) as a random variable; but since he deals constantly in the expectation of this variable, he often seems to assume that it is literally the same for every member of the population of "criminals." This assumption may suffice for the purpose of making overall estimates of the incapacitative effect of imprisonment; as Blumstein et al. (1978:69) have noted, most models of the incapacitative effect can be generalized to incorporate distributions of parameters such as λ. (See also Greenberg 1975; Cohen 1978.) But the assumption of a homogeneous λ or proneness seems mistaken, if what is wanted is to model *individuals'* criminal careers (which is in fact important for some incapacitative strategies, e.g., those aimed at "high-risk" offenders). A further complication is that what may be called the "velocity" of individuals' criminal behavior probably varies; that is, λ is probably not constant over time. For a discussion see Green (1978); and for a discussion of models which can be used to tackle such problems, see Fienberg (1977); Bartholomew (1973).

The point being made here, however, is that $p(T)$ is itself a function not merely of individuals' propensities to behave in certain ways (i.e., $p(T|O)$, which is the "real" proneness at issue), but also of their opportunities to exercise those propensities. The social and spatial distributions of those opportunities need to be taken into account, in any full attempt to describe or explain criminal behavior.

A similar consideration applies if the crime rate is to be used as a measure of the system of social control, or as a dependent variable in a criminological theory or its testing. In each case, what the crime rate is supposed to measure is (approximately) the tendency of the population to behave in certain illegal ways, under specified control arrangements (e.g., a particular set of penalties) or specified social-structural or other conditions (e.g., a given level of unemployment, status integration, or relative deprivation). Plainly variations in opportunities must be controlled for, if changes in crime rates are to be interpreted correctly: a sharp decrease in opportunities for crime, for example, could be expected to lead to a decrease in

Table 2. Cars registered, population, and thefts from motor cars, England and Wales, 1938-1961; thefts standardized for population and stocks of cars; and another variable

1 Year	2 Thefts from cars	3 Cars registered ('000)	4 Population ('000)	5 Thefts/population/cars	6 "X"
1938	25,281	1,944	41,125	100.0	1,527.3
1940	16,849	1,423	39,889	94.1	2,397.6
1941	15,672	1,503	38,743	85.3	1,984.5
1942	12,180	858	38,243	117.6	1,672.2
1943	11,084	718	37,818	129.4	1,513.1
1944	14,509	755	37,785	161.2	1,441.2
1945	26,520	1,487	37,916	149.1	1,294.6
1946	32,546	1,770	40,759	143.0	1,462.2
1947	33,984	1,943	41,786	132.7	1,872.3
1948	32,665	1,961	43,296	122.0	2,449.0
1949	30,297	2,131	43,595	103.4	2,205.4
1950	33,156	2,258	43,830	106.2	2,301.3
1951	43,127	2,380	43,815	131.1	2,465.2
1952	41,125	2,508	43,955	118.2	2,412.8
1953	39,739	2,762	44,109	103.4	2,173.6
1954	33,398	3,100	44,274	91.0	2,194.1
1955	43,304	3,562	44,441	87.6	2,330.8
1956	50,782	3,888	44,667	92.7	2,575.6
1957	54,937	4,187	44,907	92.6	2,969.0
1958	68,466	4,549	45,109	105.7	3,419.4
1959	79,899	4,966	45,386	112.4	3,512.3
1960	92,704	5,526	45,755	116.2	4,153.5
1961	112,671	5,979	46,166	129.4	4,846.6

Wilkins (1964:55).

criminal behavior independently of any changes in presumed causal factors or the social-control system, merely because it was no longer possible for some people to commit crimes which they would otherwise have committed.

Finally, variations in opportunity must be taken into account in assessing the risk of victimization. If, for example, the number of cars in use doubles while the number of car thefts only goes up by fifty percent, then the average risk of a car's being stolen has declined; similarly, if people cease to go out of their houses at night, their risk of being assaulted in the street at night obviously declines.

One of the first researchers to take into account variations in opportunity was Sarah Boggs, in her study of urban crime patterns (1965). Boggs noted that

> Environmental opportunities for crime vary from neighborhood to neighborhood. Depending on the activities pursued in different sections of the city, the availability of such targets as safes, cash registers, dispensing machines, people and their possessions varies in amount and kind. These differing environmental opportunities should be reflected in the occurrence rates [of crime]. (Boggs, 1965:899)

Boggs noted correctly that a consequence of population-standardized crime rates was the production of "spuriously high" crime rates for central business districts, which contain relatively small resident populations but large amounts of merchandise, parked cars, and so on. Accordingly, she constructed "crime-specific" rates using denominators which could reflect opportunities for the type of crime in question: a business-residential land-use ratio for commercial burglary and robbery, the amount of street space available for parking, in the case of car theft, and so on. Rank correlations between these rates and rates standardized for population only, across 128 census tracts in St. Louis, were very high for highway robbery, residential burglary, rape, homicide, and aggravated assault; but they were low and in some cases negative for other offenses, e.g., $\tau = -.23$ in the case of nonresidential daytime burglary (Boggs 1965:901).

It is important to note, however, that standardization for variations in opportunities for crime is not an *alternative* to standardization for the size of the population available to commit crimes (as Boggs's analysis suggests). Instead, *both* the population of potential offenders, *and* the stock of available opportunities for crime, should be reflected in the denominator of a crime rate. Thus, an opportunity-standardized rate R_c^* would be defined by

$$R_c^* = \frac{kC}{PO} = \frac{R_c}{O} \tag{3}$$

where O = opportunities for the type of crime in question,
k = a constant chosen to give a convenient base or rate (e.g., numbers of cars stolen per 1,000 persons able to steal cars, per 1,000 cars available to steal.

(Note that R_c^* is defined as zero when O is zero, since in that case C is necessarily zero as well.) If we write R_c^{*1} for the opportunity-standardized crime rate in t_1, and R_c^{*2} for the similar rate in t_2, then $R_c^{*1} < R_c^{*2}$ if $(C_2C_1) < (O_2O_1)$; that is, the rate will decrease if the number of crimes fails to increase as rapidly as the number of opportunities. It is also evident that if $O_2 = O_1$, then

$$\frac{R_c^*{}_2 - R_c^*{}_1}{R_c^*{}_1} = \frac{R_c^*{}_2 - R_c^*{}_1}{R_c^*{}_1} \tag{4}$$

that is, if opportunities remain unchanged between t_2 and t_1, they can be ignored in the calculation of crime rates. This is also true, of course, for population. An opportunity-standardized crime rate thus uses in its denominator the elements of both crime rates and victimization rates, as these are usually calculated.

The effects of taking into account opportunities can be illustrated by considering the data in Table 2, most of which are taken from Wilkins (1964:55). Column 2 of this table gives the numbers of thefts from motor cars reported to the police in England and Wales during the years 1938-61 (excluding 1939); column 3 gives the numbers of cars registered in England and Wales during those same years. Column 4 contains the resident population of England and Wales in 1938-61 (General Register Office, 1964:Table 2). Column 5 is calculated from columns 2, 3, and 4, and represents the numbers of thefts per 100,000 population per 1,000 cars available to steal; for convenience I have indexed this series so that 1938 = 100. The column headed "X" is also taken from Wilkins (1964:55); it is in fact the numbers of thefts from shops and stalls reported to the police in England and Wales in 1938-61. It is not clear to me why these figures were given by Wilkins, since he nowhere discusses them in the text of his book; they are, however, convenient for illustrative purposes. We may consider them to represent some candidate explanatory variable, e.g., average family income in pounds sterling.

Wilkins's own discussion of the data in columns 2 and 3 of this table raises a number of questions. For one thing, he considers the numbers of thefts *from* motor cars, rather than the numbers of thefts *of* motor cars;[5] for another thing, he uses the *numbers of thefts reported to the police*, rather than the (reported) *theft rate* per 100,000 persons, though as column 4 shows the resident population of England and Wales increased about 12 percent over the years 1938-61. Inspection of the table will show that the numbers of thefts reported rose by about 350 percent, in the same years; the theft rate per 100,000 persons of course increased by somewhat less than this. The number of cars available to steal—or to steal from—rose by over 200 percent, however; the result is that the number of thefts from cars, standardized both for population and for cars, rose by only 29 percent.

[5]Wilkins has since informed me that reliable data for thefts *of* cars were not available for the period in question.

The effects of standardizing for opportunities can be quite striking. They may be quickly illustrated by comparing correlations between thefts from cars, and the same thefts standardized for both population and cars available (i.e., columns 2 and 5 of Table 2), and the (fictional) explanatory variable in the column headed "X". For the numbers of thefts and X, $r = +.93$; for the number of thefts standardized for population and cars registered, and X, $r = -.21$. Thus what at first sight seems a very strong positive correlation (+.93) is turned into a moderate *negative* correlation (–.21), when changes in the population and the stock of cars available are taken into consideration.

The effects of population and opportunities may vary, of course, and may themselves be of interest. To illustrate the estimation of those effects, we may regress the numbers of thefts in Table 2 on population and the numbers of cars. It might be argued that the effects of population and opportunities were additive; intuitively, however, it seems more reasonable to assume that they are multiplicative, so that

$$T' = kC^{x_1}P^{x_2} \qquad (5)$$

where T' = estimated number of thefts,
C = the stock of cars,
P = the population.

This can be rewritten, after taking logarithms on both sides, as

$$\ln T' = a + x_1 \ln C + x_2 \ln P, \qquad (5a)$$

which can be estimated using ordinary least squares. A somewhat similar approach was recently used by Felson and Cohen (1977) in their analysis of burglary rates in the United States in the period 1950-72. Felson and Cohen utilize what they call a "routine activity" approach to crime, based on the human ecology theories of Amos Hawley (1950); they show that a substantial proportion of the variation in burglary rates can be accounted for (corrected $R^2 = .986$) by a multiplicative combination of three variables, namely the percent of the population aged 15–24; the proportion of "primary individual" households; and an ingenious measure of the "inertia" of property targets, *viz.* the weight of the lightest television set advertised in current Sears & Roebuck catalogues. (See also Cohen and Felson 1978.)

Similarly, using the data in Table 2, the numbers of thefts from cars can quite accurately be forecast from population and the numbers of cars registered ($R^2 = .93$). As I shall argue later, however, it is often the case that neither the population at risk, nor the stock of opportunities, is of much theoretical interest; what is wanted, then, is to exclude their effects, and examine variations in crime that are "left over" after that exclusion. Thus, for example, after estimating the numbers of thefts from cars T', using equation (3a), we may calculate the residual thefts $T^* = T - T'$, and use those residuals as the dependent variable in subsequent analyses. (The correlation between those residuals, and the fictitious explanatory variable in column X of Table 2, is +.20.) It is to be noted that excluding the effects of population and opportunities in this way does not have the same effect as controlling for changes in opportunities (in this case, changes in the numbers of cars) by partial correlation. A partial correlation $r_{xy \cdot z}$ is of course a correlation between the residuals of the regression of x on z and those of the regression of y on z; the effect of the control variable z is thus removed from both x and y. In the present case, however, we are removing the effects of population and opportunities from the dependent variable only, in order to examine the residual variation in relation to some candidate independent variable or variables. (See, for a further discussion, Mosteller and Tukey 1977:269-71.) Some cases in which partial correlation and kindred techniques may be appropriate, in the analysis of criminal opportunities, will be discussed in a later section of this paper.

In a more recent paper, Gould (1969) presented data similar to those given by Wilkins, for thefts of cars and numbers of cars registered, in the United States in the years 1933–65. Gould also compared year-end amounts of cash on hand in banks with the numbers of bank robberies and burglaries, during the years 1921–65. In each case Gould, like Wilkins, compared the *numbers* of thefts with his measures of opportunity, and took no account of changes in the population able to commit the two types of theft; in view of the fact that the U.S. resident population increased by about 50 percent between 1933 and 1965, this is no small omission. (It might of course be appropriate to take into account changes in the age *structure* of the population, or to exclude the effect of some subgroup thought to be especially likely to be involved in the particular kind of crime in question—as Felson and Cohen (1977) did in their analysis of burglary.) Gould does not present, in his 1969 paper, data which would permit the calculation of doubly standardized rates of theft, or of residual theft after the exclusion of population and opportunity effects. But from his graphs of the two time series (Gould 1969:53) it is clear that both car theft and bank robbery/burglary rates declined, relative to opportunities, up to about the middle 1940's; and that, even without allowing for increases in the population, rates relative to opportunities remained virtually unchanged from the mid-1940's to the mid-1960's. (The correlation between numbers of car thefts and cars registered is +.97, for the years 1950–65; for cash in banks and bank robberies/burglaries, the correlation is +.98 for the period 1944–65.)

Still more recently, Mayhew et al. (1976) analyzed car thefts and numbers of cars registered, in London in the years 1961-74. They noted that the parallel trend noted earlier by Wilkins (for thefts *from* cars, and cars registered) did not continue, at least in London, after 1961. Mayhew et al. were mainly concerned with the effects of a preventive measure (steering column locks) in reducing opportunities for theft. Since 1970, steering column locks have been required on all cars manufactured in or imported into England; the result has clearly been a decline since that date in the stock of what Mayhew et al. call "stealable" cars. Yet they found that the number of thefts of cars has continued to rise sharply in London, especially since 1970. If theft rates since that date were standardized for (declining) opportunities in terms of the numbers of unprotected cars, the increase in theft rates would be much greater than is suggested by the data which Mayhew et al. present.

Defining and measuring opportunities for crime

What constitutes an opportunity for the commission of a crime naturally depends on the type of crime in question, and satisfactory definition—both conceptual and operational—can in some cases be very difficult. One problem, to which I see no general solution, concerns the vagueness of the borderline between difficulty and impossibility, reflected by the Mayhew et al. (1976) study just discussed. Steering-column locks may deprive amateur car thieves of the opportunity to ply their trade; but they leave the opportunities open to professional car thieves (presumably) more or less unchanged. Similarly, the design of most bank vaults deprives the majority of the population of the opportunity to steal the valuables in those vaults; professional bank burglars, equipped with explosives or thermic lances, still have their opportunities available to them. In general, if preventive measures are themselves of theoretical interest—if, for example, it is intended to assess their effects on the crime rate—then those effects should be kept separate, and not excluded from crime rates in the way suggested in the preceding section. To the extent that one's interest is in other criminological factors, however, the opportunity-reducing effects of preventive or other social-control factors may best be removed from the crime rates in question, so the residual crime rates can be examined by themselves.

There still remain many difficulties of definition. In the case of crimes of violence, opportunities are presumably created by contacts or interactions between persons. As is well-known, in a population of N persons, the maximum possible number of two-person contacts which can take place equals $N(N-1)/2$; this was in fact the base used by Boggs (1965:900) in calculating her "crime-specific" rates of homicide and aggravated assault. However, this maximum number is of potential contacts of any kind (assuming—what is not off-hand clear—that the notion of a "contact" can be defined with reasonable precision). To be even approximately satisfactory, standardization for opportunities for crimes of violence would need to take into account the nature and duration (and possibly the frequency) of interactions between different types of persons, e.g., across different types of communities or in given interpersonal relationships. Thus some years ago Svalastoga (1962) analyzed a small sample of homicide cases in Denmark; he found, as have most other researchers on homicide, that the majority involved family members and that strangers accounted for only 12 percent of the cases studied. On the basis of a small survey of students, plus some admitted guesswork, Svalastoga estimated that a Danish person might have contacts with relatives, acquaintances, and strangers in the ratios $4 \cdot 10$: $4 \cdot 10^3$: $4 \cdot 10^6$; on this basis, he calculated that the probability of being killed by an acquaintance was some 3,000 times greater than the probability of being killed by a stranger and that the probability of being killed by a family member was some 600,000 times greater. These "probabilities" assume, of course, that the numbers and types of contacts are on average the same *within* each of these three groups, which is improbable to say the least. Nonetheless, the general logic of this approach seems to me correct; for the purpose of explaining the social, spatial, or temporal distribution of violent crime, as well as for assessing the risk of it, some account needs to be taken of the distribution of opportunities (i.e., interactions between persons) which are a logically necessary condition of such victimization.

In the case of crimes against property, the choice of an adequate base for calculating an opportunity-standardized rate can also be problematic. One approach is simply to use the stock of stealable goods; this is in fact what was done by Wilkins, Gould, and Mayhew et al. in their studies of theft from, and of, cars. But for other types of theft, the matter is less clear. Thus Gould (1969), in analyzing bank robberies and burglaries, used data on amounts of cash and coin in banks. It might be argued, however, that the *number of banking offices* is a better measure of opportunity than the amounts of cash which are contained in those banking offices. (Data from the *Statistical Abstract* [U.S. Department of Commerce, 1970] show that the amount of cash and coin in banks increased about five times in the period studied by Gould; the number of banking offices increased by only about 70 percent.) Similarly, should one use the number of supermarkets and/or department stores as a base for shoplifting, or the value of those stores' inventories? The answer would seem to be that it depends on the purpose for which rates are being calculated. If the objective is the assessment of risk, then the number of institutions (stores, banks, etc.) would usually be more appropriate; if the objective is the explanation of observed patterns of theft, then the stocks of available goods might be preferred.

In the case of thefts from individuals and/or households, the choice of an appropriate base is even more complicated. In some places, estimates are available for the stocks of consumer goods owned by individuals and/or unincorporated businesses (see, e.g., Roe 1971:70-71), where some estimated values for the United Kingdom in the years 1955-66 are given; these data are shown at written-down replacement cost, though if price-index changes are taken into account an estimate of physical stocks of goods could in principle be derived from them). In this country, data are available from the *Statistical Abstract* and from a variety of trade publications, such as *Merchandising Week*, on most types of durable consumer goods. Typically, these data are for production, shipments, or sales of such goods, though estimates of stocks can be derived from them if assumptions are made about average life (or average "stealable" life). Where figures for estimated stocks of such goods are available, they almost invariably disclose massive increases over the past three decades, usually far greater than the increases in burglaries and larcenies recorded in the *Uniform Crime Reports*. Thus, according to estimates based on market research by a television network (NBC 1975), only 9 percent of all American households had a television set in 1950; by 1974 the figure was over 94 percent, with over two-thirds of those households having color television and about two-thirds having more than one set. (Similarly, trade sources estimate that the total number of radio sets in use in the United States increased by nearly 3½ times in the period 1950–74.) Within the past few years, there has evidently been an even more rapid increase in ownership of such things as stereo equipment, tape recorders and cassette players, hand-held calculators, and CB radios. The result has been to increase substantially the quantity of personal disposable property available to be stolen, and thus the opportunities for theft.[6] In the case of television sets, for example, the figures just quoted mean that in 1950 a burglar or thief had less than one chance in ten of finding a television set in an American household chosen at random; by 1974 he would have had difficulty in *not* finding one, and in two houses out of every three could have had a choice of sets (or the chance of a color set) to steal. Over the same period—*and for exactly the same reason*—the chance of any particular television set's being stolen has almost certainly decreased sharply.

[6] According to a recent newspaper account, moving companies estimate that the average household move involved 1,000 cubic feet of goods 15 years ago; now the figure is about 2,000 feet. Not surprisingly, suburbanites are said to be "worse than city dwellers" in this respect. See "In age of accumulation, moving is a trial," *New York Times*, March 19, 1978, p. R1.

Until better data are available on the amounts and types of property stolen, it does not seem worthwhile to estimate changes in opportunities for theft with greater precision or detail. Unfortunately, little such information is now available. Since 1974, data have been collected (though not published) in supplementary returns from police forces under the Uniform Crime Reporting program; these returns, which have recently been made more detailed, apply only to thefts reported to the police, and nothing is now known about their validity. The same is true of the similar data now being collected in the National Crime Surveys. The Crime Incident forms (NCS-2 and NCS-4) used in these surveys contain questions pertaining to the value (calculated in several different ways) of stolen or damaged property; but they are now coded so as to distinguish only between thefts of cash, motor vehicles and accessories, and "other" property. Given more data on the types and amounts of property stolen, it would, in principle, be possible to estimate the stocks of property from which those thefts occurred; if this were done, it would be possible to estimate rates of theft relative to opportunities, either cross-sectionally or over time.

Opportunities and criminological explanations

Thus far, this paper has been concerned with the effects of variations in opportunities for crime—for example, changes in the stock of personal disposable property, in the case of theft—on the interpretation of the crime rate: it has been argued that, given the purposes for which we commonly measure crime, it is appropriate to standardize crime rates for opportunities. I can see no argument against this kind of standardization, which would not apply with equal force to standardization for changes in the population available to commit crimes.

It remains to be considered, in conclusion, whether opportunity factors also can or should figure as separate independent variables in an explanation of criminal behavior or variations in crime rates. Recent papers by Gould and his associates (Gould 1969, 1970; Mansfield, Gould, and Namenwirth 1974) appear to treat changes in the stocks of one type of property—cars—in precisely this way. Thus Gould (1969:54) writes that "the availability of property influences the amount of theft against it," and he refers to this as a "causal sequence." He goes on to speculate that "property crime is not only related to the availability of property, but ... this relationship is itself structured by the relative scarcity or abundance of the property being stolen." (Gould 1969:56) He also notes that changes in patterns of car theft, and in the availability of stealable cars, parallel an apparent change in the population of car thieves, who (according to arrest data) are now much more likely to be juvenile or adolescent "joyriders" than they were in the 1930's or 1940's. (Gould also notes that similar changes appear to have taken place among bank robbers, with "professional" robbers having largely been replaced by inept amateurs.)

Gould suggests, then, that (at least so far as car theft is concerned) the period from the early 1930's to the early 1940's was "a period of economic scarcity" (Gould 1969:55) in which car theft was mainly an activity of "professional" thieves; and that the years after about 1942 were "marked by abundance," and the emergence of juvenile "nonprofessional" thieves. Inspection of the (graphed) data which Gould presents, however, suggests a rather different picture. From his graph of car registrations and car thefts, it appears that:

- Car registrations *rose* in the years 1933-41, though the increase would admittedly be less if increases in population were taken into account.
- Cars were possibly relatively scarce in the years 1941-45—there was a decline of about 16 percent, from about 35 million to about 30 million.
- Since 1945, the stock of cars registered rose steadily, but the numbers of car thefts fell, until about 1950.
- After 1950, the increase in car thefts roughly paralleled the increase in cars registered.

It is evident that none of these changes in stocks of stealable cars, and of car thefts, in any way necessitates the shift which appears to have taken place, from "professional" to "amateur" thieves; this change is quite independent. Moreover, the relation between cash and coin in banks, and numbers of bank robberies, is (as Gould [1969:56] puts it) "somewhat different"; to the extent that there has been a somewhat similar shift in the population of robbers, this is not paralleled by changes in the amounts of cash and coin available to steal. Using the concepts outlined earlier in this paper, we could in fact describe the situation relating to car theft in the following ways:

- In the years 1933-41, the numbers of cars registered rose fairly steadily, while the numbers of thefts of cars declined; the rate R_c^* of car thefts standardized for opportunities fell sharply. Quite probably this could have been because, though it was successively easier to steal a car (there were more of them), it was also easier to obtain one legitimately.
- In the years 1941-45, when no new cars were manufactured and the stock of stealable cars declined, the numbers of thefts (and the theft rate R_c^*) rose; because cars became relatively scarce, the opportunities for obtaining them legitimately also decreased.
- In the years 1945-50, the numbers of cars (and thus of opportunities for car theft) rose again; car thefts fell, so the car theft rate R_c^* fell even more sharply; again this could have been because cars were more easy to obtain legitimately.
- Finally, in the years 1951-65, the numbers of cars (and of opportunities for car theft) rose steadily; so did the numbers of thefts, so that the car theft rate R_c^* remained about unchanged.

These temporal patterns are evidently compatible with many different combinations of professional and amateur car thieves and different participation rates of each.

In a later paper, Mansfield, Gould, and Namenwirth (1974) expand on Gould's earlier work (and incidentally standardize both car ownership and car theft for changes in population, as Gould had not). They attempt, using data from four countries, to test a model according to which thefts of cars (and, by implication, other stealable property) are determined by the interaction of "professional" and "amateur" demand for stolen vehicles, and the supply of available vehicles. The authors admit that they are only able to carry out a partial test of their model; in particular, they have no data on the relative magnitude of "professional" and "amateur" demand for stolen vehicles, nor on the shapes of the respective demand curves. Their paper illustrates two ways in which an opportunity factor—e.g., the supply of stealable goods—may be incorporated into an explanatory theory. But it is important to note that it is not opportunities *as such*—in the sense described in this paper—that figure in such a theory. Instead, it is the relative scarcity or abundance of goods, which may affect participation in theft in, broadly speaking, one of two ways: (1) by affecting motivation to steal (e.g., by making it easier to obtain goods legitimately or conversely by increasing relative deprivation) and (2) by leading to changes in social control measures (in the broadest sense of that term, including measures for the protection of stealable property). Mansfield, Gould, and Namenwirth do mention both of these, but they make no attempt to operationalize or measure either one. The first of these explanations would seem to require (for example) some evidence about the *distribution* of stealable property, as well as the *quantity* of it; it might also require consideration of the *value* of that property since (under certain conditions) an increase in supply relative to demand may lead to a decrease in price. Thus, for example, if more cars are available, opportunities for car theft will (*ceteris paribus*) increase; but it should also be easier to obtain cars legitimately. In this case we should expect a *decrease* in R_c^*, such as that which occurred in the years 1933-41 and 1945-50. The second line of explanation might involve showing, for example, that when property is relatively abundant people are less likely to protect it against theft, or to report thefts to the police. But an increase in the stock of goods—and thus in opportunities to acquire those goods legitimately—can occur together with an increase in protective measures; an example would be a law requiring steering-column locks to be fitted to all cars. (As Mayhew et al. [1976] point out, such a law was passed in Germany in 1963.) In this case the number of thefts would be expected to fall, for two distinct reasons: first because demand for stolen cars fell, and secondly because of a decrease in opportunities.

One case in which it may be useful to treat changes in opportunities independently is the case in which they function as an intervening variable, helping to explain an observed relationship between crime or victimization and some other variable. Thus, several victimization surveys have found a negative association between age and victimization, especially for violent crimes such as assault (Sparks, Genn, and Dodd 1977:chap. 4; Aromaa 1971; Hindelang 1976:111-14). A possible explanation for this finding is that older people tend to go out less often, especially at night; they are thus less at risk of (certain sorts of) victimization. (What is opportunity from the offender's point of view, of course, is risk from the victim's.) If this factor is taken into account, the zero-order association between age and victimization may disappear, or at least be reduced. In our London victimization survey, for example, the zero-order γ between age and victimization was -.42; between age and the number of nights per week the respondent went out, $\gamma = -.18$; between nights out and victimization, $\gamma = +.20$. The partial γ between age and victimization, controlling for nights out, was reduced to -.29. A reasonable interpetation of these findings is that a part of the older respondents' lower victimization rate was due simply to the fact that they were less often at risk. A similar analysis is possible for some of the National Crime Survey data, since the NCS-6 "attitude" questionnaire administered to half of the respondents in the city-level surveys contains a question (Q.8a) asking "How often do you go out in the evening for entertainment?"; so far as I know, however, these data have not been analyzed from this point of view.

Another case in which variations in opportunity need to be considered separately occurs when those variations are (hypothesized to be) consequences of protective or social-control measures in the broadest sense of that term. In the case of steering-column locks on cars, for example, it would seem best to analyze thefts of cars with such locks separately from thefts of cars without them, taking into account the stocks of cars of each type; the hypothesis that locks were effective in preventing theft would predict that thefts of cars with locks would decrease (or increase less rapidly) relative to thefts of cars without them. Intensive (and highly visible) police patrols, anti-shoplifting devices, apartment-house security systems, burglar alarms, exact-fare buses, and the like, may all be assumed to reduce *some* potential criminals' opportunities for crime, by making the successful accomplishment of crimes more difficult for the average person (or potential criminal). To the extent that our theoretical and/or practical interests are focussed on this kind of effect, we would *not* want to remove the effects of variations in opportunity from the dependent variable (i.e., crime rates), but to treat them independently. If our interests lie elsewhere, however, we might well wish to remove those effects, so as to see better the effects of other variables on variations in crime.

Inevitably, there will be borderline cases. The "inertia" measure used by Cohen and Felson (1977), discussed earlier, is an example. The fact that many durable consumer goods have become smaller and lighter (as measured by the weight of the lightest television set advertised in current Sears and Roebuck catalogues) can certainly be interpreted as increasing (the average man's) opportunities to steal. Are we interested in the effect of *that* change on rates of theft? Or are we interested in variations in theft of durable consumer goods, *given* that change in opportunities? Each of these questions requires a different measure of theft—the former excluding, the latter including, the effects of variation due to the opportunity variable.

Suppose, furthermore, that the use of transistors, printed circuits, etc., has made some kinds of goods smaller and easier to steal (under average circumstances). Suppose, moreover, that (as is almost certainly true) those same technological factors have made those goods much cheaper (and thus generally easier to acquire legitimately). Suppose, furthermore, that—as is also certainly true—the stocks of such goods available to steal have increased sharply at the same time. In those circumstances, to treat prices, stocks, and "inertia" as separate variables in an explanation of theft rates will be likely to lead to severe problems of multicollinearity; but these will still be present, even if any one or two of the three (say, stocks of goods and "inertia") are removed, along with changes in the population, in the calculation of crime rates.

Conclusions

It seems that there are some instances in which variations in opportunities to commit crime may be substantively important in the explanation of variations in crime rates, and so should be treated independently rather than being "netted out" of the crime rates themselves. In general, however, it seems to me that variations in opportunities for crime are likely to be rather obvious and uninteresting where the explanation of crime is concerned. It is seldom useful to point out that cars could not be stolen before cars were invented; and it is not, in general, illuminating to point out that a man who never goes out of his house will never get assaulted or robbed in the street.

But though they may often be relatively trivial in themselves, variations in opportunity may nonetheless often obscure the effects of more important theoretical variables. To avoid this, the procedure of standardizing crime rates for opportunities, as outlined in this paper, seems to me appropriate. This procedure is in fact analogous to the "method of residues" proposed some years ago by Coleman (1964: chap. 15). Coleman noted that a great deal of effort had at one time gone into finding a "law of social gravity" to the effect that, say, the amount of travel or other interaction between two cities is directly proportional to the product of their populations, and inversely proportional to some function of the distance between them. Such a "law" has the unfortunate defect that it often does not fit the observed data on intercity travel very well. But it also has the even more serious defect that, where it does fit, it is utterly uninteresting. By standardizing rates of travel for populations and distance, Coleman suggested, one could calculate for any pair of cities a "residue" which would be the difference between observed travel and that expected on the basis of population and distance alone; examination of these residues might then reveal more interesting effects which would otherwise be obscured.

The same approach may often be useful in relation to crime and victimization. The *number* of homicides in New York is greater than the *number* of homicides in, say, New Orleans; but the homicide *rate* standardized for population is higher in New Orleans. The *number* of car thefts in the United States in 1965 is greater than the number in 1945; but the car theft *rate* standardized for both population and the stock of stealable cars is smaller. Only by removing the sociologically trivial effects of population and opportunities can more interesting and important effects be seen.

Acknowledgments

Many of the ideas in this paper were developed at the SSRC Workshop in 1975; earlier versions of the paper were presented at the annual meeting of the American Society of Criminology in November 1977, and at a Workshop on Quantitative Analysis of Crime and Criminal Justice, sponsored by the Institute for Social Research, University of Michigan, in August 1978. I am grateful to many people for their comments on those earlier versions: in particular to Don Gottfredson, Andrew von Hirsch, David Seidman, Howard Wainer, Colin Loftin, and Simon Singer. None of them is responsible for residual errors.

References

Aromaa, Kauko (1971)
Arkipaivan Vakivaltaa Suomessa [Everyday violence in Finland]. Helsinki: Kriminolginen Tutkimoslaitos.

Aromaa, Kauko (1973)
"Victimization to violence: Some results of a Finnish survey." *International Journal of Criminology and Penology* 1:245.

Aromaa, Kauko (1974)
The replication of a survey on victimization to violence. Helsinki: Institute of Criminology.

Avi-Itzhak, B., and R. Shinnar (1973)
"Quantitative models of crime control." *Journal of Criminal Justice* 1(3):185-217.

Bauer, Raymond A. (ed.) (1966)
Social indicators. Cambridge, Mass.: M.I.T. Press.

Bentham, Jeremy (1778)
"Observations of the Hard Labour Bill." *Works* (J. Bowring, ed.), vol. 4, 28-41.

Blumstein, Alfred, Jacqueline Cohen, and Daniel Nagin (eds.) (1978)
Deterrence and incapacitation: Estimating the effects of criminal sanctions on crime rates. Report of the Panel on Deterrent and Incapacitative Effects, National Research Council. Washington, D.C.: National Academy of Sciences.

Boggs, Sarah L. (1965)
"Urban crime patterns." *American Sociological Review*, 30:899-908.

Carr-Hill, Roy A. (1971)
The violent offender: Illusion or reality? Oxford University Penal Research Unit, Occasional Paper No. 1. Oxford: Basil Blackwell.

Cloward, Richard, and Lloyd Ohlin (1960)
Delinquency and opportunity. New York: Free Press.

Cohen, Jacqueline (1978)
"The incapacitative effects of imprisonment: A critical review of the literature." In Alfred Blumstein, Jacqueline Cohen, and Daniel Nagin (eds.), *Deterrence and incapacitation: Estimating the effects of criminal sanctions on crime rates.* Report of the Panel on Deterrent and Incapacitative Effects, National Research Council, Washington, D.C.: National Academy of Sciences.

Cohen, Lawrence E., and Marcus Felson (1978)
"Social change and crime rate trends: A routine activity approach." Working Papers in Applied Social Statistics. Department of Sociology, University of Illinois at Urbana-Champaign.

Coleman, James S. (1964)
Introduction to mathematical sociology. New York: Free Press.

Felson, Marcus, and Lawrence E. Cohen (1977)
"Criminal acts and community structure: A routine activity approach." Working Papers in Applied Social Statistics. Department of Sociology, University of Illinois at Urbana-Champaign, April 13, 1977 (mimeo.).

General Register Office (1964)
Census, 1961, England and Wales. Vol. 1: Age, marital condition, and general tables. London: Her Majesty's Stationery Office.

Gould, Leroy C. (1969)
"The changing structure of property crime in an affluent society." *Social Forces* 48:50-59.

Gould, Leroy C. (1970)
"Crime and its impact in an affluent society." In Jack D. Douglas (ed.), *Crime and justice in American society.* Indianapolis: Bobbs-Merrill.

Green, Gary (1978)
"Measuring the incapacitative effectiveness of fixed punishment." In James A. Cramer (ed.), *Preventing crime,* Sage Criminal Justice System Annuals, vol. 10, New York: Sage Publications.

Greenberg, David (1975)
"The incapacitative effect of imprisonment: Some estimates." *Law and Society Review* 9:541-80.

Greenwood, M., and G. Udny Yule (1920)
"An inquiry into the nature of frequency distributions representative of multiple happenings with particular reference to the occurrence of multiple attacks of disease or repeated accidents." *Journal of the Royal Statistical Society* 83-255.

Hawley, Amos (1950)
Human ecology: A theory of community structure. New York: Ronald.

Hindelang, Michael J. (1976)
Criminal victimization in eight American cities: A descriptive analysis of common theft and assault. Cambridge, Mass.: Ballinger Publishing Co.

Horning, Donald M. (1964)
Blue-collar theft. Ph.D. dissertation, Indiana University.

Mansfield, Roger, Leroy C. Gould, and J. Zvi Namenwirth (1976)
"A socioeconomic model for the prediction of societal rates of property theft." *Social Forces* 52:462-72.

Mayhew, Bruce H., and Roger L. Levinger (1976)
"Size and density of interaction in human aggregates." *American Journal of Sociology* 82:86-110.

Mayhew, P., R. V. G. Clarke, A. Sturman, and J. M. Hough (1976)
Crime as opportunity. Home Office Research Study No. 34. London: Her Majesty's Stationery Office.

Merton, Robert K. (1938)
"Social structure and anomie." In *Social theory and structure* (rev. ed. 1957), chap. 3, New York: Free Press.

Mosteller, Frederick, and John W. Tukey (1977)
Data analysis and regression. Reading, Mass.: Addison-Wesley.

National Broadcasting Corporation (1975)
"TV households, sets and per cent saturation." NBC Research Estimates. *TV Factbook,* No. 44 (1974-75):69.

National Criminal Justice Information and Statistics Service, Law Enforcement Assistance Administration (1976)
Criminal victimization surveys in Chicago, Detroit, Los Angeles, New York, and Philadelphia: A comparison of 1972 and 1974 findings. National Crime Survey report No. SD-NCS-C-6. Washington, D.C.: U.S. Government Printing Office.

Penick, Bettye K. Eidson and Maurice E. B. Owens III (eds.) (1976)
Surveying crime. Final report of the Panel for the Evaluation of Crime Surveys, Committee on National Statistics, Assembly of Mathematical and Physical Sciences, National Research Council. Washington, D.C.: National Academy of Sciences.

Reiss, Albert J., Jr. (1967)
Studies in crime and law enforcement in major metropolitan areas. Vol. 1, section I: Measurement of the nature and amount of crime. President's Commission on Law Enforcement and Administration of Justice, Field Surveys III. Washington, D.C.: U.S. Government Printing Office.

Reynolds, Paul Davidson, et al. (1973)
Victimization in a metropolitan region: Comparison of a central city area and a suburban community. Minneapolis: Minnesota Center for Sociological Research (mimeo.).

Roe, A. R. (1971)
The financial interdependence of the economy. London: Chapman and Hall.

Shinnar, Reuel, and Benjamin Avi-Itzhak (1973)
"Quantitative models in crime control." *Journal of Criminal Justice* 1:7.

Shinnar, Reuel (1975)
"The incapacitative function of prison internment: A quantitative approach." Department of Chemical Engineering, City University of New York (mimeo.).

Simon, Herbert A. (1957)
Models of man: Social and rational. New York: John Wiley & Sons.

Sparks, Richard F., Hazel G. Genn, and David J. Dodd (1977)
Surveying victims: A study of the measurement of criminal victimization, perceptions of crime, and attitude to criminal justice. London: John Wiley & Sons, Ltd.

Taylor, Ian, Paul Walton, and Jock Young (1975)
Critical criminology. London: Routledge and Kegan Paul.

U.S. Department of Commerce, Bureau of the Census (1970)
Statistical abstract of the United States, 1970. Washington, D.C.: U.S. Government Printing Office.

Webb, Stephen D. (1972)
"Crime and the division of labor: Testing a Durkheimian model." *American Journal of Sociology* 78:643-56.

Svalastoga, Kaare (1956)
"Homicide and social contact in Denmark." *American Journal of Sociology* 62:37-41.

Wilkins, Leslie T. (1964)
Social deviance: Social policy, action, and research. London: Tavistock Publications.

Wolf, Preben, and Ragnar Hauge (1975)
"Criminal violence in three Scandinavian countries." In *Scandinavian studies in criminology,* vol. 5, London: Tavistock Publications.

Yule, G. Udny (1924)
"A mathematical theory of evolution, based on the conclusions of Dr. J. C. Wills, F.R.S." *Philosophical Trans.* B, 213:31.

Notes on measurement by crime victimization surveys

ALBERT D. BIDERMAN
Bureau of Social Science Research
Washington, D.C.

The first part of this paper describes shortcomings in victimization surveys in providing data for the measurement of the prevalence (rather than the incidence) of crime. The second part discusses the immunizing effects of victimization experiences to future experiences, and how the concept of immunization may be applied to the assessment of risk of victimization.

The significance of measurements of events and conditions

Incidents and conditions. Among many failures of the National Crime Survey (NCS) to derive the full potential of the victimization survey method that stem from the carryover into its conception and design of Uniform Crime Reports (UCR) rationales is its emphasis on crimes as incidents occurring at points in time. The survey method is more effective as a tool for gaining respondents' reports of their current states than of recalled events in even very recent and brief periods of their life histories. To a considerable degree, the significance of crime from a disaggregated, individual-oriented perspective (as opposed to a collective perspective) resides in durable consequences of victimization, rather than events of ephemeral consequences for the individual. While we must work toward developing elaborate procedures for stimulating recall in order to provide a survey instrument that will not grossly undercount even UCR Index crimes (conventionally labeled "serious"), we are, in the design of the NCS, neglecting almost totally gaining information on criminal victimization that is more readily approached as a *condition* than as an *incident.* We are also neglecting the durable consequences, that is, the changes in condition, produced by point-in-time incidents. In other words, NCS pursues an incidence, rather than prevalence, rationale. Among the kinds of victimization that may be conceived and measured in prevalence rather than in incidence terms are various forms of continuing terrorization and extortion, for example, the worker who is kept in line by union or company "goons," school children who must regularly yield their lunch money to fellow student toughs, the merchant subject to a shakedown racket, the prostitute terrorized by her pimp, or the spouse or sexual partner kept from separating from a hated relationship by fear of violence.

To some degree, victimization surveys yield information about these kinds of situations through tabulations of what are called "series victimizations." In the NCS these are defined as three or more similar incidents of victimization mentioned by a respondent, but which, because of frequency and/or similarity, the respondent cannot individually date in time or differentiate descriptively from one another. (Were each incident of a series counted as one, they would form a very appreciable portion of all personal victimizations in the National Crime Survey; I would hazard a minimum of 20 percent of the total number of incidents.) Thus, the terrorized spouse may be identifiable in a victimization survey through repeated incidents of spouse beating, and the terrorized school child by repeated incidents of robbery.

It is not necessary for a durable condition of victimization to exist, however, for there to be many incidents, each qualifying under the definition of a criminal victimization used by the survey. To make a threat credible to the victim and to continue a state of terrorization, the terrorist must neither continually repeat his threat nor demonstrate his willingness to carry it out by actually inflicting violence.

Albert J. Reiss, Jr. (1972) also illustrates a somewhat different type of continuing victimization by the case of the tenant inhabiting a dwelling affected by building code violations. The "crime" of the landlord in this instance is similarly a state, rather than incident form of crime, that continues in duration through time, so long as the condition of the structure remains uncorrected. Bigamy has the same continuing character and involves a victimization where the bigamist keeps a partner ignorant of the other. Such victimization states are subject to incidence measurement with regard to points of entering or leaving the state, but prevalence measures are applicable to the observation of such victimization in a population.

The series form of incident may also be an indicator of a condition of victim proneness, that is, a person vulnerable to offenses of a similar character by different offenders on frequent occasions. Among such conditions mentioned in victimization survey results are the shopkeeper in a high-crime area or the resident of a highly burglary-prone dwelling unit or the person who is forced to park his automobile where it is regularly subject to vandalization.

While the NCS utilized the panel technique primarily to institute a control on "telescoping," the value of the panel feature probably will reside more in the elucidation of those forms of victimization best characterized in terms of prevalence, rather than incidence, measures. As indicated earlier, because inquiry can be made of current conditions of victimization, recall problems are avoided. Conditions are more accessible to survey detection than past events. In addition, their very duration or frequency in the individual life space makes the former more significant for individuals than are the many incidents with highly ephemeral consequences for individuals with which victimization surveys have been preoccupied. Finally, as Reiss has pointed out, such victimizations usually present a much higher potential for effective system intervention than is the case with point-in-time incidents.

Another facet of victimization subject to a prevalence, rather than incidence measurement rationale, is the durable consequence of a point-in-time crime event. This is the rationale I employed in my feasibility study of injury-produced handicaps and pain prevalence among a randomly sampled population. (Biderman 1975)

Ideal control and compensation models. Elsewhere, I have discussed ideal models of society which can be posited by taking current ideas of crime control and victim compensation to their logical extremes. The "perfect" condition from which the victimization survey can be seen as measuring departures is one in which social life is so regulated that no one ever harms any other. To perfect a social order that is less-than-perfect in this regard, a perfect compensation system can be posited; that is, one in which any victim of harm receives compensation (from the doer of harm or from the state) that makes him "whole" again.

Our data problems can perhaps be understood in this way: we are able to construct intellectual models of social welfare that are much more integrated, comprehensive, and logical than are any institutions or social welfare we can actually bring into being. If we had (or could possibly have) the ideal welfare state with a perfect system of comprehensive social insurance, all harms befalling individuals would be recorded and measured as part of the process of social compensation (or state-enforced recompense by offenders). (Under these circumstances, indicators from the data of compensation would, of course, no longer be of individual concern, since the costs would be socialized by perfect compensation. The social costs also would greatly exceed the aggregate value of individual harms because of the expense of operating the compensation system.)

In current (as well as in conceivable) reality, we deal with limited systems of compensation. Theoretically, victimization surveys would be needed primarily to measure the gaps, in scope and degree, between the actual occurrence of harms and the reach of the system of compensation. In actuality, however, the systems of social compensation are terribly poor at generating good, comprehensive statistics. In the United States, the factors of decentralization and the mixture of public and private systems make social and casualty insurance data close to useless for the purposes of social indicators of the "output" type. These defects of agency data place an additional burden on the victimization survey—that is, the survey is the only means currently available for gaining data on harms dealt with by systems of compensation, even though, with only modest perfections of the statistical operations of such institutions, such phenomena should be revealed with far greater scope and accuracy by data from transactions of the systems themselves.

Two aspects of our earlier discussion bear on victimization surveys as a source of guidance for victim compensation systems. Data from the NCS have already been used to estimate the costs of operating crime victim compensation systems. The orientation of casualty compensation, with an exception which will be discussed shortly, is entirely toward specific incidents of loss. To be administered economically, a crime compensation system would have to operate with a lower limit of liability, since by far the largest proportion of all incidents entail losses that are too minor to be worth the costs of processing them. The bulk of the data gathered by a victimization survey would fall below a reasonable threshold and therefore prove to be irrelevant to a feasible system of compensation.

The discussion earlier showed that "trivial" crimes can be significant, however. One significance they can have is when they are not isolated events but rather are associated with a status of the person making him vulnerable to continuing or repeated victimization. Public victim compensation systems, like private casualty insurance systems, are exclusively incident oriented and would be blind to such cases, however. Compensation would also not deal with such costs to the victim as anxiety and the effects on behavior proceeding from anxiety. Private casualty insurance does have one form of interest in victim vulnerability, multiple victimization, and the hazards of local social environments—its interest in avoiding insuring anyone who might fall within a high risk class.

A similar rationale applies to agencies for social control. The victimization survey should be necessary for data only with regard to instances of failures of (transgressions against) social controls which result in people being victimized by events that control systems aim to prevent, but which go totally unnoticed by these systems. In actuality, the survey is useful precisely because it helps compensate for the imperfections of control agencies as recorders and processors of information on their own transactions.

Individualistic and reductionistic contradictions. The recall problems of the victimization survey point to two contradictions in using the victimization survey for social indicators, where one seeks indicators which have the value standard of concrete *harm to specific individuals.*

First, the relatively trivial and ephemeral consequences of most individual incidents of victimization which the surveys (and other offense statistics) seek to measure are precisely a major source of the recall difficulty that affects survey data and that makes expensive, brief-reference-period, large sample panel techniques necessary for their conduct. As was pointed out, were the survey oriented solely to victimizations having serious and lasting consequences to individuals, many of its methodological problems would be greatly alleviated. Surveys restricted in this way could be conducted far more economically and reliably than those which attempt to bring into their scope even just those incidents that qualify as "serious" by being in the set used by Uniform Crime Reports for a "Crime Index," or which are felonious or indictable under law. An early finding of the victimization survey method was that the bulk of such incidents entailed relatively minor effects on the health or economic well-being of victims, relative, that is, to a host of other common hazards of daily life. They nonetheless were clearly serious from the standpoint of the offense given to legal, moral, and social norms. Aggregates of such incidents, furthermore, are of greater interest. For vulnerable individuals repeated victimization can be of major consequence as can be the fear of repeated victimization. The former—repeated victimization—directs attention, not to individual incidents, but rather to a pattern of vulnerability to victimization over spans of individual lives. While the latter—that is fears which may lead to costly alterations of people's life patterns—can result from a single and even minor incident of victimization, they arise when one's victimization is seen as related to a persisting condition of vulnerability specific to the individual or general to the community. For *social* indicators, particularly important are fears resting on perceptions of the extent to which norms and laws with regard to respect for persons and their property do or do not control the conduct of fellow inhabitants of one's community.

One way in which we may seek to avoid criminal victimization is by individual action; that is, by being guarded in exposing one's property and one's person to others; by retreating, figuratively, into a constricted life behind defensible walls. To the extent that such guarded and constricted behavior becomes general and is successful, indexes of offenses and victimization would be kept down, but at vast costs in real freedom and in the richness of life.

Another manner of protection is collective means; that is, through a normative system internalized by the processes of socialization and reinforced by both private and public sanctioning of deviance from the norms. It is upon these collective means, rather than individual guardedness, that we are mostly dependent for organized life in society. The role of crime counts as indicators of this social condition, of the vitality of the normative order, is what has elevated them as indexes to the major position as "social indicators" that they have always occupied since men first sought quantitative measures of the state of their society. Indicators of the extent to which people and organizations are guarded or trusting in their relations with each other, or feel they must be, would serve this purpose as well as offense or victimization data do. To look for the significance of crime exclusively in the concrete harm resulting from a given event to a specific individual misses most of the significance of most crime to the individual.

There is a second contradiction in the individualistic, disaggregated orientation to victimization indicators which the problems of recall also bring to our attention. It is implicit in an assumption about motivation inherent in the survey approach. Cooperation with a public survey by a respondent makes little rational sense from the standpoint of individual self-interest for most respondents. For a few, participating in a lengthy interview on the subject of victimization may be a welcome intellectual and social diversion in lives otherwise barren of such experiences, but to most respondents, presumably, the survey constitutes an intrusion into a preferred round of activities. Cooperation with a survey does make rational sense, however, as a civic act; a citizen duty. If an interview is couched solely in terms of *private* significance, what should motivate respondents to recall and report in that interview events of relatively small private significance? Indeed, is it not a contradiction in itself to conduct surveys which depend upon a collective sense and a civic regard, and which are in themselves of value only in a collective way, yet have these surveys guided by a rationale to which nothing makes sense but private self-interest?

Immunizing effect of exposure

Early observations in victimization surveys found empirical distributions of the number of persons (Ni) victimized a given number of times (Vi) in a time interval in which N_{vi} was lower for all $Vi > 1$ than would be expected on the basis of a probability model that assumed an equal or random distribution of risk among the entire population and given the mean incidence for population. While refinements of the victimization survey method have produced distributions of n's of victimization having different shapes than this, and while there are many plausible post hoc interpretations in terms of both phenomena and method for the distributions of multiple victimizations that are observed, nonetheless, the problem invites consideration of a possible immunizing effect of victimization.[1]

[1]We are concerned here only with "natural immunization."

There is an explanation for an immunizing effect of interspecies victimization that is captured in everyday speech by the phrase, "Once bitten, twice shy." There may be some virtue, however, in exploring the application of a medical analogy to criminal victimization. Considering such an analogy suggests some additional significant elements of immunization that are not suggested by the folk wisdom we have cited. Indeed, it is questionable whether avoidance behavior is a good analogy to the concept of immunization in medicine. Learning, or even unconscious avoidance, as a result of slight negative reinforcements does indeed produce less vulnerablility to more severe perils from the same harmful agency or class of agencies as was the source of the original insult. For it to fit the medical analogy better, however, the victimizing experience has to result in some change in the subject that makes it less vulnerable to actual attacks by the criminal agent, rather than a behavioral change in the direction of avoiding that agent. Locks on doors, bars on windows, a burglar alarm system and other so-called "target hardening measures" as well as such steps as training in self-defense measures, carrying weapons on the part of individuals, and various adjustments on the part of merchants with regard to rearranging patterns of customer movement and merchandise display could all be illustrative here.

Another form of immunization is the immunization from exposures which controls overreactive responses of a self-defensive sort. An analogy here can possibly be made to immunization from exposure to the effects of allergies. Experience with some forms of victimization, for example, petty theft or minor assaults, can develop the reaction, "It is not as bad as all that," and actual exposures to crimes of low consequence may eliminate relatively incapacitating adjustments such as extreme fear and behavioral circumspection. Residents and merchants in areas of fairly pervasive disorder may have such immunizing exposures against overreactive defense mechanisms.

As in the case of medicine, however, both analysis and treatment in criminology confront the problem that a given exposure can lead to the mobilization of either appropriate or inappropriate defense reactions, depending upon complex relationships between the individual and his exposure to the agent.

Thus far we have touched only on individual-level immunization. Two kinds of social or collective perspectives can also be taken. The first is the development of a population immunity. This takes place when a sufficient number of individuals reach a sufficiently high level of immunity so that the survival of the offending species is affected. It takes place only when the offending agent is dependent for its survival on a population of hosts subject to immunization. It requires the creation of a very high level of long lasting immunity of a very large percentage of the population. This removes the opportunities for the criminal agent to survive and reproduce so that it approaches extinction. Levels of immunization in extent and efficacy below this level of population immunity can generate increased vulnerability.

Population immunities also develop by the differential survival of its more immune members. This can occur with or without change in the prevalence of the agent in the territory of the community or the mean virulence of the agent. An example of selective survival resulting in elevated population immunity would be decreased ecological burglary and robbery incidence rates occurring in a community as the most vulnerable small business establishments fail because of crime losses. There may be no associated decrease in the prevalence of offenders in the territory of that community (they may prey more outside its confines) and the offenders remaining may be more likely to use extreme measures in the offenses they do commit among the remaining "harder" targets.

Another way in which the social analogy can be pursued is by regarding the community's defenses as the immunizing effects of exposure rather than individuals'. Most of the defensive adaptions to crime that are possible are collective ones taken at the community rather than the individual level. "Social defense" presumably requires prompt, effective immunizing reactions to low levels of attack.

Most of the attention in criminology has been given to the effects of measures of social defense. "Victimology," however, concentrates its attention on the individual. Some recent evaluation studies, for example, in Portland, Oregon, have attempted to analyze the consequences of individual defensive measures for aggregate levels of crime incidence. Such studies logically present the statistical requirement for taking into account the effects of victimization on subsequent individual and population vulnerability to victimization—a difficult and as yet untouched analytic problem.

Risk of victimization

Consideration of the concept of immunizing victimization may also help guard against tempting fallacies inherent in the frequent tendency to equate incidence rates with "risk."

"Risk," first of all, conveys a future, rather than a past, temporal implication. If there is a high immunizing effect of exposure, the very individuals who contribute to incidence during any period of observation may be precisely those who will be found to have been at low risk during some later period of observation (i.e., there will be few instances of repeated victimization). This may be the case even with no change, or even some increase, in the total rate for the community or class to which the victimized individuals belong, provided only that the supply of those vulnerable to victimization is very large relative to the actual incidence of victimization.

Secondly, statistical risk inferred from an incidence rate for a population is expressed as a probability value for a member of that population. Distributions of individual risk, in the sense of future vulnerability, can be very unevenly distributed among the population, however. If the concept of risk is to be associated with an incidence rate, either there must be empirical and theoretical grounds for assuming a high degree of homogeneity of vulnerability among the population or care must be taken to use words that suggest the "risk" of aggregates, rather than individuals.

This suggests additional difficulties with the risk concept. If there is a high concentration of victimization among some portion of a community, an inaccurately low measure of incidence for the entire community may more accurately reflect the risk rate for the large majority of the members of the population (or the median or modal "risk") than would an incidence rate based on a more exhaustive measure of incidence. There are good indications that this is indeed the case with rates based on victimization survey data in that small, but highly vulnerable, components of the population are most subject to underenumeration. For an aggregate risk level indicator most often most applicable by members of a community, that is, one descriptive of the "chances of my being victimized if the future resembles the past," an incidence rate that neglects the experience of both the most and least vulnerable components of the community would serve best. From the standpoint of social problem concerns, it may be only or mostly those who suffer frequent victimization and who benefit from no immunizing effects, who suffer appreciably at all. This is true where the effects of any one victimization are small, but repeated victimization may be costly or fatal.

Determining which components of a community may be those most vulnerable presents very different statistical problems if a large proportion of all the victimization (in the real world or in that time slice of it that we can capture in our measurements) stems from many repeated victimizations of the same individuals than it is where repeated victimizations are rare or nonexistent (as in the case with homicide "risks"). The last case is analogous to the fatal or usually fatal disease. The only possible immunizing effect is one of population immunization by selective elimination of the more vulnerable. Rare multiple victimization can also be a function of extremely low levels of relatively randomly distributed incidence. High multiple victimization can be found associated with either very high but widely dispersed total incidence or with highly concentrated (nonrandom) incidence.

If multiples are rare, then we have no means for estimating a given individual's vulnerablility other than by categorically associating him with some class that has an observed rate of incidence. (If multiples are rare because of fatality, we must use a statistical logic of nonreplacement.) If multiples are common, however, and if we are capable of making sufficiently long-term observations of individuals, then we can state empirical vulnerability rates for individuals and observe the stability of predictions of future victimization from observations of past victimization for each individual. If multiples are common because of concentration, this may yield only slight improvement over "risk" estimates based on aggregate incidence for subclasses, providing the correlations of incidence with the classifications available are very high. If there is great dispersion, however, individual measures will be required for accuracy, as they will also be if there are immunizing effects of any great consequence.

References

Biderman, Albert D. (1975)
"Victimology and victimization surveys." In I. Drapkin and E. Viano (eds.) *Victimology: A new focus* (vol. III, Crimes, victims, and justice), Lexington, Mass.: Lexington Books.

Reiss, Albert J., Jr. (1973)
"Surveys of self-reported delicts." A paper presented for the Symposium on Studies of Public Experience, Knowledge, and Opinion of Crime and Justice, Washington, D.C., March 17-18, 1972, revised July 1973.

Victimization and the National Crime Survey: Problems of design and analysis

STEPHEN E. FIENBERG
Department of Statistics
and Department of Social Science
Carnegie-Mellon University

1. Introduction

Crime and its impact on society have long been the subject of public interest and social concern. While the study of crime has proved profitable to social scientists over the years, the limitations of police crime statistics (e.g., see Biderman and Reiss, 1967) have always been viewed as being so great as to make it virtually impossible to measure criminality in a population. Hood and Sparks (1970) note that "Questions about criminality, like those about sexual behavior, are especially liable to distorted and untruthful answers." Thus it was with great anticipation that the social science community heralded the adoption of survey research methods to find the victims of crime, and to learn of their experiences. As a result of some small-scale attempts at victim surveys in the United States and Great Britain, and after considerable planning and preparation, the Law Enforcement Assistance Administration (LEAA) initiated a major new social statistics series based on a national victimization survey.

The primary purpose of the national victimization survey, as stated in a planning document developed by LEAA, is "to measure the annual change in crime incidents for a limited set of major crimes and to characterize some of the socioeconomic aspects of both the reported events and their victims (Penick and Owens 1976, p. 220)." Henceforth, we refer to this survey as the National Crime Survey (NCS), but the reader should bear in mind that the focus of the NCS is upon *victims* and *their* experiences with crime, not on crime itself.

Actually the NCS consists* of four separate surveys:

(1) A continuing national survey of household locations.
(2) A continuing national survey of commercial establishments.
(3) A separate set of single or duplicated surveys of household locations in selected cities.
(4) A set of city commercial surveys to parallel (3).

In this chapter we restrict our attention *solely* to the continuing national survey of household locations.

The NCS has been designed and executed for LEAA by the U.S. Bureau of the Census and it includes personal interviews at six-month intervals with individuals in up to 65,000 households. Given the magnitude of the NCS and the massive files of data collected since the initial field work began in mid-July of 1972, it is remarkable that the NCS has received so little attention from professional statisticians outside of the Bureau of the Census.

Central to an examination of victimization and the concepts underlying the NCS is the notion of a crime or criminal incident and how it gets recorded by various criminal justice agencies. The dictionary definition of crime offers little in the way of a starting point. For example, a recent edition of the Random House Dictionary defines crime as

> an action or an instance of negligence that is deemed injurious to the public welfare or morals or to the interests of the state and that is legally prohibited.

To shed some light on this matter, Section 2 describes in detail a single criminal incident, and notes how it would be recorded in statistics gathered by the police and in the NCS.

Section 3 contains a brief summary of the survey and questionnaire design of the NCS, and describes some aspects of its execution. Special attention is focused on the panel structure of the survey design, with a rotation plan for households. The major shortcomings of the design are then noted. Section 4 is brief and it summarizes the published analyses from the NCS. The lack of LEAA resources devoted to the statistical analyses of NCS data was one of the principal findings of the Panel for the Evaluation of Crime Surveys appointed by the Committee on National Statistics (Penick and Owens 1976, p. 3). This report contains considerably more detailed descriptions of the NCS survey and questionnaire design than we provide here. It describes the developmental research behind the design, and it suggests areas for further investigation. The conclusions of the report overlap considerably with ours regarding the need for extensive ongoing methodological research.

Any assessment of the NCS must look closely at its objectives and determine to what extent they are being met. The primary purpose of the NCS as described actually has several components:

(1) To measure the incidence of crime.
(2) To measure the changes in crime rates over time.
(3) To characterize socioeconomic aspects of criminal events and their victims.

Closely related to item (3) are the aims

(4) To identify high-risk subgroups in the population and to estimate the rate of multiple victimization.
(5) To provide a measure of victim risk.

From its inception, the NCS was viewed as a multipurpose survey that would produce not only the general-purpose victimization rates already described, but also data for policy-oriented problems; for example,

(6) To calibrate the Uniform Crime Reports data produced by the FBI.
(7) To index changes in reporting behavior.
(8) To measure the effectiveness of new criminal justice programs (the city surveys were initiated for exactly this reason).

To determine if the NCS properly fulfills aims (1)–(4) special attention needs to be

*Since this paper was prepared, all but the continuing national survey of households has been discontinued.

 ISBN: 0-12-513350-2. Reprinted by permission, with minor changes, from *Survey Sampling and Measurement* (K. Namboodiri, ed.).

focused on questions that utilize the longitudinal structure of the NCS. Section 5 outlines a number of substantive questions regarding victimization and victim-survey methodology that in principle should be answerable by analysis of NCS data. A major stumbling block to the successful completion of these analyses is the highly complex NCS survey structure, designed to produce descriptive statistics rather than data amenable to analytical studies of interrelationships and their changes over time. Although the NCS is a rotating panel in form, the primary purposes of the panel structure are to get more stable rate comparisons from one period to the next, and to bound the time frame under consideration.

2. Recording crime

Criminal incidents are events or social encounters involving one or more offenders and one or more victims, in one or more locations for specific periods of time. The duration of a single criminal incident may be 10 minutes, an hour, a day, a week, or even a month. Nonetheless, when put into a larger timeframe a criminal event is quite profitably viewed as the realization of a point process distributed over time and space, and we do so in Section 5. What complicates the modeling of a large number of crimes is the interpenetrating social networks linking offenders and victims, both within a single incident and across several incidents, and giving rise to multiple offending and multiple victimization. Reiss (1980) describes some of the impact of such networks and associated group structures on crime rates with special attention to the implications for measuring the effects of deterrence and incapacitation. The stochastic structure of criminal social networks and the resulting lack of independence of criminal incidents also has potentially important implications for both the design and analysis of victimization surveys. It is for this reason that we discuss some first steps in the stochastic model of victimizations for individuals over time in Section 5.

How one records crime is a function of one's perspective. A single criminal incident or social encounter can involve one or more offenders, one or more victims or possibly no victims at all, and multiple violations of the law leading to multiple indictments of a single offender or several offenders who have participated in the event. There may even be mutual offending and victimization, for example, in cases of assault. Thus a particular configuration of crimes aggregated over a given time period may well look dramatically different when viewed from the perspective of offense rates as opposed to victimization rates, and neither set of rates is likely to reveal the true nature of the criminal events that have taken place.

A single hypothetical example can illustrate the complexity associated with criminal incidents and the manner in which they are recorded. A young couple living in the household of the woman's parents in Stamford, Connecticut, go to New York City on December 31 to celebrate New Year's Eve. They park their car in a lot on the east side of Manhattan and have a leisurely dinner at a nearby restaurant. After dinner when they return to their car, they are accosted by five young males just outside the parking lot and are taken into an adjacent alleyway, at approximately 11:00 p.m. One of the youths threatens the couple with a revolver, and the other four take turns raping the woman. When the woman resists, one of the youths assaults her with a knife, and then he also assaults the man. Following the acts of rape the youths take the woman's purse and the man's wallet, and they appear to flee. It is now about 1:00 a.m., January 1. The couple have to travel several blocks to report the incident to the police. When they finally return to the parking lot with a police officer at 3:00 a.m., they discover that their automobile is missing. A week later three young males are stopped by the police in Newark, New Jersey, driving the couple's car through a red stoplight and they are arrested.

The incident just described involved five offenders, two victims, three arrests, and numerous offenses including forcible rape, robbery, aggravated assault, and motor vehicle theft. It spanned several hours (and two calendar years!) and took place in at least two locations. How would it be classified by various recording systems?

Let us begin with the police record of the event as it is transmitted to the FBI for use in its *Uniform Crime Reports* (UCR). In a multiple offense situation, the police classify each offense, and then locate the offense that is highest on the list of what is known as Part I Offenses (the ranking is criminal homicide, forcible rape, robbery, aggravated assault, burglary, larceny-theft, and motor vehicle theft). The highest offense is entered and the others are ignored. Multiple offenses need to be separated in time and place to lead to multiple entries in the UCR. The exeption to this rule involves crimes against the person (criminal homicide, forcible rape, and aggravated assault) where one offense is entered for each victim. Thus the UCR record will contain one offense of forcible rape (against the woman) and one offense of aggravated assault (against the man). Had the youths only robbed but not assaulted the man, there would only be one offense entered. These offenses would be recorded by the New York City police, and I am unclear as to which day (and thus which year) they will be attributed. The UCR record will also show that the offense(s) have been cleared (i.e., "resolved") by the arrest of the three youths in New Jersey. Although this event led to one or two UCR offenses, it might well lead to the prosecution of the five youths on up to a total of five counts of rape, 10 counts of aggravated assault and robbery, and five counts of motor vehicle theft.

Suppose now that the couple's household is chosen as part of the NCS so that the event will also be recorded from the victim's perspective. Both the man and the woman would be interviewed separately and the NCS would record two victimizations in December: one for the woman "assaultive violence with theft—rape," one for the man "assaultive violence with theft—serious assault with weapon." Even if the man had only been robbed but not assaulted there would still be two victimizations recorded (as compared with a single offense). Moreover, because of the separation of household victimizations from individual victimizations, when the woman's father reports the household victimizations, he may well report the theft of the car separately, and the month of victimization may be given as January, and thus it could go into a separate calendar year.

In summary, our single criminal incident involving five offenders and two victims, leads to one or two offenses recorded in New York and two or three victimizations recorded in Connecticut. The perspectives are clearly different, and so too are the records of the event.

Because a large proportion of criminal incidents is never reported to the police, the discrepancy between all criminal offenses and those reported to the police has been described by Biderman and Reiss (1967) as the "dark figure" of crime, and one of the original purposes of victimization surveys was "to bring more of the dark figure to statistical light." Biderman and Reiss go on to note:

> In exploring the dark figure of crime, the primary question is not how much of it becomes revealed but rather what will be the selective properties of any particular innovation for its illumination. As in many other problems of scientific observation, the use of approaches and apparatuses with different properties of error has been a means of approaching truer approximations of phenomena that are difficult to measure.
>
> Any set of crime statistics, including those of the survey, involves some evaluative institutional processing of people's reports. Concepts, definitions, quantitative models, and theories must be adjusted to the fact that the data are not some objectively observable universe of "criminal acts," but rather those events defined, captured, and processed as such by some institutional mechanism [pp. 14–15].

Much controversy has centered on the comparability of police statistics on offense rates and NCS survey statistics on victimization rates (e.g., see Biderman 1967; Biderman and Reiss 1966; Penick and Owens 1976, pp. 152–154; U.S. Department of Justice 1976b), but the utility (or lack thereof) of the NCS data for such comparisons should not obscure the richness of information about victimization available in the NCS. It is for this reason that the NCS data must be collected and organized in a manner that will make it amenable to standard forms of statistical analysis. Otherwise the rich veins of information on such topics as high-risk segments of the population and multiple victimization, or the way that deviance is perceived and dealt with in various social contexts, may never be mined.

3. Design of the NCS

a. Sample design

The NCS is a sample survey of households and their occupants, and as such it closely resembles the Current Population Survey (CPS), which is also conducted by the Bureau of the Census, in almost all aspects. In fact, descriptions of the designation of housing units for the CPS (e.g., see Thompson and Shapiro 1973) are almost identical to those for the NCS (e.g., see U.S. Department of Justice 1976a,b), the major exceptions being the sample sizes, the interview schedules, and the panel and rotation group structures.

The structure of the NCS is that of a stratified multistage cluster sample. The first stage consists of dividing the United States into approximately 2000 primary sampling units (PSUs) comprising counties or groups of contiguous counties. The PSUs are then separated into 376 strata and one PSU is selected from each stratum with probability proportional to population size. Within each PSU so selected, a systematically chosen group of enumeration districts is selected, and then clusters of approximately four housing units each are chosen within each enumeration district. For 1973, this process led to the designation of about 80,000 housing units, and interviews were obtained from occupants of about 65,000. Most of the remaining designated housing units were vacant or otherwise ineligible for inclusion in the NCS.

The basic sample is divided in six subsamples or rotation groups of a little over 10,000 households each. (Actually there are seven subsamples, but the data for the newest one are not incorporated into the reported rates. Rather these data are used for bounding purposes, as described in Section 3, *a*.) The occupants 12 years of age or older are interviewed at six-month intervals for a total of three years. Every six months a new rotation group enters the sample and the "oldest" existing rotation group from the previous sample is dropped. Each rotation group is divided into six panels, with one panel being interviewed in each month of the six-month period.

For estimating various rates, a series of weights and adjustment procedures are applied to the raw data. The weighting procedures are standard practice for surveys of this sort and are basically designed to adjust for the differential probabilities of including various household locations in the survey, and to reduce bias and variance of sample estimators. The final adjustment involves the use of ratio estimation so that the distribution of individuals (or households) in the sample is in accord with independent estimates of the current population in each 72 age-sex-race categories.

By reporting only adjusted rates, for both the NCS and the CPS, Census has removed from public scrutiny many of the actual defects of the sample design when it is actually implemented. Since all aggregate counts have essentially the same totals for various categories, we can never tell when a given sample is badly off the mark, nor in what directions.

Although the NCS is basically a *sample of household locations*, at the same time it yields both a *sample of households* or families and a *sample of individuals*. Household locations are of little substantive interest in the study of victimization. While the NCS allows for the study of differential rates of victimization by type of household location (e.g., house, apartment, rooming house, mobile home), not one of the 100 tables in the LEAA report for 1973 (U.S. Department of Justice 1976b) deals with such information. The primary reason that the NCS is a sample of locations rather than households or individuals appears to be because Census has available a detailed frame only for locations.

The NCS primarily measures victimization while the CPS primarily measures employment and unemployment. Since both unemployment and victimization are relatively rare phenomena, a naive person might suggest that a sample design that has proved successful for measuring unemployment should, with only minor modifications, do a good job of measuring victimization. Such a suggestion is naive because, among other things, it ignores the considerable knowledge we have available regarding crime and its physical as well as socioeconomic characteristics. In central cities, crime rates vary dramatically from block to block, and a limited amount of fieldwork might lead to cluster boundaries that differ dramatically from those that would seem appropriate for unemployment. It may well be that the NCS sampling plan is most sensible given budgetary constraints, but an exploration of alternatives and variants to the current plan should probably be included in the research, development, and evaluation program of the Bureau of the Census.

b. *Questionnaire design*

The questionnaire administered every six months at each household consists of two parts: a basic screen and crime incident reports. The basic screen includes household location information, household or family information, the personal characteristics of all of the individuals in the household (who may change from interview to interview), plus household or individual screen questions on crime. The report of the Panel for the Evaluation of Crime Surveys (Penick and Owens 1976) gives a detailed critique of the basic screen, and we refer the interested reader to their discussion. For each crime incident detected by the screen, a crime incident report containing answers to almost 100 questions is completed.

The questionnaire distinguishes between individual identifiable incidents and series of at least three similar incidents which the respondent is unable to separate in time and place of occurrence. For individual victimizations, the questionnaire records the month in which the crime took place, but for series victimizations the respondent only needs to indicate the quarter(s) in which the incidents took place (i.e., spring, summer, fall, winter), the number of incidents (3–4, 5–10, 11+, or don't know), and the details for the most recent event in the series. We discuss the distinction between single and series victimization in more detail in Section 4, where we note how the Bureau treats series victimizations and why we believe series victimizations should be the topic of extensive and analytical investigation. What is unclear to us from published documents and various unpublished memoranda is the extent to which series victimization is a true phenomenon or an artificial construct resulting from the NCS questionnaire design.

Not only does the NCS questionnaire solicit information on the details of an incident, the offender, and any resulting physical injury and how it was treated, but it also inquires whether the incident was reported to the police and if not, why not.

c. *Reference period and bounding*

One of the most crucial problems in the design of a victimization survey is eliciting accurate information on the time of occurrence of criminal incidents. The problem has at least two components:

(1) *Recall decay*. The longer the time lapse between a criminal incident and the date of interview, the greater the probability that the event will not be reported to the interviewer.

(2) *Telescoping*. Events occurring in one time period can be reported as occurring in a different one. The displacement of telescoped events can be forward or backward in time.

It is especially difficult to model recall decay and telescoping, since such evidence seems to point to differential rates of decay and telescoping for different types of crimes, and for different types of respondents. Moreover, there can be no check on a crime that has never been reported, either to the police or the NCS. Thus the only way to get a handle on these two phenomena is via a sample of crimes reported to the police and the subsequent inclusion of victims of these *reported* crimes in a victim survey. Such "reverse record checks" were part of the pretests of the NCS survey instrument (see U.S. Department of Justice 1972, 1974). The problem with drawing inferences from reverse record checks is that they are aimed at data which are missing from the victimization survey, but which are *not* missing at random (see Rubin [1976] for a discussion of the importance of the missing at random assumption).

A consideration of both recall decay and telescoping is necessary for the determination of the optimal reference period for a victim survey. The NCS reference period is six months, and Census uses the first interview and six-month period of a household location for *bounding*, that is, establishing a time frame to avoid duplication of incidents in subsequent interviews. For a detailed study of the effects of bounding on telescoping, see Murphy and Cowan (1976). A major problem in the design of the NCS arises because the bounding procedures bound household locations, not households or individuals. If one household replaces another during the course of the three-year period during which a location is included in the NCS sample, there is no bounding for the new household or for its members as individuals. Murphy and Cowan (1976) report that unbounded households in returning rotation groups comprise (for 1974–1975) 13.3 percent of the interviewed sample. In addition, only about 95 percent of the interviews in the bounded households are themselves bounded due to considerable transience for households in heavily urban areas. As a result, as few as 20 percent of the individuals over a three-year period in a given set of household locations may produce complete victimization records for the period. These design characteristics drastically impair the utility of the NCS data for longitudinal analysis of individual victimization profiles.

Considerable methodological interest is centered on the differences in victimization experience for migrants and nonmigrants. In addition to follow-up studies of out-migrants (which are quite costly), it seems reasonable to do special analyses of the in-migrants to the sample locations since their data are already in the NCS (see Penick and Owens 1976; Reiss 1977b). For every out-migrant household there is an in-migrant one. Of course the current lack of bounding for in-migrants would complicate such analyses, but it should be feasible to do a special study of in-migrants where a bounding period would be included along with additional interviews beyond the standard three-year period for the household location.

4. Published analyses of the NCS data

Not only does the formal responsibility for the design and execution of the NCS lie with the U.S. Bureau of the Census, but the analysis of the collected data is also the responsibility of a small staff of Census employees. This analysis by LEAA and Census involves the periodic preparation of two- and three-dimensional cross-tabulations of estimated victimization rates and estimates of their standard errors. The cross-tabulations produced are basically those requested in advance by professional staff at LEAA, and not as a result of a more detailed and complex statistical analysis.

Suppose for simplicity that NCS employed a simple random sample and that the data (which are primarily categorical in nature) for any year were analyzed using some variant of loglinear model analysis for a k-dimensional cross-classification (e.g., see Bishop et al. 1975). Then one of the implications of finding a model that gives a good fit to the data would be that the k-dimensional table may be succinctly summarized by a series of tables of smaller dimension, from which the original table can be reconstructed with essentially zero information loss. Such analyses can thus provide a rationale for reporting certain cross-tabulations and not others. This point is described in more detail by Fienberg (1975). Even though the NCS does not employ simple random sampling, the idea of careful statistical analyses leading to the choice of cross-tabulations to be published is one which should be considered more seriously by LEAA and Census.

How many reports has LEAA published on the results of the NCS national household sample? As of December 1976*, several preliminary but only two final reports had been released: a 162-page report on the 1973 survey (U.S. Department of Justice 1976b), and a much briefer 73-page report comparing findings for 1973 and 1974 (U.S. Department of Justice 1976a). Since both final reports also contain data on separate commercial surveys, the interested reader is left with very slim pickings from what appeared to be a sumptuous meal. Moreover, these two reports contain only weighted data or proportions. No raw counts are available. Thus it is almost impossible for the skilled statistician to do extensive secondary analysis of the published data.

When preliminary versions of the 1973 report were distributed by LEAA, several investigators noted that series victimizations were not included in the computation of any published rates or calculations. Thus all reported numbers and rates of victimization may be severe underestimates. For example, LEAA estimated for 1973 (U.S. Department of Justice 1976b) that there were approximately one million series victimizations in the personal sector and just over 20 million victimizations not in series. A series consists of *three or more* victimizations, and an average of five victimizations per series is likely an underestimate for the NCS data. (Some calculations based on an unpublished tabulation suggest that the average is in excess of six victimizations per series.) This then means that at least 20 percent of all victimizations in the personal sector have been excluded from the reported calculations. This matter becomes even more serious when we note that, in 1973, 46.3 percent of all personal series victimizations involved crimes of violence while only 26.6 percent of all victimizations not in series. Thus, series victimizations may have accounted for over one-third of all crimes of violence.

*The final draft of this paper was completed then.

We note that despite the panel structure of the survey, LEAA has yet to make use of the full longitudinal structure of the data base. The construction of a panel tape tracking individuals and households over time was not deemed to be a central goal of the NCS, and the preparation of such a tape was only belatedly arranged through a contract with a group at a private university. It might be argued that the panel structure of the NCS sample is intended to handle certain technical problems and to give more accurate year to year comparisons, and not for longitudinal analysis of individual files. This can be true only in this narrowest of senses because without a detailed longitudinal analysis we can never know whether the aggregate annual reported victimization rates are at all accurate. For example, Reiss (1977), reporting on some preliminary longitudinal analyses, notes that highly victimized individuals are much more likely to be out-migrants than those with low victimization rates, and series victims are more likely to move than nonseries victims. Moreover, a high percentage of individuals reporting series victimizations in a given six-month period report no victimizations in the subsequent six-month period. These observations call into question the accuracy of the published victimization rates.

5. Modelling victimization

To understand reported annual victimization rates and the implications of changes in them from one year to the next, we need a detailed understanding of how victimization varies among individuals and subgroups within the population. This detailed understanding will necessarily have to come from the analysis of disaggregated data, and of individual victimization records over time. Such analyses will be complicated by the complex structure of the NCS sample design, but the effects of stratification and clustering on analyses will vary greatly from problem to problem. For many problems the use of unweighted data may well simplify the modelling process. This is clearly the case if we are interested in the structure of individual reported victimization patterns over time.

The Panel for the Evaluation of Crime Surveys gives several suggestions for analytic research on the existing NCS data. One of these suggestions deals with the relationship between series victimization and multiple victimization, a topic we discussed in Section 4. To investigate this relationship, however, we need models for the occurrence of victimizations over time, and we propose one such model in Section 5*b*. A second suggestion deals with analyses to investigate under- and over-reporting of incidents as they relate to the month of incident and the month of interview. In Section 5*a* we take up some aspects that need to be considered in such analyses.

a. *Reporting biases and time-in-panel*

For several characteristics on which data are collected in the Current Population Survey, Bailar (1975) notes that there is a higher level for the first interview than for succeeding ones, and so on. The effect of such variation is usually referred to as "rotation group bias," and there is reason to expect such biases in the NCS data as well. In the NCS the rotation group bias problem is compounded by several factors including the elapsed time between the incident and the interview (recall that interviews provide data for the preceding six-month period).

What we would like to do is develop a model which compares the victimization rates for specific crimes for a series of reference months as a function of the number of interviews, the time-lag from incident to interview, and other possibly relevant temporal variables. We build up to this in stages.

In Table 1 we show the list of panels being interviewed by month of collection for a full three-year collection cycle, where the months have been labeled from 31 to 66. Panels 1–6 form a subsample that was first interviewed in months 1–6 (we ignore the initial interview for bounding purposes here) and leaves the sample after the interviews in months 31–36. Note that the difference between the month of collection and the number of a panel being interviewed equals the number of months the panel has been in the sample (time-in-panel). All three variables bear examination in terms of their effects on reported rates. The time-in-panel variable yields the rotation group bias information, while month of collection measures seasonality and other unique temporal effects, and panel number represents temporal characteristics and effects unique to those that entered the sample at the same time. The formal identity linking these three variables is the same as that linking age, period, and cohort as described by Fienberg and Mason (1978), and any model using all three as independent variables needs to take into account the identification problem associated with the linear components of the effects.

Since each interview collects data for the preceding six-month period, for each *reference* month there are a total of 36 distinct panels which provide data. For example, panels 1, 7, 13, 19, 25, and 31 provide data with a one-month lag for month 30 during collection month 31; panels 2, 8, 14, 20, 26, and 32 provide data with a two-month lag during collection month 32; and so on. Thus the ensemble of 36 victimization rates for a given reference month can be modeled as a function of month of collection, time lag to reference month, panel number, and time in panel (as well as various additional independent variables such as education and race if we wish to compare subgroups of the sample).

Of course we need to model several reference months simultaneously if we are to use all of the independent variables at once. If we in addition use reference month as an independent variable, then we have an additional identification problem related to the identity involving reference month, collection month, and time-lag until interview.

To analyze and model data using the variables just described, we need to know whether we can treat the data for different reference months from the same panel as being independent. Moreover, it is unclear whether we should use rate as the response variate or counts of victimization (e.g., the number of respondents with 0, 1, 2 . . . victimizations), and whether we should use weighted or unweighted data.

Models of the sort we have just described need to be explored carefully if we are to get a proper handle on such problems as rotation group bias and memory decay associated with recall. Modeling these phenomena separately (e.g., Bailar 1975; Finkner and Nisselson, Chapter 5, 1978) when they in fact occur simultaneously should only be the first step in an analysis, since it may lead to improper inferences unless there are order-of-magnitude differences in the sizes of their effects. What is especially troublesome with any attempt to model these phenomena is that we can deal only with individual victimizations, and not series, even though the latter may make up a sizeable proportion of the total reported victimizations in a given period.

Table 1. An illustration of the NCS panel rotation structure

Collection month	Panels being interviewed										
31	1	7	13	19	25	31					
32	2	8	14	20	26	32					
33	3	9	15	21	27	33					
34	4	10	16	22	28	34					
35	5	11	17	23	29	35					
36	6	12	18	24	30	36					
37		7	13	19	25	31	37				
38		8	14	20	26	32	38				
39		9	15	21	27	33	39				
40		10	16	22	28	34	40				
41		11	17	23	29	35	41				
42		12	18	24	30	36	42				
43			13	19	25	31	37	43			
44			14	20	26	32	38	44			
45			15	21	27	33	39	45			
46			16	22	28	34	40	46			
47			17	23	29	35	41	47			
48			18	24	30	36	42	48			
49				19	25	31	37	43	49		
50				20	26	32	38	44	50		
51				21	27	33	39	45	51		
52				22	28	34	40	46	52		
53				23	29	35	41	47	53		
54				24	30	36	42	48	54		
55					25	31	37	43	49	55	
56					26	32	38	44	50	56	
57					27	33	39	45	51	57	
58					28	34	40	46	52	58	
59					29	35	41	47	53	59	
60					30	36	42	48	54	60	
61						31	37	43	49	55	61
62						32	38	44	50	56	62
63						33	39	45	51	57	63
64						34	40	46	52	58	64
65						35	41	47	53	59	65
66						36	42	48	54	60	66

b. *A model for multiple victimizations over time*

Most of the models that have been proposed for victimization assume that each individual has an "annual" victimization rate λ_i for crime type *i*, and that the expected number of victimizations the individual will experience for crime type *i* in a fixed period of time *T* is simply $\lambda_i T$. This is, of course, the expected number if we assume that victimizations follow a Poisson process. Since victimization is a rare event, in order to test the Poisson model we need to pool individuals into groups expected to have similar values of λ_i. Those victimization studies that have looked at victimization distributions for fixed periods of time and for subgroups of the population typically find that the Poisson model gives a poor fit. This may be an artifact of the data collection procedure, it may be a result of not using a fine enough disaggregation, or it may in fact be the result of the inappropriateness of the Poisson process.

One more general structure for modeling victimization as a point process is the semi-Markov process, which includes the Poisson process as a special case. In this structure we view victimization as a point process $\{Y(t), t > 0\}$, where $Y(t) = j$ if the individual were last a victim of crime type j. If the process is semi-Markov (see e.g., Çinlar 1975), then it has transition probabilities

$$p_{ij}(t) = \Pr\{Y(t) = j \mid Y(0) = i\}, \tag{1}$$

where i and j run over the possible types of crimes, say $1 \leqslant i, j \leqslant r$. These transition probabilities can be expressed directly in terms of two sets of quantities:

(1) A matrix of one-step transition probabilities governing a discrete-time Markov chain, $\mathbf{M} = \{m_{ij}\}$, which represent an individual's "victimization propensities" given his current victimization state.

(2) A family of waiting time distributions, $\mathscr{F} = \{F_1(t), F_2(t), \ldots, F_r(t)\}$, characterizing the intervictimization intervals and depending on the last type of victimization.

The transition probabilities are the unique solution of the system of equations

$$p_{ij}(t) = \delta_{ij}[1 - F_i(t)] + \sum_{k=1}^{r} \int_0^t f_i(s) m_{ik} p_{kj}(t-s)\, ds, \tag{2}$$

where $i, j = 1, 2, \ldots, r$,

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j, \end{cases} \tag{3}$$

and $f_i(t)$ is the probability density corresponding to the distribution function $F_i(t)$. When the distributions $F_i(t)$ are exponential, the process reduces to a time-homogeneous Markov one, and when, in addition, the probabilities $\{m_{ij}\}$ do not depend on i, that is, the rows are homogeneous, we get a set of Poisson processes.

In order to use this general semi-Markov model for the NCS data, we need to see how the one-month-at-a-time data collection framework of the NCS can be embedded in the structure of the continuous time model. This problem resembles one explored by Singer and Spilerman (1974, 1976a,b), who have used the semi-Markov process model of Eqs. (1) and (2) for investigating occupational mobility. In their work they have placed special emphasis on the embeddability of fragmentary multiwave panel data into a class of continuous time Markov models, and the identification problem within that class of models.

The use of this class of models in the context of the NCS is complicated by the fact that as few as 20 percent of all individuals have full three-year records. Moreover, it is unclear whether we need to take into account the complexities of the sample design when we try to model the victimization histories of individuals with common sociodemographic and geographic characteristics. A final complication in the NCS data is the existence of series victimizations, which illustrate a strong propensity for rapid and repeated victimization of a specific type. Analyses based on underlying continuous time models certainly should include both series and separate individual victimizations.

6. Discussion

The two models described in the preceding section have not been explored with the NCS data, even in a preliminary form. They do, however, illustrate the problems involved in the analysis of data from the NCS when the purpose of the analysis is to provide estimates of aggregate victimization rates. While some have argued that modeling of this sort is unrelated to the primary objectives of the NCS, we disagree. First, we believe that an understanding of the basic structure of the panel data produced by the NCS is crucial to a proper evaluation of aggregate victimization rates. Second, the detailed stochastic modeling of individual records is required to directly meet one of the NCS objectives described in the introduction of this chapter: to identify high-risk subgroups and to estimate the rate of multiple victimization. Third, a reading of various documents about the NCS makes clear that it is in fact a multipurpose survey, and substantive issues and concerns need to be properly articulated so that the NCS design may be appropriately modified.

Because the NCS is similar in sample design to many other large-scale social surveys such as the CPS, the Annual Housing Survey, and the National Assessment of Education Progress, it shares with these other surveys various methodological problems associated with data analysis and inference. For example, the weighting procedures used to get aggregate victimization rates and estimates of standard errors are not necessarily appropriate for other analytical purposes. To solve these problems, statisticians must develop variants of various multivariate techniques appropriate for the analysis of data from complex surveys. At the same time they must work toward the development of survey designs that are especially amenable to classes of analytical purposes, or at least to specific forms of analysis.

Our evaluation of the NCS is well summarized by the following excerpt from the Report of the Panel for the Evaluation of Crime Surveys (Penick and Owens 1976):

> The panel has found much to commend, and much to criticize, in the design and execution of the NCS to date. We have argued that a very great amount of methodological and developmental research must be done, and many changes in existing procedures must be made, if certain of the specific initial objectives of the surveys are to be accomplished. The panel also maintains, however, that those objectives themselves need further scrutiny and that a subtle but fundamental change in the official concept of victimization surveying is necessary if the potential value of this relatively new research method is to be fully realized [p. 152].

Acknowledgments

This article grew out of material discussed in the Workshop on Criminal Justice Statistics held in Washington, D.C., July 1975, and sponsored by the Social Science Research Council Center for Coordination of Research on Social Indicators and the Law Enforcement Assistance Administration. The author is indebted to several of the participants of the Workshop whose ideas and suggestions inevitably have found their way into this article. In particular, thanks are due to Albert D. Biderman, Kinley Larntz, Albert J. Reiss Jr., David Seidman, and Richard Sparks.

References

Bailar, B. A. (1975)
"The effects of rotation group bias on estimates from panel surveys." *Journal of the American Statistical Association* 70:23-30.

Biderman, A. D. (1967)
"Surveys of population samples for estimating crime incidence." *The Annals of the American Academy of Political and Social Science* 374:16-33.

Biderman, A. D., and A. J. Reiss, Jr. (1967)
"On exploring the 'dark figure' of crime." *The Annals of the American Academy of Political and Social Science* 374:1-15.

Bishop, Y. M. M., S. E. Fienberg, and P. W. Holland (1975)
Discrete multivariate analysis: Theory and practice. Cambridge, Mass.: M.I.T. Press.

Çinlar, E. (1975)
Introduction to stochastic processes. Englewood Cliffs, N.J.: Prentice-Hall.

Fienberg, S. E. (1975)
"*Perspective Canada* as a social report." *Social Indicators Research* 2:153-174.

Fienberg, S. E., and W. Mason (1978)
"Identification and estimation of age-period-cohort models in the analysis of discrete archival data." In *Sociological methodology 1979* (K. Schuessler, ed.), pp. 1-67. San Francisco: Jossey-Bass.

Finkner, A., and H. Nisselson (1978)
"Some statistical problems associated with cross-sectional surveys." In *Survey sampling and measurement* (K. Namboodiri, ed.), New York: Academic Press.

Hood, R., and R. Sparks (1970)
Key issues in criminology. New York: McGraw-Hill.

Murphy, L. R., and C. C. Cowan (1976)
"Effects of bounding on telescoping in the National Crime Survey." *American Statistical Association Proceedings of the Social Statistics Section*, part II, pp. 633-638.

Penick, B. K., and M. E. B. Owens (eds.) (1976)
Surveying crime. Panel for the Evaluation of Crime Surveys. Washington, D.C.: National Academy of Science.

Reiss, A. J. (1980)
"Understanding changes in crime rates," this volume.

Reiss, A. J. (1977)
Personal communication.

Rubin, D. B. (1976)
"Inference and missing data." *Biometrika* 63:581-592.

Singer, B., and S. Spilerman (1974)
"Social mobility models for heterogeneous populations." In *Sociological Methodology 1973-74* (H. Costner, ed.), pp. 356-401, San Francisco: Jossey-Bass.

Singer, B., and S. Spilerman (1976a)
"The representation of social processes by Markov models." *American Journal of Sociology* 82:1-54.

Singer, B., and S. Spilerman (1976b)
"Some methodological issues in the analysis of longitudinal surveys." *Annals of Economic and Social Measurement* 5:447-474.

Thompson, M. M., and G. Shapiro (1973)
"The current population survey: An overview." *Annals of Economic and Social Measurement* 2:105-129.

U.S. Department of Justice, Law Enforcement Assistance Administration (1972)
San Jose methods test of known crime victims. Statist. Tech. Rep. No. 1. Washington, D.C.: U.S. Government Printing Office.

U.S. Department of Justice, Law Enforcement Assistance Administration (1974)
Crimes and victims: A report on the Dayton–San Jose pilot survey of victimization. Washington, D.C.: U.S. Government Printing Office.

U.S. Department of Justice, Law Enforcement Assistance Administration (1976a)
Criminal victimization in the United States: A comparison of 1973 and 1974 findings (No. SD NCP-N-3). Washington, D.C.: U.S. Government Printing Office.

U.S. Department of Justice, Law Enforcement Assistance Administration (1976b)
Criminal victimization in the United States 1973 (No. SD NCP-N-4). Washington, D.C.: U.S. Government Printing Office.

Victim proneness in repeat victimization by type of crime*

ALBERT J. REISS, JR.
Department of Sociology
Yale University

This paper explores whether victimization by crime is a random occurrence or whether some households and persons in a population are more vulnerable or prone to victimization and repeat victimization than are others. To do so, a crime-switch matrix is constructed for repeat victimizations and that distribution is compared with one expected under a simple assumption of random occurrence of repeat victimization by type of crime. Both households and their members are considered as units at risk.

The paper also explores the error structure in a crime-switch matrix. The major types and sources of error are:

(1) Errors in reporting month of occurrence of incidents and the absence of information on date of occurrence.

(2) The absence of information on occurrence of series incidents so that their occurrence may be ordered in time and in relation to nonseries incidents.

(3) Multiple reporting from members of a household.

In examining observed switches in type of crime in repeat victimization, the following are noted:

(1) Among repeat victims, given prior victimization by one of the most frequently occurring crimes (assault, personal larceny without contact, household larceny, and burglary) the most likely next victimization is by the same type of crime; for repeat victimization for the less frequently occurring crimes (motor vehicle theft, robbery, purse-snatching/pocket-picking, and rape), the most likely next victimization is the most frequently occurring of all major crimes, personal larceny without contact.

(2) Nonetheless, there is substantial household proneness to victimization by the same type of crime.

(3) Excluding proneness to repeat victimization by the same type of crime, any crime occurs about as frequently with any other type of crime over time as would be expected from the joint probability of their occurrence. Four symmetrical pairs of crimes occur less frequently than expected, however: assault with personal larceny, assault with household larceny, personal larceny with burglary, and personal larceny with household larceny. These are significant changers in a mover-stayer model.

(4) The order of occurrence of a type of crime has no effect on the probability of occurrence of a pair.

(5) The probability of consecutive victimization by the same type of crime varies directly with the probability of occurrence of that type of crime among all victims.

The problem

Considerable interest attaches to the question of whether victimization by crime is a random occurrence or whether some households and persons in a population are more vulnerable or prone to victimization and repeat victimization than are others. The question is not easily answered since theoretically one would want to model the distribution and behavior of both victims and offenders. Too little information is available at present to construct such a model. There is too little information on the distribution of offenders in a population and almost none on their networks or selection of victims. Do offenders, for example, select victims for their vulnerability to victimization? Similarly, while it is known that the risk of victimization varies considerably across territorial space, among different social aggregates, and over time, there is too little information on repeat victimization and the behavior proneness of victims.

There are two major related and competing explanations for differences in the risk of victimization and repeat victimization: victim proneness and victim vulnerability. The victim proneness explanation selects personal, social, and behavior characteristics of persons as potential victims and their relationship to offenders as explanatory variables while the victim vulnerability explanation selects situations and characteristics of offenders, their networks, behavior and relationships to potential victims as explanatory variables.

Simply put, victim proneness models explain high risk of victimization and repeat victimization by victim behavior and relationships with potential offenders that precipitates crimes or increases their vulnerability to potential offenders. The vulnerability models are more offender oriented, explaining repeat victimization in terms of such factors as the offender's prior relationship with victims, selection of criminal opportunities, and the organization of networks of offending. The considerable overlap of the models not only makes it difficult to test them as competing explanations but argues for a more general model. A more precise general model, however, depends upon yet to be acquired information on the behavior of both victims and offenders.

*This paper is based on research supported by LEAA Grant #SS-99-6013.

We can begin to answer our initial question of whether victimization by crime is a random occurrence or rather whether some persons or households in a population are more prone (or vulnerable) to victimization by crime by investigating the extent to which repeat victimization within the U.S. population conforms to a random model of occurrence. The National Crime Survey (NCS) makes it possible to examine repeat victimization and the characteristics of repeat victims to determine whether certain kinds of victims are more at risk than others. This paper limits the examination to whether in repeat victimization of a household and its members, or of households only, there is a propensity to repeat victimization by the same type of crime.

The answer to the question, by what type of crime is repeat victimization most likely to occur, given victimization by a prior type of crime, is obtained by constructing a crime-switch matrix for repeat victimizations and determining whether it conforms to an expected model of occurrence under a simple assumption of random occurrence of repeat victimization by type of crime. The crime-switch matrix consists of pairs of victimization where a preceding victimization by type of crime is compared with a following victimization by type of crime.

When the effect that victimization by a previous type of crime has on the occurrence of the next type of crime is examined in repeat victimization, the question arises whether the order in which the prior event occurs has an effect on the next reported type of crime. For example, in sequential or repeat victimization by a pair of any two types of crime such as personal larceny and robbery, is one as likely to be victimized by a personal larceny following a robbery as by a robbery following a personal larceny, given that the probability of occurrence of either a robbery or a personal larceny remains the same at both points in time? Or, is one more prone to robbery following a personal larceny than to a personal larceny following a robbery? We shall answer this question by investigating the degree of symmetry in paired victimizations by type of crime.

Risk of victimization and victim proneness

The question of what unit is at risk in victimization by any given type of crime or in repeat victimization by crime is not easily resolved. Among major types of crime measured in the NCS, a household is considered at risk for the offenses of burglary, household larceny, and auto theft while its individual members are considered at risk for offenses against the person, viz., rape, robbery, assault, larceny with contact (purse-snatching and pocket-picking), and personal larceny without contact. Yet, where each member of the household may purchase and own an automobile from individually earned income, an auto theft may symbolically represent more of a personal than a household crime. This may be particularly the case where the household member is a young person who has purchased an automobile for his journey to work and personal use. Larceny of a paycheck from the principal wage-earner of a household, on the other hand, may symbolically represent a household rather than a personal victimization. Similarly, an assault that incapacitates the principal wage-earner may be regarded as a family or a household victimization rather than a single member's victimization.

Many, though not all, households consist of families who appear to regard any household crime as a victimization in which they share. On the average any adult member of the household can report more accurately the crimes against the household than the personal crimes against all members of the household. The NCS thus uses only a single member of the household to respond to the Household Screen Questionnaire to obtain information on crimes against the household while it asks each member the questions in the Individual Screen Questionnaire to obtain information on crimes against the person. Nevertheless, the NCS provides evidence that the household respondent does not always report all of the household crimes and additional household crimes, particularly burglary and motor vehicle theft, are reported by individual members in response to the Individual Screen Questionnaire or the Crime Incident Report (Dodge 1975). These matters of how information on personal and household crimes can most accurately be obtained aside, the evidence suggests that symbolically household members often experience a household crime as a personal as well as a household victimization. Where the household is a single-person household, a not uncommon condition, it seems unlikely that the member distinguishes the household collectively from his or her individual membership in it. Household and person crimes are probably experienced as personal victimizations in single-person households.

Selection of the unit at risk in repeat victimization then is no simple matter. From one perspective the unit at risk is a household and its members; from another, it is its individual members. Yet in selecting either as the unit at risk, somewhat different assumptions are possible. Selecting the household as the unit at risk, one assumption is that it is at risk only for household and not person victimizations. An equally plausible assumption, however, is that the household as a collectivity is at risk so that household and person victimizations of its members are regarded as household victimizations. Selecting the person as the unit at risk, one assumption is that each member of the household experiences both household and person crimes as person victimizations so that household victimizations should attach to each member of the household. An equally plausible assumption perhaps is that only the person victimizations should attach to the person as the unit at risk.

Each of these assumptions can be examined as the unit at risk for repeat victimization. Refinements in the unit at risk are also possible, such as by size of household (single-person, two-person, and three-or-more-person households) or for their composition in terms of families and unrelated individuals. The preliminary inquiry reported in this paper is limited to examining the household as the unit at risk. Crime switch matrices are constructed for the household as the unit at risk using all offenses against the household and its members (Tables 2–5) and for the household using only household crimes (Tables 7–8).

Construction of the crime-switch matrices and their error structure

The crime-switch matrix is a cross-classification of pairs of crimes in the order of their occurrence in a sequence of victimizations over time. The number of pairs (p) in a sequence of two or more victimizations is one less than the number of victimizations (n), i.e., $p = (n - 1)$. There were 57,407 pairs of victimizations for households reporting sequences of two or more person and household victimizations and 16,884 pairs of victimizations for households reporting two or more household-only victimizations.

All incidents for households reporting two or more victimizations between July 1, 1972, and December 31, 1975, are included in our matrices. Since households and their members were in sample for varying lengths of time—from one to seven interviews—the probability of repeat victimization depends upon the length of time in sample. This fact is not taken into account in the construction of the matrices. Both bounded and unbounded incident data are included.

Respondents initially define crime events in response to stimulus items in the Basic Screen Questionnaire. For each incident reported in the Basic Screen Questionnaire, a Crime Incident Report is completed that includes a large number of facts about the actual event. These facts are used to classify events into particular types of crime victimization. Although there undoubtedly are some errors in respondent reporting, it is doubtful that they materially affect the classification by type of crime. When such errors occur, they should produce more error in detailed types than in major types of crime classifications.

Reverse-record check studies show that errors occur in month of reporting the incident. These errors can have an impact on the ordering of events in time and thus affect the order of types of crimes in a pair in the crime-switch matrix. The procedure followed by the NCS, however, reduces the likelihood of these errors having a substantial effect since Crime Incident Reports are taken in the order of occurrence of events, beginning with the "first" incident.

The Crime Incident Report of the NCS records only the month and not the date of reported occurrence of a nonseries victimization. For series incidents, the month in which the first incident in the series is reported as occurring is recorded; when that month lies outside the reference period for which crime incidents are being recalled, the series is recorded as beginning in the sixth month of the reference period or that farthest from the time of interview. The seasons of the year during which the series incidents took place also are recorded. When more than one incident is reported taking place within the same month, they are not ordered by time of occurrence within the month.

These survey questionnaire procedures make it difficult to order victimizations precisely. Where there is multiple victimization, crime incidents can only be ordered by month of occurrence and not within the month. Moreover, the incidents in a series victimization cannot be ordered by time of occurrence. Where both series and nonseries victimization are reported, series and nonseries incidents cannot be ordered by time of occurrence. Given these survey limitations, a set of procedures was adopted for this report that orders victimizations in time.[1] These procedures are described below.

First, using a random number generator, a date of occurrence was assigned to each nonseries incident within the month of the reference period for which it was reported. Where there was more than one incident reported within a given month, the assignment of the date of occurrence was independent for each event.

Second, a series incident was treated as only a single incident. The random number generator was used to assign a date of occurrence within the first reported month of occurrence of the series incident within the reference period. The type of crime assigned to the series incident was that reported for the most recent incident in the series since that is the only incident for which type of crime is reported.

Third, merging all person and household series and nonseries incidents, the incidents were first ordered for each household and then reordered for household incidents only. The first ordering produced a household crime-switch matrix using all offenses against the household and its members and the second a household crime-switch matrix for offenses against the household only.

[1] Other procedures may be adopted in later work. For example, for series incidents given the month of first reported occurrence of a series incident, the span of months (from seasonal data) for which they occur can be determined. Taking an estimate of the number of incidents in the series, we can then estimate an average time between series incidents. If we randomly assign dates of occurrence to these series incidents, they can then be merged with nonseries incidents.

This procedure produces an unknown amount of error in ordering events within a month since the order of events *within* a month is determined from the randomly assigned dates of occurrence.

Excluding all but one incident of a series also produces errors of two types. First, excluding series reduces the diagonal cells in a crime-switch matrix, i.e., the "stayers" in a mover-stayer model. A substantial proportion of all series victimizations would fall on the diagonal since on first reporting any victimizations, 76 percent of all persons and 71 percent of all households report only series victimization within the 6–7 month period. Not all cells of the diagonal would be affected equally since series victimization varies by type of crime. Second, in the remaining cases where both series and nonseries or two or more series victimizations are reported, there are misclassifications in preceding and following types of crime in order of occurrence. Such errors affect only the "changer" cells in a mover-stayer model. On balance there is far more of the first type than of the second type of error since the number of events in a series is large relative to reporting of nonseries events within the same reference period and series only reporting predominates over series and nonseries reporting within the same reference period. Were all series to be included, then, one would expect even greater victim proneness (or stayer propensity) in our observed model.

Table 1 presents the percent distribution by type of crime for all first and last reported crime incidents where the household and its members reported one or more crime incidents and the first and second incidents of a pair for all sequences of two or more victimizations. First and last reported crime incidents include the *same* crime incident when only a single victimization is reported but all *first* reported victimizations also include the *first* reported incident in a sequence of victimizations and all *last* reported incidents include the *last* incident in a sequence. Where there are only two crime incidents in a sequence of victimizations, the first (preceding) and second (following) incidents may be of the same or different type of crime. In sequences of three or more victimizations, however, the second incident of a first pair becomes the first incident of the next pair in a crime-switch matrix, and so on. Thus, *all* incidents in a sequence other than the *last* will appear as a *first* incident in a pair and all but the *first* incident in a sequence will appear as the *second* incident in a pair.

Table 1. Percent of each type of crime by order of reporting victimizations of households and their members: All household and person victimizations, July 1, 1972, to December 31, 1975

Type of crime	Reported as first incident		Reported as last incident		Reported as preceding incident of a pair		Reported as following incident of a pair		All reported incidents*	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Rape	63	0.1	67	0.1	77	0.1	81	0.1	144	0.1
Attempted rape	161	0.3	161	0.3	196	0.3	197	0.3	358	0.3
Serious assault	959	1.7	932	1.7	1,346	2.3	1,316	2.3	2,279	2.0
Minor assault	1,028	1.9	994	1.8	1,473	2.6	1,437	2.5	2,471	2.2
Attempted assault	3,646	6.6	3,775	6.9	5,767	10.1	5,892	10.3	9,558	8.5
Robbery	653	1.2	594	1.1	799	1.4	741	1.3	1,395	1.2
Attempted robbery	429	0.8	384	0.7	649	1.1	606	1.1	1,037	0.9
Purse snatch/pocket picking	752	1.4	715	1.3	610	1.1	573	1.0	1,328	1.2
Attempted purse snatch	120	0.2	115	0.2	93	0.2	87	0.1	209	0.2
Personal larceny, $50 and over	5,481	10.0	5,525	10.0	5,054	8.8	5,100	8.9	10,613	9.4
Personal larceny, under $50	14,712	26.7	14,658	26.6	15,999	27.9	15,941	27.8	30,732	27.3
Attempted personal larceny	1,428	2.6	1,483	2.7	1,463	2.6	1,517	2.6	2,954	2.6
Burglary, forced entry	3,357	6.1	3,331	6.1	3,026	5.3	2,995	5.2	6,362	5.7
Burglary, no force	4,646	8.4	4,689	8.5	4,375	7.6	4,417	7.7	9,074	8.1
Attempted burglary	2,300	4.2	2,257	4.1	2,194	3.8	2,153	3.8	4,458	4.0
Household larceny, $50 and over	3,350	6.1	3,307	6.0	3,311	5.8	3,269	5.7	6,621	5.9
Household larceny, under $50	8,997	16.4	9,046	16.4	8,234	14.3	8,271	14.4	17,286	15.3
Attempted household larceny	858	1.6	886	1.6	827	1.4	854	1.5	1,713	1.5
Motor vehicle theft	1,331	2.4	1,403	2.6	1,151	2.0	1,223	2.1	2,562	2.3
Attempted motor vehicle theft	730	1.3	704	1.3	763	1.3	737	1.3	1,469	1.3
Total	55,001	100.0	55,026	100.0	57,407	100.0	57,407	100.0	112,623	100.0

*The entries for "Reported as first incident" and "Reported as following incident of a pair" do not always sum to the Total, "All reported incidents," for each type of crime; there are 215 incidents for which month of occurrence is unreported.

Table 2a.

Type of crime reported as preceding incident	Rape	Att. rape	Serious assault	Minor assault	Att. assault	Robbery	Att. robbery	Purse snatch, pocket picking
Rape	6	3	8	1	5	2	2	1
Attempted rape	1	16	6	9	21	1	6	2
Serious assault	2	11	119	75	212	37	19	18
Minor assault	4	8	83	152	244	22	20	21
Attempted assault	11	29	199	228	1,685	59	81	37
Robbery	2	4	50	30	76	89	27	23
Attempted robbery	1	5	30	21	72	27	54	12
Purse snatch/pocket picking	1	2	21	19	47	26	9	46
Attempted purse snatch	-	-	3	2	10	2	3	1
Personal larceny, $50+	7	15	110	99	401	61	44	44
Personal larceny, under $50	16	33	238	292	1,314	126	136	144
Attempted personal larceny	1	3	25	34	115	22	24	15
Burglary, forced entry	7	17	90	59	199	42	25	27
Burglary, no forcible entry	9	11	85	107	319	45	34	41
Attempted burglary	-	8	46	46	166	30	15	23
Household larceny, $50+	4	8	62	61	238	43	24	30
Household larceny, under $50	7	18	89	154	554	75	49	64
Attempted household larceny	1	4	11	8	74	4	11	3
Motor vehicle theft	1	1	29	24	78	20	11	16
Attempted motor vehicle theft	-	1	12	16	62	8	12	5
Total number	81	197	1,316	1,437	5,892	741	606	573

One would expect, therefore, considerable similarity in the type of crime distributions for the first and second incidents in pairs of crime incidents since they can differ only in the distributions for first and last incidents in a sequence of victimizations. Any selective influence in proneness to type of crime for multiple or repeat victimization of a household and its members will be reflected largely in differences between the distribution for all first reported incidents and that for the first incident of a pair or the distribution for last reported incidents and that for the last incident of a pair or in differences with the distribution for all reported victimizations.

What is most apparent in Table 1 is the striking similarity in the distributions by type of crime for all first and last reported and all preceding and following incidents of a pair with that for all reported incidents. Although this similarity is determined to some degree by the relative frequency of types of crime among all victimizations, it is apparent that *multiple or repeat victims of households are generally no more or less prone to victimization by certain types of crime than are all victims.* There are some exceptions. Assaults, particularly attempted assaults, occur more frequently among repeat than among all victims and by deduction they occur less commonly among persons reporting single victimizations. Actual burglaries fit the opposite pattern, being less common among repeat than all victims. In the aggregate, crimes against the household occur only somewhat less frequently among repeat than among all victims while the opposite is the case for crimes against household members, i.e., crimes against persons.

The degree to which there are switches in victimization by type of crime for a household and its members in multiple or repeat victimization depends in part upon the system of classification for crime incidents. The procedure followed was to classify each crime event by the most serious offense reported. A rape with theft, for example, was classified only as a rape and not as two separate incidents of a rape and theft. Sequences of incidents that occur within the same crime event are thus excluded from the crime-switch matrix.

NCS survey procedures separate the reporting and classification of household and person incidents. Whenever a single event of victimization involves both a household and a person crime, e.g., a burglary followed by a rape, they are reported as two separate incidents. Our procedure will err in assigning them as separate events in time. Moreover, NCS procedures take a separate incident report from each member of the household who was "robbed, harmed, or threatened" in the same event. Since these incidents are identified only by month of occurrence, our procedure again will err in assigning them as separate events in time. Although members to a common event can be classified in a separate detailed crime category, e.g., an assault on one and an attempted assault on the others, in general such events will contribute disproportionally to the diagonal cells of the crime-switch matrix. Relative to the aggregate of all multiple victimization over time, these events are relatively uncommon so that the amount of error should be small. These sources of error could be eliminated were NCS procedures to clearly separate incidents that occur in the same event in time from those that do not and also identify members of the household to a common event.

Number of crime incident pairs of preceding and following detailed major types of crime reported by a household and its members: All households reporting two or more victimizations while in survey, July 1, 1972, to December 31, 1975

Type of crime reported as following incident

Att. purse snatch	Personal larceny $50+	Personal larceny under $50	Att. personal larceny	Burglary, force	Burglary, no force	Att. burglary	Household larceny $50+	Household larceny under $50	Att. household larceny	Motor vehicle theft	Att. motor vehicle theft	Total number
-	7	16	-	5	2	4	2	9	1	3	-	77
3	11	43	5	7	8	13	7	27	2	5	3	196
2	100	233	26	78	92	53	68	145	16	32	8	1,346
4	93	315	30	64	83	32	82	163	16	25	12	1,473
3	371	1,244	141	196	301	184	230	557	72	87	52	5,767
1	76	130	21	41	54	23	41	78	5	15	13	799
-	55	156	21	26	37	16	27	63	7	13	6	649
4	48	141	22	38	42	26	30	48	8	17	15	610
10	9	18	5	4	4	1	3	10	2	6	-	93
9	859	1,385	173	231	337	159	317	544	55	133	71	5,054
13	1,428	7,067	474	436	960	363	637	1,749	158	247	168	15,999
4	156	446	149	53	72	47	68	142	19	35	33	1,463
4	212	450	53	666	237	198	183	394	53	79	31	3,026
7	354	917	74	258	993	134	253	550	58	88	38	4,375
-	151	395	43	192	172	360	116	327	39	33	32	2,194
4	350	636	65	179	261	121	475	552	68	82	48	3,311
14	544	1,779	133	378	597	297	555	2,526	171	149	81	8,234
2	65	161	24	33	49	47	62	164	73	16	15	827
2	134	253	25	74	76	48	74	127	8	122	28	1,151
1	77	156	33	36	40	27	39	96	23	36	83	763
87	5,100	15,941	1,517	2,995	4,417	2,153	3,269	8,271	854	1,223	737	57,407

The crime-switch matrix in Tables 2–5 presents shifts in type of crime for repeat or multiple victimization of households reporting two or more victimizations between July 1, 1972, and December 31, 1975. The matrix is presented in two different ways. Table 2 presents the frequency distribution for the crime-switch matrix. Table 3 displays the percent of each type of crime that follows a preceding reported type of crime in a sequence of eight major types of crimes against households and persons.

The reader should bear in mind that these reports of victimization are for a household and its members. Where crimes against persons are involved, the two incidents in a pair might be for the same *or* for different members of the *same* household. Thus, a robbery followed by robbery could be reported for the same or for different members of the same household. Unfortunately, we are unable to distinguish victims *within* the same household to a *common* event, e.g., an assault, and our procedure for ordering events will treat them as separate victimizations occurring at different, though closely related, points in time. There is, therefore, a confounding of some preceding and following events in the crime-switch matrix. The reader is cautioned, in any case, against interpreting the ordering of personal crime incidents as applying to the same person.

Observed switches in type of crime in repeat victimization of a household and its members

There is considerable mutliple and repeat victimization by crime. The longer the time interval for which victimization is measured, the greater the propensity to repeat victimization.

Given the distribution of all reported victim incidents or of that for repeat victims only (Table 1), the most likely victimization by a major type of crime is a personal larceny (39 percent) with household larceny (22 percent), burglary (17 percent) and assault (15 percent) among the next most likely events. The other major crimes against persons or households occur much less frequently: motor vehicle theft (3 percent); robbery (2 percent); purse-snatching or pocket-picking (1 percent); and the least frequent, rape (0.4 percent). Yet, *among repeat victims, given prior victimization by one of the most frequently occurring crimes (assault, personal larceny without contact, household larceny, and burglary), the most likely next victimization is by the same type of crime.* Roughly a third of all assaults and of all burglaries, 38 percent of all household larcenies and 54 percent of all personal larcenies are preceded and followed by reports of victimization by that same type of crime (Tables 2 and 3). *In repeat victimization for the less frequently occurring crimes (motor vehicle theft, robbery, purse-snatching or pocket-picking, and rape), the most likely next victimization is the most frequently occurring crime of personal larceny without contact.* Roughly 3 in 10 of these less frequently occurring crimes are followed by a personal larceny (Table 2).

Table 2b.

Type of crime preceding next reported type of crime	Rape	Att. rape	Serious assault	Minor assault	Att. assault	Robbery	Att. robbery	Purse snatch, pocket picking
Rape	7.8	3.9	10.4	1.3	6.5	2.6	2.6	1.3
Attempted rape	0.5	8.2	3.1	4.6	10.7	0.5	3.1	1.0
Serious assault	0.2	0.8	8.8	5.6	15.8	2.8	1.4	1.3
Minor assault	0.3	0.5	5.6	10.3	16.6	1.6	1.4	1.4
Attempted assault	0.2	0.5	3.5	4.0	29.2	1.0	1.4	0.6
Robbery	0.3	0.5	6.3	3.7	9.5	11.1	3.4	2.9
Attempted robbery	0.2	0.8	4.6	3.2	11.1	4.2	8.3	1.9
Purse snatching/pocket picking	0.2	0.3	3.4	3.1	7.7	4.3	1.5	7.5
Attempted purse snatching	-	-	3.2	2.1	10.8	2.1	3.2	1.1
Personal larceny, $50 and over	0.1	0.3	2.2	2.0	7.9	1.2	0.9	0.9
Personal larcency, under $50	0.1	0.2	1.5	1.8	8.2	0.8	0.9	0.9
Attempted personal larceny	0.1	0.2	1.7	2.3	7.9	1.5	1.6	1.0
Burglary, forced entry	0.2	0.6	3.0	2.0	6.6	1.4	0.8	0.9
Burglary, no forcible entry	0.2	0.2	1.9	2.4	7.3	1.0	0.8	0.9
Attempted burglary	-	0.4	2.1	2.1	7.6	1.4	0.7	1.0
Household larceny, $50 and over	0.1	0.2	1.9	1.8	7.2	1.3	0.7	0.9
Household larceny, under $50	0.1	0.2	1.1	1.9	6.7	0.9	0.6	0.8
Attempted household larceny	0.1	0.5	1.3	1.0	9.0	0.5	1.3	0.4
Motor vehicle theft	0.1	0.1	2.5	2.1	6.8	1.7	1.0	1.4
Attempted motor vehicle theft	-	0.1	1.6	2.1	8.1	1.1	1.6	0.7
All next reported incidents	0.1	0.3	2.3	2.5	10.3	1.3	1.1	1.0
Number next reported incidents	81	197	1,316	1,437	5,892	741	606	573
No incidents reported								

Table 3. Percent distribution by next reported type of crime for major preceding types of crime reported by households with two or more victimizations of the household and its members: All household and person victimizations, July 1, 1972, to December 31, 1975

Type of crime[1] preceding next reported crime	Next reported type of crime[1]								Total	
	Rape	Assault	Robbery	Purse snatch, pocket picking	Personal larceny	Burglary	Household larceny	Motor vehicle theft	Percent	Number
Rape	9.5	18.3	4.0	2.2	30.0	14.3	17.6	4.0	99.9	273
Assault	0.8	34.9	2.8	1.0	29.7	12.6	15.7	2.5	100.0	8,586
Robbery	0.8	19.3	13.6	2.5	31.7	13.6	15.3	3.2	100.0	1,448
Purse snatching, pocket picking	0.4	14.5	5.7	8.7	34.6	16.4	14.4	5.4	100.1	703
Personal larceny	0.3	11.7	1.8	1.0	53.9	11.8	16.4	3.1	100.0	22,516
Burglary	0.5	11.6	2.0	1.1	27.6	33.5	20.6	3.1	100.0	9,595
Household larceny	0.3	10.1	1.7	1.0	30.4	15.9	37.6	3.2	100.1	12,372
Motor vehicle theft	0.2	11.6	2.7	1.2	35.4	15.7	19.2	14.0	100.0	1,914
Total	0.5	15.1	2.3	1.1	39.3	16.7	21.6	3.4	100.0	57,407

[1] All types of crime include both actual and attempted crimes.

There is considerable variation in patterns of repeat victimization for the different major types of crime, however, as the following summary based on Tables 2 and 3 discloses.

(1) *Despite the fact that rape is an infrequent event accounting for only 0.4 percent of all reported victimizations, 9 percent of all rapes occurring in households reporting two or more victimizations were followed by a rape.* There appear to be differences in victim proneness to actual and attempted rape. Most rapes preceded or followed by an attempted rape are also attempted rapes, and a substantial majority of rapes preceded or followed by an actual rape are actual rapes.

Since rape is a form of assault, some interest attaches to victim proneness to rape and other types of assault (serious, minor, and attempted assault). *Somewhat more rapes than expected in households reporting two or more victimizations were*

Percent distribution by next reported type of crime for major preceding types of crime reported by households with two or more victimizations of the household and its members: All household and person victimizations, July 1, 1972, to December 31, 1975

Next reported type of crime												
Att. purse snatch	Personal larceny, $50+	Personal larceny, under $50	Att. personal larceny	Burglary, force	Burglary, no force	Att. burglary	Household larceny $50+	Household larceny under $50	Att. household larceny	Motor vehicle theft	Att. motor vehicle theft	Total
-	9.1	20.8	-	6.5	2.6	5.2	2.6	11.7	1.3	3.9	-	100.1
1.5	5.6	21.9	2.6	3.6	4.1	6.6	3.6	13.8	1.0	2.6	1.5	100.1
0.2	7.4	17.3	1.9	5.8	6.8	3.9	5.0	10.8	1.2	2.4	0.6	100.0
0.3	6.3	21.4	2.0	4.3	5.6	2.2	5.6	11.1	1.1	1.7	0.8	100.0
0.1	6.4	21.6	2.4	3.4	5.2	3.2	4.0	9.7	1.2	1.5	0.9	100.0
0.1	9.5	16.3	2.6	5.1	6.8	2.9	5.1	9.8	0.6	1.9	1.6	100.0
-	8.5	24.0	3.2	4.0	5.7	2.5	4.2	9.7	1.1	2.0	0.9	100.1
0.7	7.9	23.1	3.6	6.2	6.9	4.3	4.9	7.9	1.3	2.8	2.5	100.1
10.8	9.7	19.3	5.4	4.3	4.3	1.1	3.2	10.8	2.1	6.5	-	100.0
0.2	17.0	27.4	3.4	4.6	6.7	3.1	6.3	10.8	1.1	2.6	1.4	100.1
0.1	8.9	44.2	3.0	2.7	6.0	2.3	4.0	10.9	1.0	1.5	1.0	100.0
0.3	10.7	30.5	10.2	3.6	4.9	3.2	4.7	9.7	1.3	2.4	2.3	100.1
0.1	7.0	14.9	1.8	22.0	7.8	6.5	6.0	13.0	1.8	2.6	1.0	100.0
0.2	8.1	21.0	1.7	5.9	22.7	3.1	5.8	12.6	1.3	2.0	0.9	100.0
-	6.9	18.0	2.0	8.7	7.8	16.4	5.3	14.9	1.8	1.5	1.4	100.0
0.1	10.6	19.2	2.0	5.4	7.9	3.7	14.3	16.7	2.0	2.5	1.5	100.0
0.2	6.6	21.6	1.6	4.6	7.2	3.6	6.7	30.7	2.1	1.8	1.0	100.0
0.2	7.9	19.5	2.9	4.0	5.9	5.7	7.5	19.8	8.8	1.9	1.8	100.0
0.2	11.6	22.0	2.2	6.4	6.6	4.2	6.4	11.0	0.7	10.6	2.4	99.9
0.1	10.1	20.5	4.3	4.7	5.2	3.5	5.1	12.6	3.0	4.7	10.9	100.0
0.1	8.9	27.8	2.6	5.2	7.7	3.8	5.7	14.4	1.5	2.1	1.3	100.0
87	5,100	15,941	1,517	2,995	4,417	2,153	3,269	18,271	854	1,223	737	57,407

followed (18 percent) by an assault. Of some interest also is the fact that actual rapes were far more likely than expected to occur in households that next reported a serious assault (10 percent).

(2) *Repeat victims are only slightly more prone to assault than are victims reporting only a single victimization.* For households reporting two or more victimizations, 15 percent of all victimizations were an assault while about 13 percent of all single victims reported an assault.

There is substantial proneness to assault among victims reporting an assault. Given the occurrence of an assault on a household member, 35 percent of all preceding and following incidents reported for the household are an assault. Serious and minor assaults were about as likely to be followed by a serious or minor as by an attempted assault, but attempted assaults followed by an assault are ordinarily followed by an attempted assault.

(3) *Repeat victims are only slightly more prone to robbery than are single-time victims.* Of all reported victimizations, 2.1 percent are robberies while 2.4 percent of all incidents reported by multiple victimized households are robberies.

There is substantial proneness to robbery among victims of robbery. Given the relatively low probability of a robbery, a substantially greater proportion (14 percent) of all robberies reported against household members in households with two or more victimizations were preceded or followed by a robbery of a member. Moreover, actual robberies are generally preceded and followed by actual robberies and attempted by attempted robberies.

Among repeatedly victimized households reporting robbery, their risk of repeat victimization by robbery is less than their risk of victimization by the major crimes of personal larceny without contact (32 percent) or assault (19 percent), but about the same as that for the much more frequently occurring crimes of burglary and household larceny.

(4) Purse-snatching and pocket-picking are relatively infrequent events accounting for only 1.4 percent of all crime incidents reported in the victim survey and an even smaller proportion of those for multiple victimized households (1.1 percent). Yet, much as in the case for the infrequent event of rape, *there is considerable proneness to repeat victimization among those victimized by purse-snatching or pocket-picking;* 9 percent of all purse-snatching or pocket-picking incidents were followed by some member of the household reporting the same kind of victimization. There is a strong tendency for reports of *actual* to be followed by *actual* purse-snatching or pocket-picking and *attempted* to be followed by *attempted* purse-snatching or pocket-picking.

Among repeat victim households reporting a purse-snatching or pocket-picking by one of their members, the risk of victimization is greatest for personal larceny without contact (35 percent), with burglary, household larceny, or assault being equally likely as next reported events.

(5) *There is considerable victim proneness for victims of personal larceny without contact.* Although personal larceny without contact is the most frequently occurring major crime against a household and its members accounting for 39 percent of all victimizations, 54 percent of all personal

Table 4.

Type of crime reported as preceding incident	Rape and att. rape	Serious assault	Minor assault	Att. assault	Robbery
Rape and attempted rape	*+461*	+10	+1	–*	–*
Serious assault	+6	*+252*	+51	+40	+22
Minor assault	+3	+72	*+360*	+ 57	+*
Attempted assault	–5	+34	+49	*+2019*	–3
Robbery	+1	+55	+5	–*	*+600*
Attempted robbery	+3	+15	+1	+*	+41
Purse snatch/pocket picking and attempts	–*	+4	+1	–3	+39
Personal larceny, $50 and over	–*	–*	–6	–27	–*
Personal larceny, under $50	–11	–45	–29	–65	–32
Attempted personal larceny	–1	–2	–*	–8	+*
Burglary, forced entry	+6	+6	–4	–40	+*
Burglary, no force	–*	–2	–*	–38	–2
Attempted burglary	–1	–*	–1	–16	+*
Household larceny, $50 and over	–1	–3	–6	–31	+*
Household larceny, under $50	–6	–53	–13	–100	–9
Attempted household larceny	+*	–3	–8	–1	–4
Motor vehicle theft	–2	+*	–1	–14	+3
Attempted motor vehicle theft	–2	–2	–*	–3	–*
Rape and attempted rape	*2.6*				
Serious assault		*1.4*			
Minor assault			*2.1*		
Attempted assault				*11.6*	
Robbery					*3.4*
Attempted robbery					
Purse snatch/pocket picking and attempts					
Personal larceny, $50 and over					
Personal larceny, under $50					
Attempted personal larceny					
Burglary, forced entry					
Burglary, no force					
Attempted burglary					
Household larceny, $50 and over					
Household larceny, under $50					
Attempted household larceny					
Motor vehicle theft					
Attempted motor vehicle theft					

χ^2 = 17,461; d.f. = 289
* = less than .5
\+ = Observed $>$ expected frequency; – = $<$ expected frequency.
Goodman-Kruskal Index (G) .20 ± .004.

larcenies without contact were followed by that same type of crime. Major ($50 or more), minor (under $50) and attempted personal larcenies without contact are all most likely to be followed by a minor personal larceny without contact in repeat larceny victimization. Nonetheless, a major personal larceny is more than twice as likely to be followed by another major personal larceny as would be expected from the risk of major personal larceny for all victims.

Minor personal larcenies show the greatest propensity to repeat victimization by minor personal larceny and are less likely than the other types of personal larceny without contact (major and attempted) to be followed by an attempted or major personal larceny. This propensity perhaps reflects differences in the victim composition of larceny victims by type of larceny. A substantial number of minor personal larcenies involve school-age victims where repeat victimization involves minor personal larcenies at school.

(6) *There is likewise considerable victim proneness among victims of burglary.* About a third of all burglaries of households reporting two or more victimizations were preceded or followed by a burglary.

Attempted burglaries were about as likely to be followed by an actual as an attempted burglary but actual burglaries with or without force were quite likely to be

Actual and percent cell contribution to chi square of pairs of victimizations in a crime-switch matrix for repeat household victims of detailed person and household crimes: All household and person victimizations for households reporting two or more victimizations, July 1, 1972, to December 31, 1975

Type of crime reported as following incident

Att. robbery	Purse snatch, pocket picking and att.	Personal larceny, $50+	Personal larceny, under $50	Att. personal larceny	Burglary, forced entry	Burglary, no force	Att. burglary	Household larceny, $50+	Household larceny, under $50	Att. household larceny	Motor vehicle theft	Att. motor vehicle theft
+9	+3	−2	−4	−1	−*	−6	+4	−3	−*	−*	+1	−*
+2	+1	−3	−53	−3	+1	−1	+*	−1	−12	−1	+*	−5
+1	+4	−11	−22	−2	−2	−8	−10	−*	−11	−2	−1	−2
+7	−10	−39	−80	−1	−37	−46	−5	−29	−90	−2	−11	−7
+41	+24	+*	−38	−*	−*	−1	−2	−*	−12	−4	−*	+1
+325	+3	−*	−3	+1	−2	−3	−3	−3	−10	−1	−*	−1
+3	*+347*	−*	−7	+4	+1	−1	+*	−1	−18	−*	+4	+4
−2	−*	*+374*	−*	+12	−4	−7	−5	+3	−47	−5	+6	+4
−6	−4	+*	*+1550*	+6	−190	−60	−94	−82	−134	−27	−26	−7
+5	−*	+5	−4	*+315*	−7	−15	−1	−3	−22	−*	+*	+11
−1	−*	−12	−181	−9	*+1635*	+*	+63	+1	−4	+1	+3	+2
−3	−*	−3	−73	−15	+4	*+1280*	−5	+*	−10	−1	−*	−6
−3	−*	−10	−75	−4	+52	+*	*+937*	−1	+*	+1	−4	+*
−3	−*	+11	−87	−6	+*	+*	−*	*+435*	+12	+7	+2	+1
−16	−3	−48	−113	−33	−6	−2	−*	+16	*+1513*	+19	−4	−6
+1	−2	−1	−20	+*	−2	−3	+8	+5	+17	*+300*	−*	+2
−*	+2	+10	−14	−1	+3	−2	+1	+1	−9	−5	*+387*	+12
+2	−1	−1	−15	+8	−*	−6	−*	−*	−2	+12	+24	*+547*
1.9												
	2.0											
		2.1										
			8.9		1.1							
				1.8								
			1.0		*9.4*							
						7.3						
							5.4					
								2.5				
									8.7			
										1.7		
											2.2	
												3.1

Table 5. Actual and percent cell contribution of chi square of pairs of victimizations in a crime-switch matrix for repeat household victims of major household and person crimes: All household and person victimizations for households reporting two or more victimizations, July 1, 1972, to December 31, 1975

Type of crime reported as preceding incident	Type of crime reported as following incident							
	Rape	Assault	Robbery	Purse snatching, pocket picking	Personal larceny without contact	Burglary	Household larceny	Motor vehicle theft
Rape	*+460*	+2	+3	−3	−6	−1	−2	−*
Assault	+13	*+2,246*	+7	−2	−200	−84	+137	−20
Robbery	+4	+17	*+782*	+22	−21	−8	−27	−*
Purse snatching, pocket picking	−*	−*	+33	*+346*	−4	−*	−17	+8
Personal larceny, without contact	−11	−172	−25	−3	*+1,223*	−319	−283	−9
Burglary	+1	−74	−5	−1	−334	*+1,624*	−5	−2
Household larceny	−5	−201	−25	−4	−251	−5	*+1,460*	−2
Motor vehicle theft	−4	−16	+1	+*	−7	−1	−5	*+635*
	Percent of total chi square							
Rape	*4.1*							
Assault		*20.1*			1.8		1.2	
Robbery			*7.0*					
Purse snatching, pocket picking				*3.1*				
Personal larceny, without contact		1.5			*10.9*	2.8	2.5	
Burglary					3.0	*14.5*		
Household larceny		1.8			2.2		*13.0*	
Motor vehicle theft								*5.7*

χ^2 = 11,190; d.f. = 49; Goodman-Kruskal Index (G) = .26± .005.
*Less than .5
+ = Observed > expected frequency

followed by the same kind of actual burglary. Thus, 22 percent of all actual burglaries with force were followed by an actual burglary with force compared with 8 percent followed by burglary with no force and 7 percent followed by attempted burglary. Similarly, 23 percent of all actual burglaries without force were followed by an actual burglary without force compared with 6 percent followed by an actual burglary with force and 3 percent followed by attempted burglaries.

(7) *Households victimized by household larceny are prone to repeat victimization by household larceny.* While 22 percent of all households reported a household larceny, 38 percent of all households victimized by household larceny reported a preceding or following incident of household larceny. As for personal larceny, any household larceny followed by a household larceny is most likely to be a minor household larceny with a loss of $50 or less. But a substantially greater proportion of major household larcenies ($50 and over) are followed by a major household larceny than expected. A similar pattern holds for attempted household larcenies.

(8) *Motor vehicle theft victims also show a propensity to repeat victimization by the same type of crime.* This propensity holds for both attempted and actual motor vehicle theft. As for other infrequently occurring crimes, motor vehicle theft is most likely to be preceded or followed by a personal larceny without contact but the other major types of crimes are about as likely events as is motor vehicle theft in their repeat victimization.

This brief description of victim proneness in repeat victimization of a household and its members has emphasized the proneness to repeat victimization by the same type of crime despite substantial differences in the probability of victimization among the types of crime. The test of proneness has been the departure of these patterns from expected frequency of occurrence, given the distribution of types of crime among all repeat victimized households. The test of divergence between the actual crime-switch matrix and an expected one based on the types of crime reported by repeat victims used was the chi-square test for homogeneity of proportions. The chi-square test for an 18 by 18 detailed type of crime matrix is presented in Table 4 and for an 8 by 8 major type of crime matrix in Table 5.

Given the substantial number of pairs, 57,407, in the matrices, the chi-square is expected to be significant. Our attention in these tables, therefore, focuses on the cells in the crime-switch matrix that contribute significantly to the chi-squared.

The diagonal cells in Tables 4 and 5 are those where the same type of crime is reported in a pair. Each of these cells contributes significantly to the value of chi-squared and they are the *only* cells in these tables where the observed frequency is *substantially greater* than the expected frequency. Indeed, the diagonal cells in Table 4 account for 78.1 percent of the total value of chi-squared and in Table 5, for 78.4 percent. We can conclude then that *the chances a household or one or more of its members will be successively victimized by the same type of crime are greater than one would expect given the chance all households and their members have of being victimized by that type of crime.* By this measure, then, *in repeat victimization, there is household victim proneness to the same type of crime.* The chances of next being victimized by a robbery, for example, are greater if the previous victimization was a robbery than if it was some other type of crime.

Readers should take note that proneness to victimization by the same type of crime does not predict the most likely next victimization by a major type of crime. That is, proneness to victimization by the same type of crime does not say that whenever a household or one of its members is victimized by any major type of crime, it has a greater chance of next being victimized by the same rather than by some other type of crime. The chances that a victim of a robbery or a motor vehicle theft will next be victimized by a personal larceny are greater, for example, than victimization by the same type of crime. Rather, for any major type of crime, the chances of next being victimized by any major type of crime are greater than chance only for the *same* type of crime.

There are eight cells in Table 5 where the observed frequency of pairs is significantly *less* than expected, given the type of crime distribution for all multiple victimized households and their members. These cells account for an additional 16.8 percent of the total value of chi-squared or all but 5 percent of the remaining variance. They involve selected combinations among the four most frequently occurring crimes of personal and household larceny, burglary, and assault. The eight cells can be regarded as four symmetrical pairs with *(1) assault with personal larceny, (2) assault with household larceny, (3) personal larceny with burglary, and (4) personal larceny with household larceny occurring less frequently than expected.* Thus an assault, for example, is less likely to be followed by a personal larceny than expected and a personal larceny is less likely to be followed by an assault than expected. Just why multiple victimized households and their members should be less prone to these four symmetrical patterns of victimization is not apparent.

All other pairs of victimization by type of crime occur about as often as expected. A rape followed by an assault or an assault followed by a rape, for example, occurs about as often as expected.

When examining the effect that victimization by a type of crime has on the chances of victimization by any other type of crime in repeat victimization, the question arises whether the *order* of victimization by a type of crime has an effect. Such effects could be presumed to occur were victims to alter substantially their behavior to avoid repeat victimization by a serious type of crime, and take precautions that reduce the likelihood of victimization by even less serious types of crime which would not be taken were one previously victimized by the less serious crime. Thus, if one or more household members were victimized by robbery, the propensity to stay home might reduce the risk of it being followed by a robbery.

An answer to the question of the effect of prior or subsequent victimization by type of crime is provided by examining the extent of symmetry in paired victimizations by type of crime. When any two types of crime are paired by their order of occurrence, the pairs are symmetrical if the probability of their occurrence is the same and it differs from other crime pairs. Even if the full conditions of symmetry are not met but the probability of occurrence remains the same for paired orders, there would be little evidence that prior victimization has had a substantial effect on type of crime victimization *in repeat victimization.*

Among the 12 person and 8 household detailed types of crime in Table 2 that can be reported as victimizations by households and their members reporting two or more victimizations, there are 400 possible pairs of incidents by type of crime when each incident is paired with the next succeeding incident by type of crime. Twenty of these pairs occur when the *same* type of crime is reported for two consecutive incidents. The remaining 380 pairs of consecutive incidents could comprise 190 symmetrical pairs. Similarly, among the five person and three household major types of crime in Table 3, there are 64 possible pairs of incidents; eight of these occur when the same type of crime is reported for two consecutive incidents. The remaining 56 pairs of consecutive incidents could comprise 28 symmetrical pairs. When the 400 and the 64 pairs are ordered by their actual rate of occurrence among 10,000 pairs among households reporting two or more victimizations, *the order of occurrence of a type of crime in a pair has little effect on the probability of occurrence of a pair.* One of the conditions of symmetry with respect to pair order thus is satisfied; any two types of crime have essentially the same probability of occurrence regardless of the order of victimization. For example, the probability of a household larceny less than $50 being followed by a personal larceny less than $50 is 310 in 10,000 pairs, while it is 305 for a personal larceny less than $50 being followed by a household larceny less than $50. There is no evidence then that victims or offenders have any effect on the order of victimization by different types of crime for repeat victimization. Where there is repeat victimization by the same type of crime, however, it is quite possible that victims and offenders have an effect, given their significant departure from chance occurrence.

The second condition for symmetry is that the probability of any two pairs that differ only by order of occurrence differ from that of all other pairs. This condition of symmetry is more or less satisfied for major types of crime though it is less apparent for many of the pairs for detailed types of crime. *Generally, however, there is symmetry in the pairs involving major crimes against persons and households.*[2]

There is considerable variation, however, in the probability that any two types of crime incidents will be consecutive victimizations. The probability of consecutive victimization by the same type of crime varies directly with the probability of occurrence of that type of crime among all victims. Moreover, the greater the probability of victimization by any type of crime, the more likely it is to occur in consecutive victimization with any other type of crime.

Switches in type of household crime in repeat victimization of households

We shall next consider briefly the household as the unit at risk in repeat victimization for household crimes only—those of burglary, household larceny, and motor vehicle theft. In general, repeat victims of household crimes are no more or less prone to victimization by certain household crimes than are all victims (Table 6). Motor vehicle theft may occur somewhat less often among repeat than among all victims, but the difference is very small.

Among major crimes against households, the most likely household victimization is that of household larceny (52 percent) with burglary somewhat less common (40 percent) and motor vehicle theft least common (8 percent). Actual household crimes occur with greater frequency than do attempted ones. Attempted household crimes are particularly uncommon for household larceny where only about 1 in 14 household larcenies is reported as an attempt. The comparable odds for burglary are that 1 in 4 is an attempt while almost 1 in 3 motor vehicle thefts is an attempt. The odds are about the same for repeat as for all victims during the three and one-half years for which victimization was reported.

[2]A more succinct test for symmetry is presented in the companion paper by Fienberg.

Table 6. Percent of each type of crime for order of reporting household victimizations: All household victimizations, July 1, 1972, to December 31, 1975

Type of household crime	Reported as first incident		Reported as last incident		Reported as preceding incident of a pair		Reported as following incident of a pair		All reported incidents*	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Burglary, forced entry	4,140	12.7	4,102	12.6	2,257	13.4	2,215	13.1	6,362	12.8
Burglary, no force	5,978	18.3	6,005	18.4	3,062	18.1	3,087	18.3	9,074	18.3
Attempted burglary	2,914	8.9	2,891	8.9	1,565	9.3	1,541	9.1	4,458	9.0
Household larceny, $50 and over	4,359	13.4	4,323	13.2	2,295	13.6	2,262	13.4	6,621	13.4
Household larceny, under $50	11,387	34.9	11,403	34.9	5,881	34.8	5,889	34.9	17,286	34.9
Attempted household larceny	1,107	3.4	1,126	3.4	587	3.5	606	3.6	1,713	3.4
Motor vehicle theft	1,740	5.3	1,816	5.6	740	4.4	816	4.8	2,562	5.2
Attempted motor vehicle theft	999	3.1	970	3.0	497	2.9	468	2.8	1,469	3.0
Totals	32,624	100.0	32,636	100.0	16,884	100.0	16,884	100.0	49,545	100.0

*The entries for "Reported as first incident" and "Reported as following incident of a pair" do not always sum to the total, "All Reported Incidents," for each type of crime because there are 37 incidents for which month of occurrence is unreported.

There is some variation in patterns of repeat victimization for the different detailed and major types of household crime (Table 7).

There is substantial proneness to repeat victimization by the same major type of household crime. While the odds of household larceny are roughly 5 in 10 for all household victims, they rise to between 6 and 7 in 10 for households previously victimized by a household larceny. Similarly, while the burglary odds are 4 in 10 for all household victims, they are almost 6 in 10 for households previously victimized by burglary. The odds for victimization by motor vehicle theft are less than 1 in 10 for all victimized households, but almost 3 in 10 when a motor vehicle theft is previously reported.

Regardless of the type of burglary, the odds are that the next household victimization will be a burglary. There is, moreover, a substantial propensity for victimization by the same type of burglary. The same pattern holds for household larceny. For motor vehicle theft, however, the odds are that the next victimization will be a household larceny or burglary, though there is nonetheless a substantial propensity for repeat victimization by motor vehicle theft. The odds on repeat victimization by motor vehicle theft are four times greater for those previously victimized by motor vehicle theft than they are for all household victims and more than seven times greater for attempted motor vehicle theft.

The chances that a household will be victimized by the same type of household crime, then, are substantially greater than one expects given the chances all households have for victimization by that type of crime. Further evidence for this is found in Table 8. The propensity to repeat victimization by the same type of crime (stayer or diagonal cells in Table 8) is substantial. These are the only cells where the observed frequency is signficantly greater than the expected and they account for 81 percent of the variance in chi-squared for repeat victimization by detailed types of crime and 68 percent of the variance in major types of crime. One other pattern is worth noting. Actual burglaries occur significantly less often with actual household larcenies than one would expect given the chances of victimization for all households. This is surprising given the close relationship between the two types of crime and the possibilities for misclassification.

As one might expect from the substantial differences in the probability of victimization among different types of household crimes, there is considerable variation in the probability that any two types of household crime will be consecutive victimizations. The probability of consecutive victimization by the same type of household crime generally varies directly with the probability of occurrence of that household crime among all victims. Moreover, the greater the probability of victimization by any type of household crime, the more likely it is to occur in repeat victimization with every other type of household crime.

For every major type of household crime also, the probability that it will occur consecutively in repeat victimization with any other type of crimes is a function of their joint probability of victimization. That is, excluding victimization by the same type of crime, for every major type of household crime, the probability that it will occur with every other major type of household crime in repeat victimization varies directly with their expected probability of occurrence based on the experience of all victims.

Conclusion

Within a population of victims, there is considerable multiple or repeat victimization. Evidence on repeat victimization makes it clear that victimization is not a random occurrence but that there is proneness to repeat victimization. Moreover, in repeat victimization, there is a proneness to repeat victimization by the same type of crime. Apart from a marked proneness to victimization by the same type of crime in repeat victimization, most patterns of victimization by type of crime in repeat victimization occur about as often as expected.

The order of occurrence of a type of crime in a pair has little effect on the probability of occurrence of a pair. There is, moreover, a general symmetry in the pairs involving major types of crimes against persons and households. The probability of consecutive victimization by the same type of crime varies directly with the probability of occurrence of that type of crime among all victims. The greater the probability of victimization by any type of crime, the more likely it is to occur in consecutive victimization with any other type of crime.

Reference

Dodge, Richard W. (1975)
National Crime Survey: Comparison of victimizations as reported on screen questions with their final classification, 1974. Washington, D.C.: U.S. Bureau of the Census Memorandum, December 1, 1975, p. 5.

Table 7. Percent distribution by next reported for preceding detailed and major types of household crime for households reporting two or more household victimizations: All household victimizations, July 1, 1972, to December 31, 1975

Percent distribution by following detailed type of household crime

Type of crime reported as preceding incident	Type of crime reported as following incident: Burglary force	Burglary no force	Att. burglary	Household larceny $50 and over	Household larceny under $50	Att. household larceny	Motor vehicle theft	Att. motor vehicle theft	Total percent	Total number preceding incidents
Burglary, forced entry	34	14	11	10	22	3	5	2	101	2,257
Burglary, no force	11	39	6	11	25	2	4	2	100	3,062
Attempted burglary	15	13	26	9	28	3	3	3	100	1,565
Household larceny, $50 and over	10	15	7	26	30	4	5	3	100	2,295
Household larceny, under $50	8	13	6	12	51	4	4	2	100	5,881
Attempted household larceny	7	12	10	15	36	14	3	3	100	587
Motor vehicle theft	13	14	9	14	23	2	20	5	100	740
Attempted motor vehicle theft	9	11	8	10	25	5	10	22	100	497
All next reported incidents	13	18	9	13	35	4	5	3	100	16,884

Percent distribution by following major type of household crime

Type of crime reported as preceding incident	Type of crime reported as following incident: Burglary	Household larceny	Motor vehicle theft	Total: Percent	Number of preceding incidents
Burglary	57	37	6	100	6,884
Household larceny	29	65	6	100	8,763
Motor vehicle theft	32	40	28	100	1,237
All next reported incidents	40	52	8	100	16,884

Table 8. Actual and percent cell contribution to chi square of pairs of household crimes in a crime-switch matrix for repeat household victims of crime: All detailed household victimizations, July 1, 1972, to December 31, 1975

Type of household crime reported as preceding incident	Type of household crime reported as following incident: Burglary, forcible entry	Burglary, no forcible entry	Attempted burglary	Household larceny, $50 and over	Household larceny, under $50	Attempted household larceny	Motor vehicle theft	Attempted motor vehicle theft
Burglary, forcible entry	+749	−25	+7	−15	−107	−5	−*	−11
Burglary, no forcible entry	−11	+753	−31	−16	−95	−9	−4	−14
Attempted burglary	+2	−26	+515	−18	−23	−1	−10	+*
Household larceny, $50 and over	−18	−9	−12	+254	−13	+*	+*	+*
Household larceny, under $50	−115	−81	−47	−6	+444	−*	−22	−17
Attempted household larceny	−15	−12	+*	+1	+*	+176	−4	+*
Motor vehicle theft	−*	−7	−*	+*	−30	−6	+371	+9
Attempted motor vehicle theft	−8	−15	−1	−4	−14	+4	+33	+631
	Percent of total chi square							
Burglary, forcible entry	*15.5*				2.2			
Burglary, no forcible entry		*15.6*			2.0			
Attempted burglary			*10.7*					
Household larceny, $50 and over				*5.3*				
Household larceny, under $50	2.4	1.7			*9.2*			
Attempted household larceny						*3.7*		
Motor vehicle theft							*7.7*	
Attempted motor vehicle theft								*13.1*

χ^2 = 4,826; d.f. = 49; Goodman-Kruskal Index = .22 ± .008.
*Less than .5

Statistical modelling in the analysis of repeat victimization

STEPHEN E. FIENBERG
Department of Statistics
and Department of Social Science
Carnegie-Mellon University

The National Crime Survey's panel structure allows for the longitudinal analysis of individual victimization records. In this paper, further details are provided on a semi-Markov model for such longitudinal analysis proposed in Fienberg (1978). The relationship between this model and some analyses of Reiss (1980) on data for repeat victimization are noted. Finally, Reiss's data are reanalyzed using a variety of log-linear models. The resulting models are interpreted in the context of victim vulnerability or proneness.

1. Introduction

The National Crime Survey (NCS) has a longitudinal structure that potentially allows for the analysis of individual and household victimization records over time. In particular, this longitudinal structure can be used for the examination of the extent of repeat victimization, and for an exploration of the concept of victim proneness. Fienberg (1978) briefly outlined a proposal for the use of a semi-Markov model for multiple victimization over time. In Section 2 of this paper we give more details for this model and point out some of its implications for the analysis of longitudinal victimization records. This paper presumes a knowledge of various concepts and phenomena related to the NCS and victimization surveys, as presented e.g., by Fienberg (1978) or Penick and Owens (1976).

An interesting and well-known consequence of the use of the semi-Markov model of Section 2 is the separation, for estimation purposes, of the transitions from one victimization to subsequent ones, from the timing of the victimizations (i.e., information on the inter-victimization time intervals). If the Markov-chain component of the semi-Markov process is of order one, then we only need to examine a one-step transition matrix for repeat victimization. Thus a version of the model of Section 2 is consistent with the analysis of the crime-switch matrix presented by Reiss (1980) in the preceding paper. The link between Reiss's approach and the use of semi-Markov models is outlined in Section 3. Then in Section 4 we present a reanalysis of Reiss's crime-switch matrices using multiplicative (or loglinear) models of the sort discussed in Bishop, Fienberg, and Holland (1975). This method of analysis leads to a more systematic exploration of patterns of repeat victimization and risk than has been proposed to date. The results of our reanalyses lead to a somewhat different and more parsimonious interpretation of the crime switch data than that presented by Reiss (1980).

2. Semi-Markov models for multiple victimization

The basic idea suggested by Fienberg (1978) was the modelling of victimization for an individual or household as a point process $\{Y(t) > 0\}$, where $Y(t) = j$ at time t if the individual was last a victim of crime type j. A starting point for such modelling is the semi-Markov process with transition probabilities,

$$P_{ij}(t) = \Pr\left\{Y(t) = j \mid Y(0) = i\right\}, \tag{1}$$

where i and j run over the possible crime types under consideration, say $1 \leqslant i, j \leqslant r$. (For details, see Çinlar 1975).

These transition probabilities can be expressed directly in terms of two sets of quantities:

(1) A matrix, $\boldsymbol{M} = \{m_{ij}\}$, of one-step transition probabilities which represent the individual's "victimization propensities" given his current victimization state, i.e., his value of $Y(t)$.

(2) A family of waiting time distributions, $\mathcal{G} = \{G_{11}(t), G_{12}(t), \ldots, G_{1r}(t)$ $G_{21}(t), \ldots, G_{rr}(t)\}$, that characterize the inter-victimization intervals (i.e., the times between victimizations) and which depend on the future victimization state as well as the current state of the Markov chain.

For simplicity, we replace $\mathcal{G}$ by the family of waiting times, $\mathcal{F} = \{F_1(t), \ldots, F_r(t)\}$, which depend only on the *current* state of the Markov chain. We also note that this modelling approach can be extended to handle more complicated processes involving m-step (for $m > 2$) transition probabilities in place of $\boldsymbol{M}$, and corresponding waiting time structures in place of $\mathcal{G}$ or $\mathcal{F}$.

A useful way to think of this semi-Markov process is in terms of a sequence of random variables

$$(X_1, S_1, X_2, S_2, \ldots, X_n, S_n) \tag{2}$$

where n itself may be a random variable. The X_i's take values $1, 2, \ldots, r$, and they form a discrete-time, first-order Markov chain with transition matrix $\boldsymbol{M}$, where

$$Y(t) = X_k \text{ for } \sum_{i=0}^{k-1} S_i \leqslant t < \sum_{i=0}^{k} S_i. \tag{3}$$

The S_k's are positive random variables with

$$\begin{aligned} &\Pr\left\{S_k > t \mid X_1, S_1, X_2, S_2, \ldots, X_{k-1}, S_{k-1}, X_k = i\right\} \\ &= \Pr\left\{S_k > t \mid X(k) = i\right\} \\ &= 1 - F_i(t), \end{aligned} \tag{4}$$

for $1 \leqslant i \leqslant r$.

Corresponding to the sequence of random variables in (2) is a vector of observed values

$$(x_1, s_1, x_2, s_2, \ldots, x_n, s_n), \tag{5}$$

and the likelihood of this vector given the model is

$$\ell = \left(\prod_{j=2}^{n} m_{x_{j-1}, x_j} \right) \left(\prod_{j=1}^{n} f_{x_j}(s_j) \right) \tag{6}$$

where $f_i(t)$ is the probability density function corresponding to the distribution function $F_i(t)$. Here we have ignored x_0, the unobserved victimization state at the time we begin to observe the process, and so, the corresponding inter-victimization interval which is observed only in a censored form. If n, the number of observed victimizations, is substantial, then the loss of information by ignoring the time lapse s_0^* ($< s_0$) from the beginning of the observation to the first recorded victimization is negligible.

The important feature of the likelihood of expression in (6) is that it factors into two parts, the first involving the transition matrix M, and the second involving the parameters underlying the interval distributions $\{F_i(t)\}$. If we estimate the underlying parameters of M and the $\{F_i(t)\}$ by the method of maximum likelihood, we can estimate the parameters of M separately from those underlying the $\{F_i(t)\}$. The first component of the likelihood can be rewritten as

$$\prod_{i=1}^{r} \prod_{j=1}^{r} m_{ij}^{n_{ij}} \tag{7}$$

where n_{ij} is the number of transitions observed from victimization state i to victimization state j. Since there are $n - 1$ such transitions in the observed sequence, (5), we have that

$$\sum_{i=1}^{r} \sum_{j=1}^{r} n_{ij} = n - 1. \tag{8}$$

Since $M = \{m_{ij}\}$ is a matrix of transition probabilities, we have that

$$\sum_{j=1}^{r} m_{ij} = 1 \quad \text{for } i = 1, 2, \ldots, r, \tag{9}$$

so that the likelihood function in (7) consists of a product of r multinomials.

For each individual or household in the NCS we can, in principle, collect full information of the form (6) for a full 3-year period. Unfortunately, the data records for most individuals are fragmentary, with as few as 20 percent of all sample members having complete 3-year records. The observed data are further complicated by several features of the NCS:

(a) Time is discretized, i.e., measured by months.

(b) Since victimizations are recorded only by month of occurrence, multiple victimizations within a given month are not sequenced in order of occurrence.

(c) Series victimizations (see Fienberg 1978 or Penick and Owens 1976 for a detailed discussion of this phenomenon) yield no information about individual inter-victimization times, only about their sum (at least within quarters of the year).

(d) When both series and nonseries victimizations are reported during a given interval, the incidents cannot be ordered by time of occurrence.

In the following section we briefly outline the approach adopted by Reiss (1980) for the handling of these problems to construct an observed crime-switch matrix in order to make inferences about the transition probabilities in the matrix M.

3. Constructing a crime-switch matrix

As Reiss (1980) notes we can adopt any one of four choices regarding the unit at risk for our analysis of repeat victimization:

(1) The household, using both personal victimizations for all household members and household victimizations.

(2) The household, using only household victimizations.

(3) The individual, using only personal victimizations.

(4) The individual, using both personal victimizations and household victimizations.

In Section 4 we analyze a crime-switch matrix constructed by Reiss for category (1), households with all victimizations.

The two tables reanalyzed in Section 4 are taken from Reiss (1980) and are based on household data for all households reporting two or more victimizations between July 1, 1972, and December 31, 1975. Thus the data have been aggregated across households, and no attempt has been made to control for the size of the household or for any other household characteristics that might affect the transition probabilities. While this aggregation implicitly assumes a homogeneity of households so that the transition matrix M of Section 2 does not depend at all on household characteristics, neither Reiss nor we in fact believe in such assumptions. Rather the aggregation of households has been done for the purposes of exploratory analysis and illustration. More refined analyses are not only possible, but also desirable.

In constructing the crime-switch matrices used in Section 4, Reiss (1980) needed to handle many of the problems related to the recording of multiple victimizations discussed at the end of the last section. In particular he chose to record series victimizations as single incidents in the month of first reported occurrence, or at the beginning of a reference period if that month lies outside it. Whenever multiple unordered incidents were recorded, the incidents were ordered using a randomization procedure described by Reiss (1980). The randomization procedure clearly introduces considerably more symmetry into the data than we would expect to be present in the original unobserved victimization sequences. The symmetry in the constructed crime-switch matrices is discussed in the course of our analyses.

Finally, the data used by Reiss are based on both bounded and unbounded interviews. Thus households used in the construction of the sample may be in the sample for from one to seven interviews, each of which yields victimization information for the preceding 6-month period. Given the potential reporting biases associated with time in sample and differential dropout rates of various groups (see the discussion in Fienberg 1978), more refined analyses than those reported here should attempt to control for those factors in some way.

4. Analyses of crime-switch data

Our analyses here are carried out on two related tables of crime-switch or victimization transition data constructed, discussed, and analyzed previously by Reiss (1980). Table 1 gives the data classified by 20 detailed types of crime, and Table 2 gives an aggregated version of Table 1 using 8 major crime types.

We begin with an analysis of the data in Table 1 and treat the counts as if they were the $\{n_{ij}\}$ of expression (7) in Section 2. What we would like to do is model the matrix of transition probabilities, $\boldsymbol{M}$. The first question we ask is: Are the rows of the matrix $\boldsymbol{M}$ in fact the same, i.e., do we have homogeneity of row proportions? The answer to this question is clearly no, as can be seen from the values of the goodness-of-fit statistics reported in Table 3 for this model, $G^2 = 12{,}420$ and $X^2 = 18{,}410$ with 361 d.f.[1]

Next we explore the question of symmetry of transition probabilities, i.e.,

$$m_{ij} = m_{ji} \quad i \neq j, \tag{10}$$

subject, of course, to the constraints, (9), that the m_{ij} sum to 1 over j, i.e.,

$$\sum_j m_{ij} = 1 \quad \text{for all } i.$$

[1]This value of X^2 differs slightly from that reported by Reiss (1980), who combines the categories for rape and attempted rape, and for purse-snatching and attempted purse-snatching, yielding an 18 × 18 table. Reiss reports $X^2 = 17{,}461$ with 289 d.f. for this collapsed table.

This corresponds to the model of quasi-symmetry, discussed in Chapter 8 of Bishop, Fienberg, and Holland (1975), and from Table 3 we see that this model provides a remarkable fit to the data—$G^2 = 174.9$ and $X^2 = 167.1$, with 169 d.f. (We need to fit the model of quasi-symmetry rather than the model of symmetry because of the marginal constraints on the $\{m_{ij}\}$.) This feature of symmetry in the observed data was also noted by Reiss (1980) and partially anticipated earlier in this paper when we discussed the use of randomization to order the unordered incidents. As a consequence of the excellent fit of the symmetry model of expression (10), we treat cells above and below the main diagonal in a symmetric manner in the following analyses.

Since the model of homogeneity of row proportions does not fit the data, it is natural to explore whether the model fits a restricted subset of the data, i.e., whether some form of quasi-homogeneity holds for subsets of the cells in the table. Because of the relatively large number of observed repeat victimizations of the same type relative to the expected values for the homogeneity model, the first quasi-homogeneity model we explore is based on dropping the diagonal cells and corresponds to the well-known mover-stayer model discussed in the context of social mobility studies. While this model gives a substantially improved fit ($G^2 = 2261$ and $X^2 = 2505$ with 341 d.f.), it still does not fit the data well. Next, we used the 8 major groupings of crime from Table 2 and dropped the corre-

Table 1.

Type of crime reported as preceding incident	Rape	Att. rape	Serious assault	Minor assault	Att. assault	Robbery	Att. robbery	Purse snatch, pocket picking
Rape	6	3	8	1	5	2	2	1
Attempted rape	1	16	6	9	21	1	6	2
Serious assault	2	11	119	75	212	37	19	18
Minor assault	4	8	83	152	244	22	20	21
Attempted assault	11	29	199	228	1,685	59	81	37
Robbery	2	4	50	30	76	89	27	23
Attempted Robbery	1	5	30	21	72	27	54	12
Purse snatch/pocket picking	1	2	21	19	47	26	9	46
Attempted purse snatch	-	-	3	2	10	2	3	1
Personal larceny, $50+	7	15	110	99	401	61	44	44
Personal larceny, under $50	16	33	238	292	1,314	126	136	144
Attempted personal larceny	1	3	25	34	115	22	24	15
Burglary, forced entry	7	17	90	59	199	42	25	27
Burglary, no forcible entry	9	11	85	107	319	45	34	41
Attempted burglary	-	8	46	46	166	30	15	23
Household larceny, $50+	4	8	62	61	238	43	24	30
Household larceny, under $50	7	18	89	154	554	75	49	64
Attempted household larceny	1	4	11	8	74	4	11	3
Motor vehicle theft	1	1	29	24	78	20	11	16
Attempted motor vehicle theft	-	1	12	16	62	8	12	5
Total number	81	197	1,316	1,437	5,892	741	606	573

*Correspond to Table 2a in Reiss (1980).

sponding 8 diagonal blocks of cells.[2] The resulting quasi-homegeneity model again is an improvement over the preceding model, but still does not fit the data well ($G^2 = 1140$ and $X^2 = 1238$ with 309 d.f.).

Rather than continue the analysis on Table 1, we choose at this point to switch to Table 2, which uses only the 8 major crime types. Again, for this reduced table the model of quasi-symmetry fits well ($G^2 = 24.4$ and $X^2 = 24.0$ with 21 d.f.), while the model of homogeneity of row proportions fits poorly ($G^2 = 8762$ and $X^2 = 11{,}190$ with 40 d.f.). The quasi-homogeneity model based on the table dropping the diagonals once again gives a remarkable improvement in fit, but still does not fit the data adequately ($G^2 = 383.9$ and $X^2 = 404.3$ with 41 d.f.).

[2]The 8 groupings involved rows and columns as follows: [1,2], [3,4,5], [6,7], [8,9], [10,11,12], [13,14,15], [16,17,18], [19,20]. Since 4 diagonal blocks of 9 cells and 4 diagonal blocks of 4 cells are dropped, there are $361 - 4 \times 9 - 4 \times 9 = 309$ d.f. for this model.

We continue, dropping pairs of cells above and below the diagonal (i.e., (i,j) and (j,i)), and fitting the model of quasi-homogeneity to those that remain. At each stage we drop the pair with the largest standardized residuals and, if there is no clear choice, we try a few likely candidates and take that pair which reduces the value of G^2 the most. The resulting chi-square values and corresponding d.f. are reported in Table 4. Model (g) in Table 4, which is singled out by this procedure, fits the data reasonably well ($G^2 = 62.3$ and $X^2 = 64.7$ with 29 d.f.), especially when we recall that the symmetry is still not being handled directly by the modelling here. A closely related model with the same d.f., which actually fits the data better and is more easily interpretable, is given in the last line of Table 4. What it suggests is elevated occurrences above those expected from the quasi-homogeneity model for pairs of successive crimes involving personal violence (categories 1, 2, and 3) as well as those involving theft without personal contact (categories 5, 6, and 7) plus elevated occurrences of the pairs of successive crimes involving purse-snatching and personal larceny. The reasonable fit of this model suggests that those patterns detectable in repeat victimization data are associated with crimes of similar type. This observation should lead to further investigations regarding the vulnerability or "proneness" of certain groups of households to certain types of crime.

We note in conclusion that those pairs of cells, singled out as departing from the quasi-homogeneity model in the preceding analysis, differ from those singled out by the techniques used by Reiss (1980), and the resulting model seems much easier to interpret.

5. Suggestions for future analyses of longitudinal victimization records

The semi-Markov model proposed in Section 2 leads to the separate analyses of data on victimization switches or transitions and of data on inter-victimization interval distributions. In Sections 3 and 4 we have suggested some analyses that can be usefully performed on the victimization switch data.

Repeat victimization data for 20 crime categories: Reported crimes by households with two or more victimizations while in survey, July 1, 1972, to December 31, 1975

Type of crime reported as following incident

Att. purse snatch	Personal larceny, \$50+	Personal larceny, under \$50	Att. personal larceny	Burglary, force	Burglary, no force	Att. burglary	Household larceny, \$50+	Household larceny, under \$50	Att. household larceny	Motor vehicle theft	Att. motor vehicle theft	Total number
-	7	16	-	5	2	4	2	9	1	3	-	77
3	11	43	5	7	8	13	7	27	2	5	3	196
2	100	233	26	78	92	53	68	145	16	32	8	1,346
4	93	315	30	64	83	32	82	163	16	25	12	1,473
3	371	1,244	141	196	301	184	230	557	72	87	52	5,767
1	76	130	21	41	54	23	41	78	5	15	13	799
-	55	156	21	26	37	16	27	63	7	13	6	649
4	48	141	22	38	42	26	30	48	8	17	15	610
10	9	18	5	4	4	1	3	10	2	6	-	93
9	859	1,385	173	231	337	159	317	544	55	133	71	5,054
13	1,428	7,067	474	436	960	363	637	1,749	158	247	168	15,999
4	156	446	149	53	72	47	68	142	19	35	33	1,463
4	212	450	53	666	237	198	183	394	53	79	31	3,026
7	354	917	74	258	993	134	253	550	58	88	38	4,375
-	151	395	43	192	172	360	116	327	39	33	32	2,194
4	350	636	65	179	261	121	475	552	68	82	48	3,311
14	544	1,779	133	378	597	297	555	2,526	171	149	81	8,234
2	65	161	24	33	49	47	62	164	73	16	15	827
2	134	253	25	74	76	48	74	127	8	122	28	1,151
1	77	156	33	36	40	27	39	96	23	36	83	763
87	5,100	15,941	1,517	2,995	4,417	2,153	3,269	8,271	854	1,223	737	57,407

Table 2. Repeat victimization data for eight major crime categories: Reported crimes by households with two or more victimizations while in survey July 1, 1972, to December 31, 1975 (adapted from Table 1) [Source: Reiss (1980)]

2nd victimization in pair/ 1st victimization in pair	Rape	Assault	Robbery	Purse snatching, pocket picking	Personal larceny	Burglary	Household larceny	Motor vehicle theft	Totals
Rape	26	50	11	6	82	39	48	11	273
Assault	65	2,997	238	85	2,553	1,083	1,349	216	8,586
Robbery	12	279	197	36	459	197	221	47	1,448
Purse snatching, pocket picking	3	102	40	61	243	115	101	38	703
Personal larceny	75	2,628	413	229	12,137	2,658	3,689	687	22,516
Burglary	52	1,117	191	102	2,649	3,210	1,973	301	9,595
Household larceny	42	1,251	206	117	3,757	1,962	4,646	391	12,372
Motor vehicle theft	3	221	51	24	678	301	367	269	1,914
Totals	278	8,645	1,347	660	22,558	9,565	12,394	1,960	57,407

These analyses of the first-order observed transition matrices represent only a first, and very exploratory step in the analysis of longitudinal victimization data from the NCS. In addition to disaggregating the data in the various ways suggested in Sections 3 and 4 and repeating similar analyses, we also need to address more carefully the procedures used to order the unordered victimization incidents and to handle series victimizations and how they related to the occurrences of other victimizations.

We also need to begin analyzing inter-victimization interval data and to concern ourselves with the overall usefulness and goodness-of-fit of the full semi-Markov model.

Because of the peculiarities associated with the data from the NCS (see Fienberg 1978), many of these more elaborate analyses will require new statistical methods or innovative adaptations of existing techniques.

Table 3. The goodness-of-fit for various loglinear models applied to repeat victimization data of Table 1

Model	d.f.	Likelihood ratio chi-square G^2	Pearson chi-square X^2
Homogeneity of row proportions	361	12,420	18,410
Quasi-symmetric	169	174.9	167.1
Quasi-homogeneity			
(a) dropping diagonal cells	341	2,261	2,505
(b) plus blocks of cells for 8 major crime groupings	309	1,140	1,238

Table 4. The goodness-of-fit for various loglinear models applied to repeat victimization Data of Table 2

Model	d.f.	Likelihood ratio chi-square G^2	Pearson chi-square X^2
Homogeneity of row proportions	49	8,762	11,190
Quasi-symmetry	21	24.4	24.0
Quasi-homogeneity			
(a) dropping diagonal cells	41	383.9	404.3
(b) plus (6,7) pair	39	234.9	252.2
(c) plus (2,3) pair	37	172.4	189.7
(d) plus (3,4) pair	35	128.7	131.6
(e) plus (1,2) pair	33	103.7	105.9
(f) plus (2,7) pair	31	83.5	85.4
(g) plus (2,8) pair	29	62.3	64.7
Quasi-homogeneity dropping diagonal cells plus (1,2), (1,3), (2,3), (3,4), (5,6), and (5,7) pairs	29	67.4	72.1
Quasi-homogeneity dropping diagonal cells plus (1,2), (1,3), (2,3), (3,4), (5,6) and (6,7) pairs	29	51.6	53.2

Acknowledgments

The author is indebted to Albert J. Reiss, Jr., who provided the data analyzed in Section 4 and whose original analyses suggested those carried out in this paper. Thanks are also due to Dennis Jennings, who assisted with the computations and analyses.

References

Bishop, Y. M. M., S. E. Fienberg, and P. H. Holland (1975)
Discrete multivariate analysis: Theory and practice. Cambridge, Mass.: M.I.T. Press.

Çinlar, E. (1975)
Introduction to stochastic processes. Englewood Cliffs, N.J.: Prentice-Hall.

Fienberg, S. E. (1978)
"Victimization and the National Crime Survey: Problems of design and analysis." In K. Namboodiri (ed.), *Survey sampling and measurement*, New York: Academic Press, pp. 89-106. (Reprinted in this volume.)

Penick, B. K., and M. E. B. Owens (eds.) (1976)
Surveying crime. Washington, D.C.: National Academy of Sciences.

Reiss, A. J., Jr. (1980)
"Victim proneness in recent victimization by type of crime." (In this volume.)

Sentencing and parole decisionmaking

Quantitative approaches to the study of parole

RICHARD PERLINE
Department of Psychology
University of Chicago

HOWARD WAINER
Bureau of Social Science Research
Washington, D.C.

The first part of this paper uses Federal parole data in an attempt to develop a model for the parole system as it stands and to predict recidivism by treating it as a latent trait. The second part discusses the adequacy of Uniform Parole Report (UPR) data for states.

Omnia quae eventura sunt,
in incerto jacent.[1]

THE AREA OF PAROLE RESEARCH is rife with dissension and disagreement, which appears to us to be entirely justifiable. We would like to add to the disagreements without further obscuring the vast darkness of the topic. This paper will be broken into two parts. The first part deals with some analyses we have done on Federal parole data. In this section we will try to develop a model for the parole system as it stands and to predict recidivism by treating it as a latent trait. For those who are in a hurry, our results are at least as good with this approach as with any other, but still leave a great deal to be desired.

[1] "Everything which is to come lies in uncertainty" (from preface to Raleigh's *History of the World*, 1614).

The second part of this talk deals with parole data and their adequacy. In this section we shall discuss data from individual states gathered under the auspices of the Uniform Parole Reports project (hereafter referred to as "UPR Data"). We have recently had the opportunity to do an evaluation of the UPR data for LEAA, and the appalling results have important implications for those interested in doing parole research. We shall keep federal data separate from state data. This does not mean that federal data are better than state data, but only that we don't know about the reliability of the federal data.

Establishing a quantitative model for de facto incarceration rules

Parole standards are set to indicate to the parole board how long a particular kind of convict should stay in prison. These are in the form of upper and lower bounds for each cell in a two-way categorization. The categorization reflects two schools of thought on incarceration:

(1) *"Just desserts"*—"you did something bad so you stay in—the badder the longer."

(2) *Rehabilitation*—"you are more likely to recidivate than he is, so you stay in longer."

These two schools of thought are concretized in the parole standards. The recommended bounds are a matrix in which the rows reflect the seriousness of the offense—the least serious at the top, the most serious at the bottom. Seriousness was determined through some sort of scaling, although how and by whom, I don't know.

The columns of this matrix are determined by "Salient Factor scores," in which a set of 11 variables are scored zero-one. The sum of the convict's scores on these variables yields his salient factor score, which is supposed to be related to the probability of recidivism (r^2 = 0.07). The federal guidelines collapse these salient factor scores into four categories (0, 1, 2 = I; 3, 4, 5 = II; 6, 7, 8 = III; 9, 10, 11 = IV). The midpoints of the recommended sentence length intervals are shown in Table 1 by offender categories and by five seriousness of offense categories (in months for adults).

Table 1.

		Offender category			
		1	2	3	4
Offense severity category	1	8	10	12	14
	2	10	14	18	22.5
	3	11	15	19	23.5
	4	14	18	22	26
	5	23.5	29.5	34	39

Table 2.

	1	2	3	4	Row medians	Row effects (row median.—gr. md.)
1	.9	1.0	1.08	1.15	1.04	−.19
2	1.0	1.15	.26	1.35	1.20	−.03
3	1.04	1.18	1.28	1.37	1.23	0.00
4	1.15	1.26	1.34	1.41	1.30	.07
5	1.37	1.47	1.53	1.59	1.50	.27
	Grand median				1.23	

Table 3.

	1	2	3	4	
1	−.14	−.04	.04	.11	
2	−.20	−.05	.06	.15	
3	−.09	−.05	.05	.14	
4	−.15	−.04	.04	.11	
5	−.13	−.03	.03	.09	
	−.15	−.04	.04	.11	column medians = column effects

Table 4.

Residual from model (in logs)

	1	2	3	4
1	.01	.00	.00	.00
2	−.05	−.01	.02	.04
3	−.04	−.01	.01	.03
4	0.00	0.00	0.00	0.00
5	.02	.1	−.01	−.02

Our first interest was to identify the model underlying this structure. The sentence lengths were originally determined empirically as a measure of the de facto sentence structure for the majority of offenders. We analyzed these data using traditional two-way table decomposition methods. That is, inspection showed that an additive model was not appropriate (this is easily seen by fitting an additive model and seeing the substantial trends in the the residuals). We thus tried a multiplicative model by first calculating logs and taking out row medians; these correspond to seriousness effects (Bishop, Fienberg, Holland 1975; Tukey 1977; Neithercutt, Moseley, Wenk 1975). We then calculated column medians of the residuals and subtracted them; these correspond to offender effects. The resulting matrix of residuals corresponds to the residuals from the model. As can be easily seen, these are all small. The effects were then transferred back to months and the total model specified.

The same multiplicative model (with different values for the parameters) holds for juveniles as well. This model and its fit are shown in Table 6.

We are then reasonably well assured that the existing parole policy corresponds closely to a multiplicative model in which seriousness of offense seems to count almost three times as much as offender characteristics among adults, and about twice as much among juveniles.

Improvement?

Having identified the de facto structure of parole policy, the road toward possible improvement is clear, and involves two steps:

(1) Improve the accuracy of the estimates of seriousness of crimes, so that a more accurate classification is possible.

(2) Improve the scaling of prisoner characteristics so that the prisoner effects more closely correspond to the likelihood of recidivism.

Let us briefly look at step (1). Shown in Figure 1 and Figure 2 are plots of seriousness effects plotted against category. It is clear that categories 1–4 for both juveniles and adults are linearly related to incarceration length, but when crossing into category 5 there is a major jump. It would seem that there is room for about 4 intermediate categories between the crimes which correspond to category 4 and those in category 5. It would seem that finer discriminations are possible. Of course such fine discriminations are made as viewed by the increased width of allowable sentences for the more serious crimes. Why not make this explicit? We did not do the obvious scaling experiment, but the details of how to do such a study are clear (Torgerson 1958). Use classical scaling methodology with some Thurstonian model and have expert judges decide which of two paired crimes is the more serious.

Step (2) was one which we found interesting. Could we improve upon the results that Hoffman and his colleagues (Hoffman and Beck 1974) had obtained through their use of multiple regression? They regressed a variety of variables against the binary variable (Recidivate—yes or no). We know very well that alternative models appear, at least on the surface, to be more suitable for

Table 5.

Adults

Offender effects					
.71	.89	1.10	1.35	19.95	
.2	.2	−.1	−.8	.55	
−1.2	0.0	.7	1.2	.79	Offense seriousness effects
−.2	.2	.1	.1	1.00	
−.3	.1	.7	.3	1.29	
−1.4	−.2	−.3	−1.7	2.29	

Residuals (months)
Residuals = Data Fit
Fit = (19.95) x Offender eff. x Offense eff.

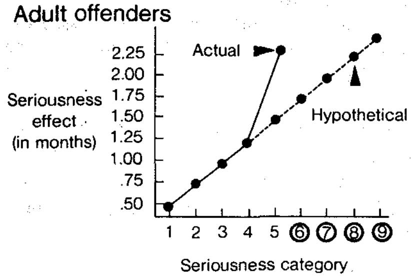
Table 6.

Youth

	Offender effects				
.7	.9	1.1	1.3	17.0	
.9	.8	.8	.7	.6	Offense seriousness effects
−.7	.2	1.2	2.6	.9	
−.9	−.3	.3	1.4	1.2	
−.3	−.4	−.4	−.5	1.2	
.9	.4	−1.5	−.3	1.9	

Residuals (months)
Fit = (17.0) x offender eff. x offense eff.

Figure 1.
Adult offenders



a binary criterion than linear regression. Also, it didn't seem that it would be that hard to improve on a method whose cross-validated accuracy only accounted for 7 percent of the variance. It should be noted that there is a severe restriction of range problem, in that we should give proper credit to parole boards for knowing something. That is that they tend to keep in the really bad characters. The ones released are at the upper end of the distribution and it is therefore harder to discriminate among them. At a guess, I would think that we psychometricians don't do all that much better; I don't know how much of the variance among Harvard students is accounted for by SAT scores, but I imagine it isn't overwhelming. This same restriction of range presumably exists in parole decision. Thus we must not be over-hasty in denigrating the 7 percent figure.

Our approach was to think of the individual's background data as a test that he takes. There were eight variables that entered into this test which can be considered "predictor variables" and a single criterion variable "parole revoked." The eight variables are shown in Table 7. There is nothing magical about the choice of the criterion variable; except that we chose it to be different from the other variables. We might just as well wish to predict one of the other variables, so we included "parole revoked" as one of the "items" and judged how well it is described by the rest of the "test" by looking at its $r_{.bis}$ with the total test. We used a Rasch logistic model (Rasch 1960) to fit these data (n = about 500) and looked at goodness-of-fit of the model. (We used differential slope models as well, but the fit wasn't increased and so we stuck with the equal slope model.)

Associated with each variable is its "difficulty," here interpreted to mean how easy it is to have that variable scored 1. Thus "planned living arrangement" is the "easiest" (86 percent had them) whereas "drug history" is the "hardest" (only about 22 percent had one). Each person also had an "ability" score (perhaps "returnability" is a better term). The difference between a person's returnability and the difficulty of the item describes the likelihood of that person having that particular characteristic. To predict recidivism we calculated the probability of a person having the characteristic "recidivate." The $r_{.bis}$ reflects the relationship between each "item" and the total test. Note that the $r_{.bis}$ associated with recidivism = .57 (squared = .32).

We can see from the results in Table 7 the general extent of fit, and those variables which seem to fit. Note that such variables as "planned living arrangement" don't seem to load too heavily on this trait. One could use such schemes to make decisions as to the importance of various kinds of information. The variables chosen and their scoring were taken precisely from the methodology Hoffman et al. developed. We were not given free access to all data, and so had to content ourselves to these.

Thus these preliminary results look promising. Upon cross validation on a neutral sample there was some shrinkage. Our best guess as to the amount of variance accounted for on a neutral sample is between

Figure 2.
Juvenile Offenders

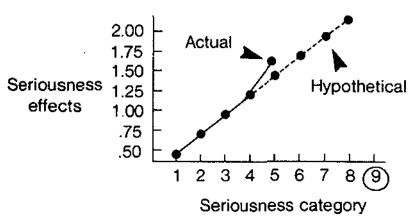


Table 7.

Variable	Difficulty	fit Mn sq.	$r_{.bis}$
Prior conviction	1.12	.61	.76
Age at first commitment	−1.36	.63	.57
Prior incarcerations	−0.07	.65	.82
Parole revoked	−0.83	.74	.57
Verified employment	−0.23	1.05	.41
Grade claimed	0.88	1.06	.38
Auto theft	−1.18	1.07	.33
Drug history	−1.57	1.36	.20
Planned living arrangement	2.20	1.73	.10

15 and 20 percent. A bit of an improvement, but nothing to write home about.

A careful analysis of fit indicated that the model could be rejected, but what does it then mean to say that there is a single underlying dimension of "likelihood of recidivism?" We tried a variety of other schemes: the most interesting one was using Kruskal's monotone analysis, which showed that the best monotonic transformation to additivity did not fit much better than the Rasch model. We conclude from this that although the Rasch model doesn't fit, it does about as well as any single factor model.

Parole data—UPR

The preceding analyses were done on Federal parole data. There exists another data set, with longitudinal records running 3 years on very large samples. These are the Uniform Parole Reports data. These data have been gathered by the National Council on Crime and Delinquency.[2] These data have been gathered for almost a decade and contain a variety of useful information. After a careful examination of these data, I decided that the most valuable thing I could do with this forum would be to issue a short warning on the use of these data. Although these data have been declared "reliable enough" by their original gatherers (Neithercutt et al. 1975), our best guess is that they are not. A number of "shoe-box" reliability studies were carried out in which a very small number of cases were recoded, and the differences between their original coding and second coding were examined (n's in these were about 160). The following results obtained: the percent agreement between the two codings for a seemingly foolproof variable like "ID#" was as low as 93 percent. This was called "high"—I would call it embarrassing. The worst variable (in terms of reliability of coding) was "other prior sentences" which averaged about 53 percent but for some agencies was as low as 27 percent. Thus if one were using such a variable in any prediction scheme accuracy of prediction would certainly suffer. Overall accuracy was 80 percent which included such (relatively) failsafe variables like "is he dead?" One agency got "sex" wrong 4 percent of the time.

There is great variability of accuracy across reporting agencies. Sadly, the reporting agencies are anonymous so you can't be sure how good your data are. These problems are being corrected, and we have reason to hope that in the future UPR data will be useful and reliable. We know for sure that the reporting agencies will no longer be anonymous.

[2]Substantial improvements have been made in this program since this paper was prepared. For current information, write to Dr. James L. Galvin, NCCD, 760 Market Street, Suite 433, San Francisco, California 94102; telephone (415)956-5651.

References

Bishop, Y., P. Holland, and S. Fienberg (1975) *Discrete multivariate analysis.* Cambridge, Mass.: M.I.T. Press.

Hoffman, P. B., and J. L. Beck (1974) "Parole decision-making: A salient factor score." *Journal of Criminal Justice* 2:195-206.

Neithercutt, M. G., W. H. Moseley and E. A. Wenk (1975) *Uniform Parole Reports: A national correctional data system.* Davis, Calif.: NCCD.

Rasch, G. (1960) *Probabilistic models for some intelligence and attainment tests.* Copenhagen: Danmarks Paedogodiske Institut.

Torgerson, W. (1958) *Theory and methods of scaling.* New York: Wiley.

Tukey, J. W. (1977) *Exploratory data analysis.* Reading, Mass.: Addison-Wesley.

Linear logistic models for the parole decisionmaking problem

KINLEY LARNTZ
School of Statistics
University of Minnesota

This paper considers the use of quantitative aids in parole decisionmaking. Current practice employs a Burgess scale which weights items equally and ignores the interval (or ordinal) quality of the data. We propose logistic regression as an alternative statistical decisionmaking guide, and illustrate use of these techniques on data from the State of Minnesota. Also discussed are the problems of selecting a few items from a pool of many items for construction of the scale.

1. Introduction

A recent trend in parole decisionmaking has been the use of objective scores, to be calculated for each potential parolee, which attempt to estimate that individual's chance of parole success. Specifically, a Federal Hearing Examiner is given a potential parolee's Salient Factor Score which, along with the type of crime, provides guidelines for sentence length. The Salient Factor Score is the sum of responses to nine items, seven of which are scored 0 or 1, and the other two are scored 0, 1, or 2. The items for the Salient Factor Score, as reported by Hoffman and DeGostin (1974), are given in Table 1. The guidelines for sentence length, also from Hoffman and DeGostin (1974), are given in Table 2. For information on the development and recommended use of these guidelines, see Gottfredson, et al. (1974).

Various states have also developed Salient Factor Scores that apply to their respective prison populations. The Federal set of nine items is not directly applicable because of the substantial differences in state and Federal crimes. In Minnesota the Parole Decision-Making Project, headed by Dale Parent, developed a Salient Factor Score for aiding in state parole decisions; see Parent and Mulcrone (1978). Table 3 gives the nine items found to be important in Minnesota. (In later implementation of the scale, three items were deleted: "juvenile commitment," "had a sustained juvenile petition," and "completed at least 10th grade.") The scale construction was based on the analysis of data collected on 931 out of about 1500 individuals released from Minnesota prisons in 1970 and 1971. The data set was divided into a construction group of 485 cases and a validation group of 446 cases. A number of regressions as well as Burgess scales were tried out before the final scale was selected. Note that the final Salient Factor Score was one of the best scales on *both* the construction and validation samples. Thus, in fact, all 931 cases played a role in selection of the final nine items.

In scale construction, each individual was classified as a "success" if, in the 2-year period following release, the individual was not convicted of a new felony. Low scores on the scale correspond to good prognosis; high scores to poor prognosis. Table 4 gives the group failure rates for the pooled construction and validation samples. Failure rates range from a low of 13 percent to a high of 54 percent. In Minnesota the Salient Factor Score is used, as in the Federal system, in conjunction with a measure of seriousness to aid in determining a release date.

Table 1. Items in salient factor score

Item A	
	No prior convictions (adult or juvenile) = 2
	One or two prior convictions = 1
	Three or more prior convictions = 0
Item B	
	No prior incarcerations (adult or juvenile) = 2
	One or two prior incarcerations = 1
	Three or more prior incarcerations = 0
Item C	
	Age at first commitment (adult or juvenile) 18 years or older = 1
	Otherwise = 0
Item D	
	Commitment offense did not involve auto theft = 1
	Otherwise = 0
Item E	
	Never had parole revoked or been committed for a new offense while on parole = 1
	Otherwise = 0
Item F	
	No history of heroin, cocaine, or barbiturate dependence = 1
	Otherwise = 0
Item G	
	Has completed 12th grade or received GED = 1
	Otherwise = 0
Item H	
	Verified employment (or full-time school attendance) for a total of at least 6 months during the last 2 years in the community = 1
	Otherwise = 0
Item I	
	Release plan to live with spouse and/or children = 1
	Otherwise = 0

Table 2. Guidelines for decision-making: average total time served before release (including jail time) — Adults

Severity of offense behavior	Offender characteristics parole prognosis (Salient Factor Score)			
	Very good (11-9)	Good (8-6)	Fair (5-4)	Poor (3-0)
Low	6-10 months	8-12 months	10-14 months	12-16 months
Low moderate	8-12 months	12-16 months	16-20 months	20-25 months
Moderate	12-16 months	16-20 months	20-24 months	24-30 months
High	16-20 months	20-26 months	26-32 months	32-38 months
Very high	26-36 months	36-45 months	45-55 months	55-65 months
Greatest	(Greater than above—however, specific ranges are not given due to the limited number of cases and the extreme variations in severity possible within the category.)			

Table 3. Final items for Minnesota Salient Factor Score

Label	Name	Category	Score
X_1	Juvenile commitment	YES NO	1 0
X_2	Number of prior parole/probation failures	2 or more 0,1	1 0
	Number of prior incarcerations	1 or more 0	1 0
X_3	Had a sustained juvenile petition	YES NO	1 0
X_4	Age at first adult conviction	19 or under 20 up	1 0
X_5	Conviction previously—this offense	YES NO	1 0
X_6	Completed at least 10th grade	NO YES	1 0
X_7	This offense was burglary	YES NO	1 0
X_8	3 or more felony convictions	YES NO	1 0

Table 4. Group failure rates for Minnesota data base

Salient Factor Score	Failure rate
0 - 1	13%
2 - 3	17%
4	29%
5 - 6	42%
7 - 9	54 %

This paper presents a reanalysis of the Minnesota data base. The objective of the reanalysis is to explore logistic regression as a tool for predicting parole success. The data base split employed by the Minnesota Department of Corrections was also used here, i.e., the same 485 construction cases and 446 validation cases. No new random splits were tried.

Section 2 of this paper discusses a published reanalysis of the Federal data. Section 3 applies the same reanalysis technique to the Minnesota data base. In Section 4 an alternative technique, logistic regression, is presented and applied to the Minnesota data. Section 5 gives a simple validation of the various procedures, and the concluding section gives the implications of this research for future studies.

2. Solomon's reanalysis of the Federal data

Solomon (1976) reanalyzed the basic data set for 2497 prisoners used in construction of the Federal Salient Factor Score. He showed that the nine items, used with equal weights, could be reduced to four items, provided unequal weighting was allowed. His final model shows decisively that the Salient Factor Score collapses groups that have substantially different rates of risk.

Solomon's final model was a logit model with four variables:

$$\log \frac{p_{ijkl}}{1 - p_{ijkl}} = w + w_{1(i)} + w_{2(j)} + w_{3(k)} + w_{4(l)}, \tag{2.1}$$

where p_{ijkl} represents the true rate of failure for individuals in the cross-classification cell (i,j,k,l), and

$$\sum_i w_{1(i)} = 0, \quad \sum_j w_{2(j)} = 0,$$

$$\sum_k w_{3(k)} = 0, \text{ and } \sum_l w_{2(l)} = 0.$$

Table 5 gives the variables, categories, and weights for his final model. Using these weights we can calculate the estimated failure rate for an individual with no prior convictions, no parole revocations, no auto theft, and planning to live with spouse as

No prior convictions	−0.6450
No parole revocations	−0.2344
No auto theft	−0.2210
Plan to live with spouse	−0.3852
Constant	−1.2675
	−2.7531

$$\text{Estimated failure rate} = \frac{e^{-2.7531}}{1 + e^{-2.7531}} = 0.0599.$$

For additional reading on logit models, see Cox (1970) or Fienberg (1977).

The important difference between Solomon's final logit scale and the nine-item Federal Salient Score is the weight assigned to each item. First, the logit model contains only four items. This has the effect of giving zero weight to the other five items. In addition, for the included variables, different weights were attached to each. In contrast, the Salient Factor Score assigns equal weights to each included item. Essentially, only the sign (positive or negative) of an item's estimated effect on failure rate is considered important. This method of forming a scale, based on equal weighting of items, is often referred to as the Burgess method.

The main contribution of Solomon's study was to offer an alternative to the Burgess method for this decisionmaking problem. Previous studies had compared the Burgess method to linear regression analysis, and found the Burgess technique superior—mainly on the grounds of robustness. However, the assumptions for linear regression are not usually met for this type of data. Solomon's proposal offers a statistical technique specifically tailored to the categorical nature of the data.

Table 5. Weights for Solomon's final logit model

	Independent variables	Categories	w
D_1:	Number of prior convictions	0	−0.6450
		1-2	+0.1054
		3 or more	+0.5396
D_2:	Prior parole revocation	No	−0.2344
		Yes	+0.2344
D_3:	Auto theft	No	−0.2210
		Yes	+0.2210
D_4:	Plan to live with spouse and/or children	Yes	−0.3852
		No	+0.3852
	Constant	—	−1.2675

Table 6. Weights and goodness-of-fit statistics for Minnesota logit models

Variable	Model I	Model II	Model III	Model IV	Model V
X_1	.68385	.66484	.78213	.82426	—
X_2	.89731	.97447	1.0480	1.0029	—
X_3	.23198	.26712	—	—	—
X_4	.13693	.71692	.74209	—	—
X_5	.56074	.54484	—	—	—
X_6	.65332	.66840	.67259	.65932	—
X_7	.45491	1.6642	1.7314	.54767	—
X_8	.07271	.01055	—	—	—
X_4X_7	—	−1.8050	−1.8080	—	—
CONST.	−2.2979	−2.7047	−2.5823	−2.1067	−.91196
X^2	132.31	120.09	126.57	141.27	196.31
G^2	151.14	137.73	141.21	155.38	225.50
df	107	106	109	111	115

3. Logit models for Minnesota data set

Following Solomon's example, logit models were fitted to the 485 cases of the Minnesota construction group. The variables used were those reported in Table 3 with the exception of "Number of Prior Incarcerations"—it was omitted because of a coding problem in our data base. Because of the zero-one nature of each of the X_i's, the logit model may alternatively be written as:

$$\log \frac{p(x_1, x_2, \ldots, x_8)}{1 - p(x_1, x_2, \ldots, x_8)} = w + w_1x_1 + w_2x_2 + \ldots, + w_8x_8, \quad (3.1)$$

where $p(x_1, x_2, \ldots, x_8)$ represents the true rate of failure (felony conviction within 2 years of release) for individuals with scores $(x_1, \ldots, x_8)$ on variables $(X_1, \ldots, X_8)$. There is a one-to-one correspondence between logit models written as in (3.1) and those written as (2.1). The estimated logits will be exactly the same irrespective of which form of the model is selected. Essentially, (3.1) is a regression analogue of the analysis of variance model (2.1). It is interesting to note that the Burgess scale (i.e., Salient Factor Score) says that $p(x_1, x_2, \ldots, x_8)$ is merely a function of $x_1 + x_2 + \ldots x_8$.

Table 6 gives the estimated weights for several logit models. Model I is the logit model with all eight variables included, i.e., $p(x_1, x_2, \ldots, x_8)$ does not depend upon $(x_1, \ldots, x_8)$. Model IV excludes four of the variables which seem to have little importance in the construction sample. Models II and III are different in that an interaction term was added to the basic Models I and IV, respectively. This is accomplished by merely adding a cross-product term to the model, e.g., for Model III,

$$\log \frac{p(x_1, \ldots, x_8)}{1 - p(x_1, \ldots, x_8)} = w + w_1x_1 + w_2x_2 + w_4x_4 + w_6x_6 + w_7x_7 + w_{47}x_4x_7. \quad (3.2)$$

To assess how well these models describe the 485 cases, goodness-of-fit statistics were calculated. For the eight independent variables there are $2^8 = 256$ possible combinations; however, only 116 patterns actually appeared in the data. Adding the dependent variable, also measured as zero-one, we can view the data in the form of a $2 \times 116 = 232$ cell contingency table. For each cell, expected frequencies can be calculated based upon a particular logit model. The following two goodness-of-fit statistics compare the actual cell counts with the expected frequencies.

$$X^2 = \sum_{\substack{\text{all}\\ \text{cells}}} \frac{(\text{Observed} - \text{Expected})^2}{\text{Expected}}, \quad (3.3)$$

$$G^2 = 2\sum_{\substack{\text{all}\\ \text{cells}}} (\text{Observed})\log\left(\frac{\text{Observed}}{\text{Expected}}\right). \quad (3.4)$$

If the expected frequencies are large, it is reasonable to assess goodness-of-fit by comparing the calculated statistic with the percentage points of a chi-squared distribution with the appropriate degrees of freedom. However, here the expected cell frequencies are small. In such a case, the Pearson statistic X^2 has a level near the nominal rate, as long as the expected cell frequencies are not too small (Larntz 1978, Koehler and Larntz 1980). The likelihood ratio statistic G^2 tends to be inflated if there are a large number of moderate (1 to 4) expected frequencies.

Our situation involves large numbers of moderate and small expected frequencies. Thus by looking at the goodness-of-fit statistics listed in Table 6, we can see that Model V certainly does not provide an adequate fit, Models II and III are clearly describing the data well, and Models I and IV are borderline cases. It is interesting to note that for the Burgess scale $G^2 = 157.12$ (114 df). Degrees of freedom for the various models were calculated by considering the data as a $2 \times 116 = 232$ cell contingency table. Thus, Model V, which tests independence in this table, has $(2-1) \times (116-1) = 115$ degrees of freedom. Other degrees of freedom were calculated by subtracting one degree of freedom for each free parameter added to the corresponding model.

The term x_4x_7 was the only one of the 28 possible interaction terms close to statistical significance at a standard level. To judge the statistical significance of the interaction, we look at differences in the G^2-statistics for models fit with and without the interaction. In our case,

$$G^2_{\text{interaction}} = G^2_{\text{I}} - G^2_{\text{II}} = 13.41 \quad (1 \text{ df})$$

or alternatively

$$G^2_{\text{interaction}} = G^2_{\text{IV}} - G^2_{\text{III}} = 14.17. \quad (2 \text{ df})$$

Table 7. Weights for the burglary X age at first conviction interaction

Age at 1st adult conviction	Burglary	
	No	Yes
20 up	0.00	1.73
19 down	0.74	0.67

In both cases, the interaction is significant at 0.05 level, even after adjusting for the fact that all 28 possible two-factor interactions were examined. Table 7 gives the weights for the interaction cells. Note that "burglary" is not an important factor for "age at first adult conviction less than 20," but has a large effect for those "20 or over at first conviction." The risk of parole is greater for burglars first convicted at an older age.

4. Logistic regression models

The logit model with X's taking on 0 – 1 values can be extended to permit continuous X variables or even a mixture of continuous and discrete variables (Nerlove and Press 1973, Fienberg 1977, Cox 1970). Table 8 lists the variables considered in this analysis. Note that a number of variables listed in Table 8 overlap with those in Table 3. However, where possible, the actual values of previously dichotomized variables are now used. Also, sex has been added as a variable and juvenile record and education have been deleted. Many more items were collected by the Parole Decision-Making Project, but data were missing on many variables for a number of cases. (Three individuals were excluded altogether because of missing data. Although it was not done in this analysis, it should be noted that, assuming the omitted data were "missing at random," the missing values can be estimated via the EM algorithm (Dempster et al. 1977). Thus, all information in the full data set could be utilized in a complete analysis.)

Table 8. Variables for logistic regressions

Label	Description
Z_1	Number of parole/probation failures
Z_2	Number of incarcerations
Z_3	Age at first adult conviction
Z_4	Number of felony convictions
Z_5	Sex 1 = Male, 2 = Female
Z_6	Age at release
Z_7	1 if offense was burglary, 0 otherwise

The logistic regression model describes the risk for individual i as

$$\log \frac{p_i}{1-p_i} = \beta_0 + \beta_1 Z_{i1} + \beta_2 Z_{i2} + \ldots + \beta_7 Z_{i7}. \tag{4.1}$$

This model has exactly the same form as the logit model (3.1), but here we are allowing the independent variables (Z's) to take on any value, while in (3.1) the independent variables (X's) were constrained to be zero or one. Recent comparisons of logistic regression with linear regression have indicated that for certain uses, such as classification, logistic regression may possess certain advantages (Press and Wilson 1978).

Table 9 gives the estimated regression coefficients based on the construction sample for five logistic regression models. Attempts at adding interactions did not prove fruitful in this case. The $-2\log f$ value gives twice the log of the likelihood ratio assuming each of the 482 observations is Bernoulli with failure rate p_i as in (4.1) (see Efron 1978). Again differences between likelihood ratio statistics can be used in comparing models; no overall goodness-of-fit test is possible here, however. From difference comparisons, Model III and Model IV emerge as candidates for the best model. Model III needs only two variables for an assessment of an individual's risk, i.e., number of parole/probation failures and age at first conviction. Thus, an individual with one failure, first convicted at age 20, would have estimated logit

$$\log \frac{\hat{p}}{1-\hat{p}} = 0.28031 + 0.22982(1) - 0.07708(20) = -1.03147$$

and estimated risk of a new felony conviction of $\hat{p} = 0.263$. Someone first convicted at age 18 with 7 prior failures would have estimated risk of 0.623.

5. A simple validation of procedures

A criticism that is often made of statistical model fitting is that although the final model fits the construction sample quite well, it does not do nearly so well when applied to new data. To avoid such criticism of their work, the Parole Decision-Making Project divided the data into construction and validation groups. The final Salient Factor Score (see below, Figure 8 and Table 11) did well on both groups. We now examine how well the logit and logistic regression models did when applied to the validation sample.

Table 9. Regression coefficients for logistic regressions

Variable	Model I	Model II	Model III	Model IV	Model V
Z_1	.21955 (2.81)	.21645 (2.76)	.22982 (3.17)	.22061 (3.03)	–
Z_2	−.11502 (−.90)	−.16966 (−1.30)	–	–	–
Z_3	−.09425 (−2.97)	−.10352 (−3.20)	−.00708 (−3.49)	−.07426 (−3.36)	–
Z_4	.09990 (.96)	.09991 (.95)	–	–	–
Z_5	−1.3631 (−1.79)	−1.1829 (−1.55)	–	–	–
Z_6	.00810 (.41)	.01809 (.88)	–	–	–
Z_7	–	.61760 (2.53)	–	.6122 (2.59)	–
CONST.	0.4901	0.3586	0.28031	0.07633	−.91339
−2 log f	525.7	519.4	531.6	525.1	577.3
df	475	474	479	478	481

Note: Approximate t-statistics are in parentheses.

Table 10. Results from applying Logit Model I in selecting 230 parolees

	Non-felon	Felon	Total
Low risk	187	43	230
High risk	134	82	216
Total	321	125	446

NOTE: $P_{NF} = \frac{187}{321} = .583$ $P_F = \frac{43}{125} = .344$

$D = P_{NF} - P_F = .583 - .344 = .239$

$P_{TOT} = \frac{230}{446} = .516.$

Table 11. Maximum D and average D for nine predictive scales

Model	D_{max}	D_{ave}
Logit I	.2480	.1399
Logit II	.2562	.1319
Logit III	.2678	.1353
Logit IV	.2729	.1349
Burgess	.2499	.1570
Logistic I	.2341	.1202
Logistic II	.2340	.1243
Logistic III	.2844	.1324
Logistic IV	.2626	.1405

There are several possible techniques for checking the degree of validation for a predictive scale. One standard method is to separate the validation sample into several groups according to *predicted risk* and see whether the proportions of felons and nonfelons differ in the groups. Ideally, the low risk group would contain predominantly nonfelons, while the high risk group would contain predominantly felons. The statistical signifancance of the degree of separation of felons and nonfelons could be evaluated using a chi-squared test. It should be noted that this technique depends upon the number of groups formed as well as the actual groupings. Different group formation could yield different degrees of validation when applied to the same data.

For comparisons in this paper we have used the following method which has the advantages of (1) not depending upon any arbitrary grouping of the validation sample, (2) allowing direct comparisons between different predictive scales, whether or not the scales actually calculate an estimated risk (the method requires only an ordering of the risk), and (3) providing a natural graphical output illustrating the degree of validation.

To illustrate the technique let us consider logit Model I. For each individual in the validation sample, calculate the estimated risk using the logit model weights given in Table 6. Order the individuals from low to high by this calculated risk. A good predictive scale would have a large proportion of the nonfelons in the "parole" group, while having a large proportion of the felons in the "nonparole" group. Now suppose, for example, that we were to "parole" 230 individuals. Table 10 summarizes logit Model I's performance in "paroling" 230 individuals. Of the 230 lowest risk scores 187 belonged to nonfelons. To measure how well the predictive scale separates the felon and nonfelon groups calculate the difference of the proportion of nonfelons assigned to the low risk group minus the proportion of felons assigned to the low risk group. In the example,

$$p_{NF} = \frac{187}{321} = 0.583, \quad p_F = \frac{43}{125} = 0.344$$

and

$$D = 0.583 - 0.344 = 0.239.$$

Note the proportion "paroled" in this sample was fixed at

$$p_{TOT} = \frac{230}{446} = 0.516.$$

Now consider calculating D for each possible "paroling strategy," i.e., for paroling $1, 2, 3, \ldots, 445$ individuals, or equivalently for $p_{TOT} = 1/446, 2/446, 3/446, \ldots, 445/446$. Figure 1 gives a graph of D vs. p_{TOT} for logit Model I. The higher D is for a given value of p_{TOT}, the better the predictive scale in separating nonfelons from felons. In Figure 1, the peak value of D is 0.2480. An "ideal scale" would have a peak value of D of 1.000.

Figures 1–9 present graphs of D vs. p_{TOT} for 4 logit models, the patterns for all figures are the same, although the peak values of D occur at different values of p_{TOT}. Table 11 gives the peak values of D and also the average values of D (this corresponds to the area under the curve) for the nine scales. Logistic Model III has the largest D value, while the Corrections Department scale has the largest mean.

We now examine the individual figures more closely. First, logit Model I (Figure 1) and logit Model IV (Figure 4) are virtually identical, except for the noise caused by the four extra variables included in logit Model I. The same is true for logit Models II (Figure 2) and III (Figure 3). Thus, as expected, there is no gain in adding nonsignificant variables to the logit model. Similarly, although the interaction term is statistically significant, when we compare Figure 1 with Figure 2 and Figure 3 with Figure 4, the validation curves indicate that inclusion of the interaction term is helpful for some values of p_{TOT} (0.2, 0.4) while not helpful for others ($p_{TOT} = 0.4, 0.6$). The Burgess scale curve (Figure 5) is comparable to the logit models and perhaps even a little higher, but it must be kept in mind that its nine items were selected on the basis of the *same* construction and validation samples being used here, i.e., the full data.

The logistic models (Figures 6–9) are clearly better when fewer variables are included, i.e., logistic Models III and IV. Logistic Model III (the two-variable equation) is particularly strong for p_{TOT} = 0.6, yielding the highest D value for any validation. However, it does not do as well as other models for p_{TOT} near 0 or 1. This may be unimportant since p_{TOT} near 0 or 1 imply strategies of paroling few or nearly all eligible prisoners.

Although not improving greatly over the Burgess scale, the better logit and logistic models do perform well on the validation sample. The two-variable logistic model did particularly well, especially considering that these two variables were selected from a pool of only seven items. The nine items in the Burgess scale were selected from a much larger pool, and since the same items were used in the logit models, the logit models' performances on the validation sample should not be surprising. In both logit and logistic cases, restricting the model to include only "significant" variables improved the validation performance.

Figure 1.
Validation — Logit I

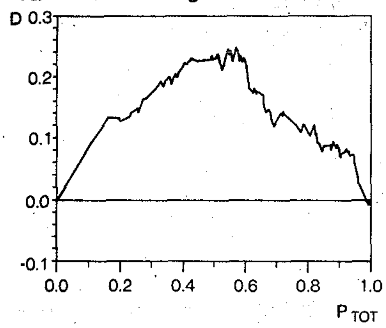


Figure 2.
Validation — Logit II

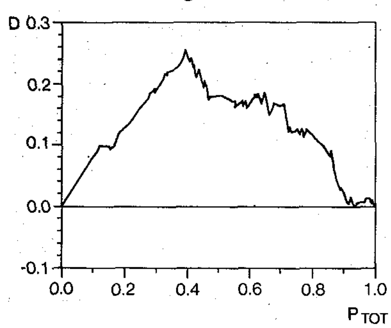


Figure 3.
Validation — Logit III

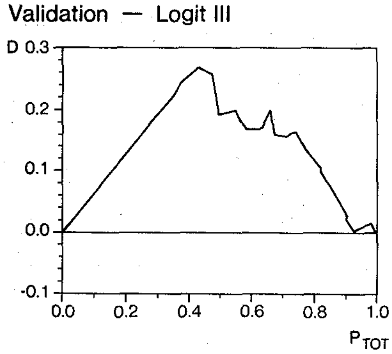
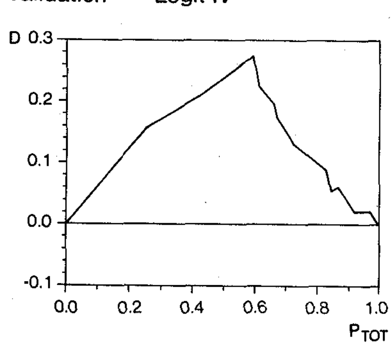


Figure 4.
Validation — Logit IV



6. Discussion

One aspect of the logit and logistic models that has been ignored in the comparisons is the *actual* estimated risk value. Does this estimated risk have a frequency interpretation? To answer this question, we conducted a sampling experiment. Using the $\hat{p}_i$'s of the validation sample, Bernoulli random variables were generated to randomly assign an individual as a felon or nonfelon with probability $\hat{p}_i$ or $1 - \hat{p}_i$, respectively.[1] Figures 10, 11, and 12 give plots of D vs. p_{TOT} for the Monte Carlo runs on logit Model I, logit Model IV, and logistic Model III. If the $\hat{p}_i$'s are good quantitative estimates of risk, the Monte Carlo curves should look the same as the original validation plots. The general shapes in all three cases are fine, but the values of D_{max} as given in Table 12 are shifted for both logit models. This is undoubtedly due to the greater degree of selection for the items comprising the logit scales—and thus also for the Burgess scale items. Note that the logistic Model III curve and D_{max} are quite similar to those seen in Figure 8. Thus, when a great deal of selection occurs as in the Burgess scale construction, the usual construction-validation splitting is not enough. Double cross-validation (Mosteller and Tukey 1977) consisting of using separate samples for (1) choosing items, (2) estimating coefficients, and (3) testing the scale likely would have reduced the overall effects of selection.

The practical implications of the data analysis results reported here are (1) to question the efficiency of the construction-validation splitting as currently practiced and (2) to offer a technique, logistic regression, whose assumption requirements are more in keeping with the data as collected. Related to the first point, we recommend use of double cross-validation when data are plentiful, and consideration of jackknifing techniques (Mosteller and Tukey 1977) when more must be squeezed from limited data. For the logistic regression versus Burgess scale question, only experience with both will decide which, if either, is better. We recommend their joint consideration in any future projects, and certainly feel that logistic regression will prove to be a better competitor than linear regression has been in the past.

[1] Computations were performed using FORTRAN programs on a CDC 6400 computer. Bernoulli random variables were generated from uniform random numbers by classifying the uniforms into one of two categories using the boundaries $(0, \hat{p}_i, 1)$. The uniform random numbers were produced by a multiplicative congruential generator using modulus 2^{48} and multiplier 5^{17}.

Figure 5.
Validation — Burgess Scale

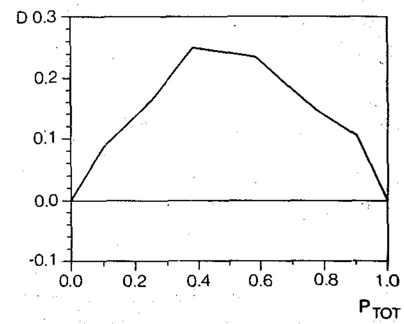


Figure 6.
Validation — Logistic I

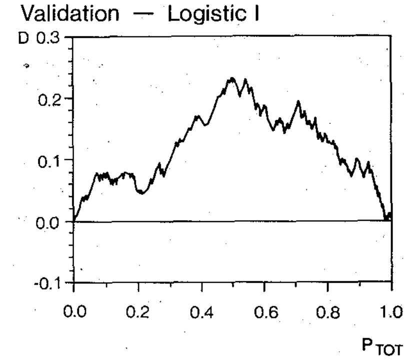


Figure 7.
Validation — Logistic II

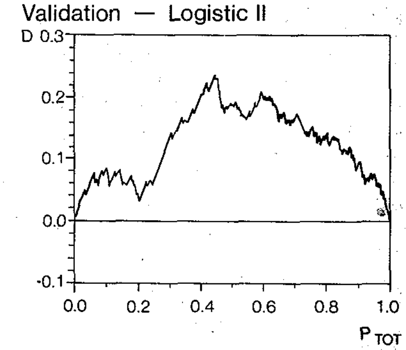


Figure 8.
Validation — Logistic III

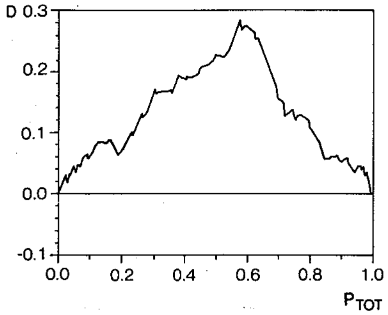
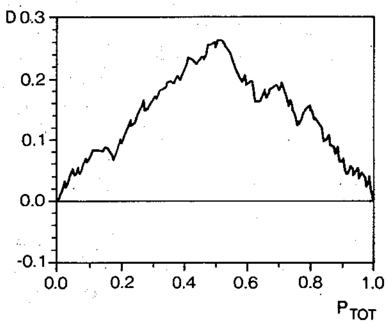


Figure 9.
Validation — Logistic IV



Acknowledgments

This research project grew out of material on the Federal Parole Project presented at the Workshop on Criminal Justice Statistics, held in Washington, D.C., July 1975. The workshop was sponsored by the Social Science Research Council Center for Coordination of Research on Social Indicators and the Law Enforcement Assistance Administration. The sponsors also provided additional financial assistance for this project. I am grateful to Sanford Weisberg, Ronald Christensen, and Sharon Yang for collaborative efforts in the early stages of this research. In addition, Stephen E. Fienberg provided valuable detailed comments on both the research and its presentation. The University of Minnesota Computer Center provided a grant for computing costs. Special thanks are owed to Dale Parent, Minnesota Department of Corrections, who provided the raw data and was particularly helpful in clarifying the coding procedures.

References

Cox, D. R. (1970)
The analysis of binary data. London: Methuen.

Dempster, A. P., N. M. Laird, and D. B. Rubin (1977)
"Maximum likelihood from incomplete data via the *EM* algorithm." *Journal of the Royal Statistical Society* 39:1-38.

Efron, B. (1978)
"Regression and ANOVA with zero-one data: Measures of residual variation." *Journal of the American Statistical Association* 73:113-121.

Fienberg, S. E. (1977)
The analysis of cross-classified categorical data. Cambridge, Mass.: M.I.T. Press.

Gottfredson, D. M., C. A. Cosgrove, L. T. Wilkins, J. Wallenstein, and C. Rauh (1978)
Classification for parole decision policy. Washington, D.C.: U.S. Government Printing Office.

Gottfredson, D. M., L. T. Wilkins, P. B. Hoffman, and S. M. Singer (1974)
The utilization of experience in parole decision-making: Summary report. Washington, D.C.: U.S. Government Printing Office.

Hoffman, P. B., and L. K. De Gostin (1974)
"Parole decision-making: Structuring discretion." *Federal Probation* (December).

Koehler, K. J., and K. Larntz (1980)
"An empirical investigation of goodness-of-fit statistics for sparse multinomials." *Journal of the American Statistical Association* 75:336-344.

Larntz, K. (1978)
"Small-sample comparisons of exact levels for chi-squared goodness-of-fit statistics." *Journal of the American Statistical Association* 73:253-263.

Mosteller, F., and J. W. Tukey (1977)
Data analysis and regression. Reading, Mass.: Addison-Wesley.

Nerlove, M., and S. J. Press (1973)
"Univariate and multivariate log-linear and logistic models." Technical Report R-1306-ECA/NIH, Rand Corporation, Santa Monica, Calif.

Parent, C. G., and T. G. Mulcrone (1978)
"The development and operation of parole decisionmaking guidelines in Minnesota." In D. M. Gottfredson, C. A. Cosgrove, L. T. Wilkins, J. Wallenstein, and C. Rauh, *Classification for parole decision policy,* Washington, D.C.: U.S. Government Printing Office.

Press, S. J. and S. Wilson (1978)
"Choosing between logistic regression and discriminant analysis." *Journal of the American Statistical Association* 73:699-705.

Solomon, H. (1976)
"Parole outcome: A multidimensional contingency table analysis." *Journal of Research in Crime and Delinquency* (July) 107-126.

Figure 10.
Monte Carlo — Logit I

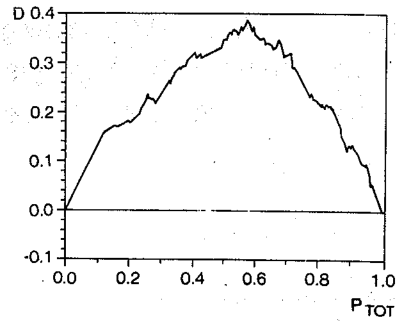


Figure 11.
Monte Carlo — Logit IV

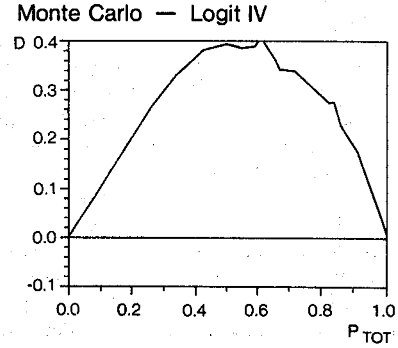


Figure 12.
Monte Carlo — Logistic III

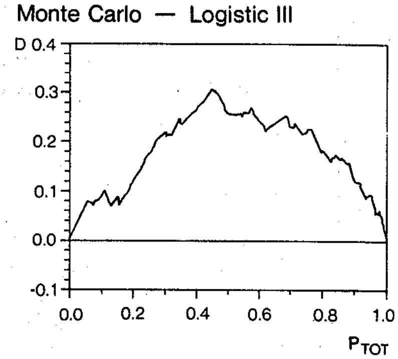


Table 12. Maximum D for Monte Carlo samples from three predictive scales

Model	Monte Carlo D_{max}	Validation D_{max}
Logit I	.3885	.2480
Logit IV	.4072	.2729
Logistic III	.3075	.2844

The effects of plea bargaining on the disposition of person and property crimes: A research note

DAVID W. BRITT
Criminal Justice Program
Nova University

KINLEY LARNTZ
Department of Applied Statistics
University of Minnesota

Many of the studies which have examined the impact of offender and offense characteristics on sentence severity have been misspecified because they have ignored (or tried to hold constant) events such as pretrial plea bargaining which took place at an earlier stage in the process but which have a critical bearing on the outcome of sentencing decisions. Our approach to the analysis of these phenomena is similar to the exploratory study by Burke and Turk (1975)in that we are considering a sample of all post-arrest dispositions. Our study differs from theirs in that we attempt to include the impact of plea bargaining on subsequent sentencing decisions.

THE TESTING OF WHAT PASSES for conventional wisdom often leads to conclusions which are apparently paradoxical. Such is the case for recent attempts to structure and analyze the bias in criminal justice systems. As Burke and Turk (1975:313) have recently noted, the conventional wisdom of criminology has held that the socially disadvantaged are more likely than the advantaged to be engaged by the criminal justice system, to be prosecuted, tried, convicted, and to suffer more severe penalties upon conviction. Yet the findings of empirical analyses have been both ambiguous and occasionally at odds with the received wisdom of common experience—let alone the alternative paradigms for viewing the operation of the criminal justice system. For example, a recent analysis of the construction of judicial decisions has observed that for a sample of judges having flexible attitudes toward law and order, blacks and those with larger numbers of current charges were more likely to receive short sentences than whites and those with smaller numbers of current charges (Hagen 1975:379). Such results are at odds both with conceptions of the criminal justice system which stress a bias against the disadvantaged *and* conceptions which emphasize offense characteristics. It is apparent that models of post-arrest disposition have omitted variables which are crucial to an understanding of the dynamics of the disposition process, and it is these specification errors which have resulted in some of the paradoxical results observed in recent studies.

Faced with the problem of disentangling the serial processes which underlie the processing of "criminals" within the criminal justice system, much current research has sought to simplify the task by examining discrete aspects of this process as a first step in the ultimate integration of these separate analyses. Such an attack on the problem would be reasonable if the same set of factors affected dispositions at the various points in the process and if the outcomes at one stage had no independent bearing on the operation of these factors at a later stage. Unfortunately, such assumptions appear to be invalid. Consequently, many of the studies which have examined the impact of offender and offense characteristics on sentence severity have been misspecified because they have ignored (or tried to hold constant) events, such as pretrial plea bargaining, which took place at an earlier stage in the process but which have a critical bearing on the outcome of sentencing decisions.

A thorough review of the entire criminal justice process as it is affected by characteristics of the offender, the offense, and variation in the structure at different points is beyond the scope of this paper. Suffice it to say that the empirical works on sentence length (Hagen 1975; Burke and Turk 1975; Tiffany et al. 1975, for example) have been unable to come up with plausible interpretations for what appear to be bizarre patterns in the operation of offender characteristics (race, sex, age, prior convictions), characteristics of the offense (severity of offense, number of current charges) and system characteristics (type of counsel, bench vs. jury trial). What is missing from all of these studies is some notion of the operation of pre-trial plea bargaining on the outcome of the sentencing decision. Some of these studies (Hagen 1975; Tiffany et al. 1975) have sought to minimize the influence of pre-trial factors by limiting the sample of cases processed by the system to those above a minimum level of seriousness which have reached trial. Since fewer than 10 percent of the initial cases reach trial in most jurisdictions (Blumberg 1967; Heismann 1975), and since cases which reach trial are biased in that they have rejected the plea-bargaining alternative (Tiffany et al. 1975), this would seem to be an unacceptable way of controlling the influence of pre-trial factors.

Our approach to the analysis of these phenomena is similar to the exploratory study by Burke and Turk (1975) in that we are

Table 1. Results of testing the hypothesis that the variable is independent of incarceration given seriousness and type of crime.

Variable	df	G^2	X^2	T^2	P-value(X^2)
Sex	6	13.31	9.23	11.66	.25> p>.10
Race	8	8.65	7.40	7.56	p>.25
Plea	15	30.58	24.48	27.01	.10> p>.05
Charge	10	25.03	20.38	21.81	.05> p>.025
No. initial charges	7	2.61	2.70	2.66	p>.25
No. convicted charges	6	6.96	4.77	5.82	p>.25
No. previous arrests	9	32.23	28.31	29.31	p<.005
Felony convictions	10	41.08	35.43	41.88	p<.005
Employment status	8	8.22	7.69	7.28	p>.25
Resisting arrest	8	12.44	10.99	10.80	.25> p>.10
Weapon usage	7	10.84	8.32	9.03	p>.25

Table 2. 138 Criminal sentences classified by type of crime, seriousness of charge plea, change in charge, previous record and incarceration

Type of crime, seriousness and change	Plea	Good previous record: No jail	Good previous record: Jail	Fair previous record: No jail	Fair previous record: Jail	Poor previous record: No jail	Poor previous record: Jail
Person, no change of charge	Guilty	0	6	1	3	2	5
	Nolo contendere	2	0	1	2	2	0
	Not guilty	2	2	0	0	0	0
Person, charge changed	Guilty	2	1	3	1	4	5
	Nolo contendere	5	1	2	0	0	1
	Not guilty	0	0	0	0	0	0
Property, low seriousness, No change of charge	Guilty	2	0	3	1	1	2
	Nolo contendere	3	0	1	1	0	0
	Not guilty	4	0	1	0	0	0
Property, low seriousness, charge changed	Guilty	3	0	2	0	1	2
	Nolo contendere	0	0	0	0	1	0
	Not guilty	0	0	0	0	0	0
Property, moderate seriousness, no change of charge	Guilty	1	0	1	0	3	1
	Nolo contendere	3	0	1	0	0	1
	Not guilty	3	0	0	0	0	1
Property, moderate seriousness, charge changed	Guilty	4	0	1	0	0	6
	Nolo contendere	2	0	1	0	1	2
	Not guilty	0	0	0	0	0	0
Property, serious crime, no change of charge	Guilty	0	0	1	1	0	1
	Nolo contendere	0	0	0	0	0	0
	Not guilty	0	0	1	0	0	0
Property, serious crime, charge changed	Guilty	1	0	1	2	1	3
	Nolo contendere	3	0	0	0	0	0
	Not guilty	1	0	0	0	0	0
Property, very serious crime, no change of charge	Guilty	0	0	0	0	0	0
	Nolo contendere	0	0	0	1	0	0
	Not guilty	0	0	0	1	0	0
Property, very serious crime, charge changed	Guilty	0	0	0	2	1	2
	Nolo contendere	1	0	0	1	0	0
	Not guilty	0	0	0	0	0	0

considering a sample of all post-arrest dispositions. Our study differs from theirs in that we are attempting to include the impact of plea bargaining on subsequent sentencing decisions.

Data

Two hundred sample cases were selected from the Denver, Colorado, courts under the auspices of State Courts Sentencing Project. Sixty-two of these cases were eliminated from this sample: 5 because of inadequate data and 57 because of their being so-called "victimless" crimes. The remaining 138 cases were analyzed to determine the major factors used by judges in reaching the decision of whether or not to incarcerate an individual convicted of a criminal offense. Thirteen variables were initially considered as possible factors:

(1) *Sex.*
(2) *Race:* White, other.
(3) *Plea:* Guilty, nolo contendere, or not guilty.
(4) *Seriousness of charge:* Low, moderate, serious,or very serious.
(5) *Change in charge:* Is the convicted charge the same as the initial charge or was there some reduction in charge?
(6) *Number of initial charges.*
(7) *Number of convicted charges.*
(8) *Number of previous arrests:* Low (0, 1, 2) or high (3 or more).
(9) *Number of adult felony convictions:* Low (0) or high (1 or more).
(10) *Employment status:* Was the individual employed or unemployed at the time of the offense?
(11) *Type of crime:* (a) Against property, (b) against person but at most slight injury, or (c) against person with serious injury.
(12) *Resisting arrest:* Did the individual resist arrest?
(13) *Weapon usage:* Did the crime involve use of a weapon?

The dependent variable was *incarceration*, i.e., whether or not the individual was sent to jail by the judge.

Analysis and results

With the number of variables that we deemed important to consider and the relatively small sample size, we opted to make an initial pass through the data to select the most important factors. Our review of the literature showed that *seriousness of charge* and *type of crime* were both important factors in explaining the number of individuals jailed. To decide which of the other variables to include, each variable was tested to see if the hypothesis of conditional (given *seriousness* and *type of crime*) independence between the variables and *incarceration* was plausible. The chi-square statistics and p-values for these tests are given in Table 1. The four variables yielding the lowest p-values were selected for further analysis.

In the initial data analysis it became clear that sentences for crimes against property had a different structure compared to sentences for crimes against persons. To deal with this structural dissimilarity, the analysis was split into two parts: crimes against property and crimes against persons. Table 2 gives the data as used in the final analysis.

Table 3. Results of logit model fittings for crimes against property

Variables included in model	df	G^2	X^2	T^2
Previous record, plea, charge seriousness	30	23.00	21.49	17.79
Previous record, plea, charge	33	29.79	22.97	23.91
Previous record, plea, seriousness	31	23.14	21.56	18.09
Previous record, charge, seriousness	32	23.02	21.57	17.75
Plea, charge, seriousness	32	57.59	49.33	46.21
Previous record, plea	34	31.11	24.21	25.35
Previous record, charge	35	29.94	23.17	24.03
Previous record, seriousness	33	23.20	21.56	18.11
Plea, charge	35	67.11	57.24	54.73
Plea, seriousness	33	57.59	49.28	46.25
Charge, seriousness	34	63.00	51.19	52.23
Previous record	36	31.12	24.22	25.35
Plea	36	68.07	57.14	55.97
Charge	37	70.74	55.28	59.60
Seriousness	35	63.61	51.73	53.06
None	38	73.40	56.59	62.89

*$X^2 = \Sigma\,(\text{Obs} - \text{Exp})^2/\text{Exp}$

$T^2 = \Sigma\,(\sqrt{\text{Obs}} + \sqrt{\text{Obs}+1} - \sqrt{4\,\text{Exp}+1})^2$

$G^2 = 2\,\Sigma\,\text{Obs}\,\log_e(\text{Obs}/\text{Exp})$

When cell expectations are small, as they are in our examples, we examine all three chi-square statistics to evaluate the model fit (Larntz, 1975).

Table 4. Results of logit model fitting for crimes against persons

Variables included in model	df	G^2	X^2	T^2
Previous record, charge, plea	7	13.82	12.81	11.09
Previous record and charge	9	19.90	16.27	18.07
Previous record, and plea	8	18.41	16.02	15.53
Charge and plea	9	13.96	13.11	11.21
Previous record only	10	24.54	19.28	22.96
Charge only	11	20.84	16.84	18.72
Plea only	10	18.47	15.86	15.76
None	12	25.13	19.72	23.33
Record x plea + charge	3	12.15	11.05	9.41
Record x charge + plea	5	10.17	8.36	8.27
Charge x plea + record	6	12.77	10.76	10.22
Charge x plea only	8	12.90	11.13	10.23

*See footnote for Table 3.

Note that crimes against persons were not classified by seriousness: this was because almost all of these crimes were classified as serious offenses. Also *number of previous arrests* and *felony convictions* were combined into a new variable, *previous record*. A "poor" previous record meant the individual had one or more adult felony convictions. A "fair" record resulted from at least three previous arrests but no felony convictions. An individual had a "good" record if there were at most two recorded arrests with no adult felony convictions.

The procedure for analyzing the data in Table 2 was to find the logit variant (Dyke and Patterson 1952, Goodman 1970) of the loglinear model (Bishop, Fienberg, and Holland 1975; Haberman 1974) that provided the closest fit to the data.

Table 3 presents the results for the models fit to the *property* crime sentences, and Table 4 similarly presents the results for the *person* crime sentences. By examining the significance of the overall goodness-of-fit chi-squares as well as the differences between the chi-square statistics when one model is a submodel of the other, inferences may be drawn as to the major factors influencing the incarceration decision.

For crimes against *property*, the logit model including only *seriousness of charge* and *previous record*—with the latter playing a more important role—provides an adequate fit to the data. Table 5 gives the estimated percentage incarcerated based on these two variables. Those with poor records who have been convicted of a very serious crime have roughly a 95 percent chance of being incarcerated. The impact of *previous record* on incarceration' chances may be better appreciated by examining two other cells in Table 5. Even if the seriousness of the present crime was quite low, if the prior record of the convicted defendant was poor, he/she had almost a 60 percent chance of going to jail. But if the prior record of the defendant was good, he/she had only negligible chance of going to jail no matter how serious the present crime.

For crimes against *persons*, where almost all crimes were categorized as being serious, the logit model also includes just two variables: *change* and *plea*. Table 6 gives the estimated incarceration rates for *person* crimes.

Of those who either chose or were not allowed to enter into plea bargaining (at least with respect to an exchange of charge reduction for a plea of guilty), those who pleaded guilty had the highest chance of going to jail (78 percent). Those who were able to plead nolo contendere had only a 39 percent chance of going to jail. Pleading guilty was associated with a 50 percent chance of going to jail. These figures are dramatically altered by the apparent exchange of plea for charge reduction. For those who plead guilty with a reduced charge, the chances of being incarcerated drop from 78 percent to 48 percent. A similar drop in incarceration chances is noted for those pleading nolo contendere: from 39 percent to 14 percent.[1]

Discussion and conclusions

Dispositions of crimes against property appear to be related to variation in the seriousness of the crime and the offender's previous record in a relatively straightforward way. As the crime becomes more serious and the offender's previous record becomes poorer, the chances of the individual's being incarcerated increase. Interestingly, as long as the previous record of the individual has not involved more than two arrests and no adult felony convictions, there seems to be a bias against sending the individual to jail—no matter how serious the present property crime is. Once the individual has been arrested more than twice previously, or has had a previous felony conviction, his chances of being incarcerated for the present offense escalate rapidly. But this appears to be the only bias in the property-crime dispositions, for neither characteristics of the offender (sex, race, employment status) nor his pleas or charge reductions appear to directly affect the decisionmaking process.

These results for property crimes overlap substantially with those of Burke and Turk (1975:326), who noted that the effect of previous incarceration on disposition severity remained important even after controlling for the nature of the offense.[2] They argued that such a pattern increased the tenability of an explanation based on ex-convicts' "vulnerability to the biases of legal control agents" and initiated an argument based on the presumed greater "propensity of ex-convicts for relatively serious crimes." In other words, even though there may be an association between prior record and offense seri-

[1]We should note here that the length of sentence has not been taken into consideration. Hence, while these data suggest a rational strategy of pledging nolo contendere or of guilty if the plea bargaining option is unavailable, a different pattern might have emerged had we been able to take length of sentence into account.

[2]Burke and Turk (1975) found that age also directly affected the dispositional process, although the impact of age was complicated and not amenable to nonspeculative, plausible explanation.

Table 5. Estimated percentage incarcerated for crimes against property

		Seriousness of charge			
		Low	Moderate	Serious	Very serious
Prior record	Good	0 (12)	0 (13)	0 (5)	0 (1)
	Fair	21 (9)	27 (3)	48 (6)	82 (5)
	Poor	58 (7)	66 (15)	83 (5)	96 (3)

Note: Estimated percentages are calculated from logit model:

$$\log \frac{P_{ij}}{1-P_{ij}} = \mu + a_i + \beta_j$$

where P_{ij} is the probability of incarceration.

ousness, each makes a contribution to the reconstruction of the observed distribution of incarceration rates.

Our results speak to the possibility of a threshold of vulnerability to bias on the part of legal control agents. Some degree of "criminality credit" appears to be granted to individuals who have had "good" prior records, so that individuals appear to be buffered from jail sentences until their contact with the machinery of the state has been relatively frequent or serious. If these results have any validity, one implication is that socially disadvantaged individuals are not being directly discriminated against at this point in the dispositional process. It is quite possible, of course, that variation in both the seriousness of crimes and poorness of records could be traced to both race and occupational status (Burke and Turk 1975). It is also possible that more complicated interactions of race and age might affect the propensity to commit serious offenses (Black and Reiss 1970) and/or the poorness of prior record. But in our admittedly small sample (as well as in that of Burke and Turk), blacks do not appear to be *differentially* vulnerable to being labelled as a "repeat offender," or "career criminal." Were this the case, we should have at least observed an interaction of race and prior record in the disposition process.

A somewhat different picture emerges regarding the structuring of dispositions of crimes against persons. To recapitulate, charges were reduced in exchange for both nolo contendere and guilty pleas, and as expected, engaging in the plea-bargaining process decreased one's chances of being sent to jail for both kinds of pleas. Those who came out of these negotiations having to plead guilty had a greater chance of going to jail than those who were able to enter nolo contendere pleas. Where plea bargaining did not take place, those who pleaded guilty had the highest chance of being imprisoned, those who pleaded nolo contendere had the least chance, and those who entered not guilty pleas were in the middle.

It must be remembered that at least for our sample of 200 cases, almost all of the crimes against persons were considered to be relatively serious. Consequently, seriousness could not play much of a role in the dispositional process. Offender characteristics, on the other hand, varied but did not affect the chances of being jailed once the charge-reduction and plea process was taken into consideration.

In sum, the dispositional dynamics for person and property crimes appear to be different, with plea bargaining having a much greater impact on the former than the latter. In neither case does the dispositional process appear to be biased once plea bargaining is incorporated into the analysis. What bias there is operates with respect to the defendant's previous record—and even here there is evidently a buffering process at work until the individual's contact with the criminal justice system has become frequent and/or felonious.

To argue that offender characteristics (sex, race, previous record) do not have a direct impact on incarceration rates for crimes against persons does not mean that they are not involved in the dispositional process. It's quite possible that a defendant's prior record, in conjunction with other indicators of his recidivism chances and the solidness of the state's case against him, set boundaries for the kinds of plea bargaining which can be struck. In our future investigations we will be examining these possibilities.

Table 6. Estimates of the incarceration rates for crimes against persons

		No change in charge	Charge reduced
Plea	Guilty	78	48
	Nolo contendere	39	14
	Not guilty	50	*

Note: Estimated percentages are calculated from logit model:

$$\log \frac{P_{ij}}{1-P_{ij}} = \mu + a_i + \beta_j$$

Where P_{ij} is the probability of incarceration.

*No information in data

Acknowledgments

This paper is an outgrowth of a workshop on criminal justice systems sponsored by the Social Science Research Council, July–August 1975. The authors would like to thank Leslie T. Wilkins for providing the data upon which this paper is based.

References

Bishop, Y. M. M., S. E. Fienberg, and P. W. Holland (1975)
Discrete multivariate analysis. Cambridge, Mass.: M.I.T. Press.

Black, Donald, and A. J. Reiss, Jr. (1975)
"Police control of juveniles." *American Sociological Review* 35:63–77.

Blumberg, Abraham S. (1967)
Criminal justice. Chicago: Quadrangle Books.

Burke, Peter J., and Austin T. Turk (1975)
"Factors affecting postarrest dispositions: A model for analysis." *Social Problems* 22:313–32.

Dyke, G. V., and H. D. Patterson (1952)
"Analysis of factorial arrangements when the data are proportions." *Biometrics* 9:1–12.

Goodman, L. A. (1970)
"The multivariate analysis of qualitative data: Interactions among multiple classifications." *Journal of the American Statistical Association* 65:226–56.

Haberman, S. J. (1974)
The analysis of frequency data. Chicago: University of Chicago Press.

Hagan, John (1975)
"Law, order and sentencing: A study of attitude in action." *Sociometry* 38:374–84.

Heumann, Milton (1975)
"A note on plea bargaining and case pressure." *Law & Society Review* 9:515–28.

Larntz, K. (1975)
"Small sample comparisons of exact levels for chi-square goodness-of-fit statistics." Technical Report No. 242. University of Minnesota School of Statistics.

Tiffany, Lawrence P., Yakov Avichai, and Geoffrey W. Peters (1975)
"A statistical analysis of sentencing in federal courts: Defendants convicted after trial, 1967–1968." *Journal of Legal Studies* 4:369–90.

The deterrent effects of punishment

Alternative estimates of the impact of certainty and severity of punishment on levels of homicide in American states

COLIN LOFTIN
Department of Sociology and Center for Research on Social Organization University of Michigan

Jack Gibbs' paper, "Crime, Punishment, and Deterrence," has been widely cited as an important part of an accumulating body of evidence which consistently demonstrates a negative relationship between crime rates and the certainty and severity of punishment. In this paper we provide a new set of estimates of the effects of the certainty and severity of punishment on homicide rates. The estimates are derived using a structural model of violent crime which was developed independently of Gibbs' research. When the effects of the other variables in the model are taken into account, the estimated effects of the punishment variables are very small. The data are consistent with a model in which certainty and severity of punishment have no effects on homicide rates.

Background

In 1968 Jack Gibbs presented the results of an influential study of the effects of certainty and severity of punishment on murder rates in American states. While he was appropriately cautious in interpreting his estimates as doing no more than challenging the "common assertion that no evidence exists of a relationship between legal reactions to crime and the crime rate" (Gibbs 1968:529–530), subsequent discussions have interpreted his study as providing support for deterrence effects (see, for example, Tullock 1974:107, Tittle and Logan 1973, Antunes and Hunt 1973, Tittle 1975). Gray and Martin (1969) and Bean and Cushing (1971) reanalyzed Gibbs' data, with slight modifications in statistical procedures and theoretical model, and concluded that the data are consistent with a model which includes direct negative effects of certainty and severity of punishment on murder rates. More recent studies of the effects of law enforcement activity on crime rates have gone considerably beyond these studies in terms of theoretical specifications and estimation procedures. Nevertheless the estimates of deterrence effects from these early studies continue to be cited in current discussions as part of an accumulating body of evidence supporting the existence of general deterrence effects. In this paper we provide a new set of estimates using a model which was specified independently of Gibbs' research, and which includes a much larger number of *etiological* (or *environmental*) variables than do any of the other deterrence studies. When the effects of these *etiological* variables have been removed from Gibbs' measures of the certainty and severity of punishment, the estimated effects of the punishment variables are very small and provide reason to question whether the data provide any support for real effects of certainty and severity of punishment on homicide rates.

Before presenting the new estimates it will be useful to briefly review the results of the other studies of Gibbs' data. Gibbs' original paper specifies the following general model for crime rates:

$$C = f(E,R)$$

where C is the crime rate; E represents *etiological* factors, that is, "extralegal conditions which are conducive to crime ..." (Gibbs 1968:517); and R represents *repressive* factors, that is "aspects of reaction to crime which operate as deterrents ..." (Gibbs 1968:517). However, his actual investigation of homicide rates fails to include any *etiological* factors. He estimates the relationship between the two repressive factors and homicide rates while ignoring *etiological* factors. The two repressive factors specified in the study are:

(1) Estimated severity of sentence received and served for criminal homicide, which he measured with the median months served on a homicide sentence by persons in a state prison on December 31, 1960 (Gibbs 1968:520–521).

(2) Estimated certainty of imprisonment for criminal homicide, which he measured as the number of persons admitted to state prisons in 1960 on a homicide sentence, divided by the average number of criminal homicides reported by the police in 1959 and 1960. The dependent variable is the average annual criminal homicide rate per 100,000 population for 1959 to 1961.[1]

Gibbs' statistical analysis is done with contingency tables by dividing each distribution at the median. From this analysis, he concludes that both certainty and severity of punishment are negatively related to state homicide rates.

[1] New Jersey's values for certainty and severity were estimated from the average of New York and Connecticut.

Table 1. Selected results of previous analyses of Gibbs' Data

Gray and Martin (1969)

(1) $H = \hat{\beta}_0 - .103\ P(I) - .37\ T + u$

$R^2 = .219$

(2) $1nH = \hat{\beta}_0 - .350\ 1nP(I) - .551\ 1nT + u$

$R^2 = .377$

(3) $1nH = \hat{\beta}_0 - 1.073\ 1n[P(I)*T] + u$

$R^2 = .359$

Bean and Cushing (1971)

(1) $H = \hat{\beta}_0 \quad - .289\ P(I) - .374\ T + u$

$\hat{\beta}/S\beta = (-2.19) \quad (-2.84)$

$R^2 = .218$

(2) $H = \hat{\beta}_0 \quad - .196\ P(I) - 2.05\ T + .717\ R + u$

$\hat{\beta}/S\hat{\beta}_0 = (-2.39) \quad (-2.33) \quad (8.31)$

$R^2 = .696$

(3) $H = \hat{\beta}_0 \quad - .179\ P(I) - .167\ T + .77\ B + u$

$\hat{\beta}/S\hat{\beta}_0 = (-22.38) \quad (-2.16) \quad (9.87)$

$R^2 = .751$

Notes: The regression coefficients are for variables in standard form, that is, they are beta weights. H = homicide rate, P(I) = probability of imprisonment (Gibbs' measure of certainty of punishment), T = median time served (Gibbs' measure of severity), R = region (South is coded 1, nonsouth is coded 0), and B = percent of population black. Gray and Martin do not provide estimates of standard errors or t values. Their coefficient for P(I) in equation appears to be an error. Our own analysis is consistent with Bean and Cushing's results.

Gray and Martin (1969) extended Gibbs' study by investigating several different forms of the relationships between certainty, severity, and homicide in an ordinary least squares regression analysis. Table 1 summarizes the major elements of their analysis which led them to conclude that the data are consistent with a model in which certainty and severity of punishment operate in a nonadditive way to reduce the homicide rate (Gray and Martin 1969: 394–395). Note that their investigation, like Gibbs', derived estimates of certainty and severity effects under the implicit assumption that *etiological* effects can be ignored. That is, no *etiological* factors were included in the estimated model.

Bean and Cushing (1971) recognized the importance of specifying a model which allows for *etiological* as well as *repressive* factors, but the models that they investigated included only one *etiological* variable at a time. They found that when a variable representing region (South versus Nonsouth) was included in a linear specification (see Table 1 above), the estimates of the certainty and severity effects were reduced, but that they were still large enough to be "consistent with the deterrence hypothesis" (Bean and Cushing 1971:289). Also they conclude that the linear specification which included region was as adequate as alternative nonlinear specifications of the model. Finally, they substitute "percent of the population that is black" for "region" in their model and find that the proportion of the variance "uniquely accounted for by the etiological factors" increased, and thus that the data are consistent with the hypothesis that the *etiological* significance of region is due to percent black (Bean and Cushing 1971: 288).

Other studies which have estimated deterrence effects for homicide, but which use data different from Gibbs, include Tittle (1969), Chiricos and Waldo (1970), Antunes and Hunt (1973), Logan (1972, 1975), Bailey, Martin, and Gray (1974). All of these studies, except Logan (1975), estimate significant negative deterrence effects, but none of them includes more than one variable to represent *etiological* factors and most of them do not include any *etiological* variables. The major exception to this generalization is Ehrlich (1973, 1975), who reports several investigations of the effects of law enforcement variables on homicide rates, all of which use models which take into account more than one *etiological* factor, and allow for simultaneous relationships between crime rates and law enforcement variables. His results are consistent with most other studies in that his deterrence variables are negatively and significantly associated with homicide rates.[2]

[2]Several recent evaluations of Ehrlich's findings raise serious questions about the sensitivity of his results to minor changes in the model (see, for example, Bowers and Pierce 1975).

New estimates of certainty and severity of punishment on homicide rates

Sociological theories of homicide, though not incompatible with utilitarian formulations which characterize most deterrence studies, suggest that homicide is an extreme manifestation of the "culture of violence" which is generated by high levels of personal frustration and extreme socioeconomic deprivation (Wolfgang 1958, Wolfgang and Ferracuti 1967, Brearly 1932, Coser 1967). Some crimes of violence may be incidental to participation in other forms of illegal activity, but most homicides are known to occur among friends, family, and acquaintances and are not byproducts of other crimes (Wolfgang 1958, 1968). They are likely to be acts of passion that grow out of high levels of frustration and subcultural reinforcement of interpersonal violence.

A previous study by Loftin and Hill, drawing on studies done by Gastil (1971) and Hackney (1969), specified a model which is derived from sociocultural explanations of homicide rates. The key element in this model is a six-variable index referred to as an index of "structural poverty" which redundantly measures the proportion of a state's population at the extremely low end of the socioeconomic class distribution. The components of that index are:

(1) Infant Mortality Rate (Grove and Hetzel 1968: Table 41).

(2) Percent of persons 25 years old and over with less than 5 years of school (Renetzky and Greene 1970:38).

(3) Percent of families with income under $1,000 (U.S. Bureau of the Census 1964a:Table 137).

(4) Percent of the population illiterate (Greene and Renetzky 1970:38).

(5) Armed Forces Mental Test Failures (August 1958–December 1965) (*American Education* 1966:9).

(6) Percent of children living with one parent (U.S. Bureau of the Census 1964b:68).

These variables account for 87 percent of the variance in 1959–1961 mean homicide rates in Loftin and Hill's analysis.[3]

[3]Loftin and Hill's homicide rate is based on the U.S. Vital Statistics rather than on the Uniform Crime Report and thus is slightly different from Gibbs' homicide rate. See Loftin and Hill (1974:718 and 720) for exact definitions. The correlation between the two homicide rates is 0.976.

Table 2. Correlations between Gibbs' measures of the homicide rate, the certainty of punishment, the severity of punishment and the six components of the Loftin-Hill Index

	Homicide rate	Certainty of punishment	Severity of punishment
Infant mortality rate (IMR)	.832	−.222	−.332
Percent of persons with less than 5 years of education (LOWED)	.825	−.166	−.241
Percent of population illiterate (ILLIT)	.768	−.176	−.291
Armed forces mental test failures (TESTFAIL)	.836	−.126	−.292
Percent of families with less than $1,000 income (LOWINC)	.713	−.160	−.065
Percent of children living with one parent (ONEPARENT)	.877	−.292	−.399

Table 2 shows the correlations between the components of the index and Gibbs' three variables (homicide rate, certainty of punishment, and severity of punishment). Note that the components are all positively correlated with the homicide rate and negatively correlated with the certainty and severity of punishment. This suggests a model, which Gibbs acknowledges (1968: 528) but does not investigate, in which sociocultural variables, such as those measured in the Loftin-Hill index (*etiological* factors in Gibbs' terminology), operate to increase homicide rates in some states and simultaneously to reduce the certainty and severity of punishment. Little is known about the determinants of variables such as Gibbs' certainty and severity, but it is plausible that they are reduced when a significant proportion of the population is extremely poor and uneducated, both because of the direct effects of the sociocultural factors and because of common causes. This suggests that estimates of deterrence effects, to the extent that they have not taken social and cultural variables into account, may have overestimated the size of these effects.

The complete specification of the homicide model that Loftin and Hill use and which we will use in the present analysis includes six other variables:

(7) Dye's Gini Coefficient of income inequality for American states (Dye 1967).

(8) Region (a binary variable coded 1 for former Confederate states, 0 otherwise).

(9) Percent of population nonwhite (U.S. Bureau of the Census 1964a:Table 56).

(10) Percent of the population age 20 to 34 (U.S. Bureau of the Census 1964b:23).

(11) Percent of the population living in rural territory (U.S. Bureau of the Census 1964a:Table 21).

(12) Number of hospital beds per 100,000 population (U.S. Bureau of the Census 1962:80).

To derive our estimate of the effects of Gibbs' repressive factors we simply use Gibbs' homicide rate as the dependent variable and add his measures of the certainty and severity of punishment as independent variables, along with those used in the previous study (Loftin and Hill 1974). It should be noted that while there is an element of arbitrariness in the selection of the sociocultural variables to be included in the homicide model, the model that we use in this analysis was developed independently of research on deterrence, and thus no selection of variables which maximize or minimize the effects of the repressive factors was possible.

Our use of ordinary least squares regression analysis to estimate the parameters of the model implies that the homicide rate is not simultaneously determined with the sanction variables. While this may not generally be the case it is plausible that, at least in the case of homicide and for the range of variation represented by American states in about 1960, repressive variables are unresponsive to homicide rates. Homicide is a rare crime with a very high clearance rate. In 1959 the clearance rate for murder and nonnegligent manslaughter for all geographic divisions of the United States, according to the Uniform Crime Reports, was 92.7 with a standard deviation of only 3.1 across the nine Uniform Crime Reports' geographic divisions.[4] Large increases in the amount of violent crime might reduce the effectiveness of police investigations and strain other resources of the criminal justice system, but the analogy with a riot situation where the probability of arrest and prosecution declines as the number of crimes increases does not seem appropriate for the situation in the United States in about 1960. Contemporary Northern Ireland and Lebanon probably fit the model, but they represent very extreme cases far outside of the range of variation represented in Gibbs' data.

It is also unlikely that judges, juries, and parole boards adjust sentences on the basis of the homicide rates. It is more plausible that these decisions are made primarily on the basis of judgments of the seriousness of the crime and the probability that offenders are committed to violent crime as a way of life, along with the particular circumstances of the crime (see, for example, Wilkins 1974:244–245). Thus the degree of criminal intent and extenuating circumstances as judged by juries and parole boards are probably the key factors, not the general level of homicide.

Certainly general public sentiments and public opinion will have an impact on punishment levels, but again public opinion is probably responsive primarily to social and cultural variables rather than to the homicide rate.

The hypothesis that homicide rates influence certainty and severity of punishment remains an open question which should be examined carefully in future research. Our argument is that, for our present purposes, it is just as reasonable to derive estimates assuming no simultaneous relationships between homicide rates and repressive factors as it is to derive simultaneous estimates under the extremely restrictive assumptions that would be necessary for such an analysis. Without pursuing a detailed critique of simultaneous estimations that have been made in deterrence studies, it should be noted that many of the variables that have been treated as instrumental variables and thus excluded from crime functions and included in punishment functions seem very arbitrary and inconsistent with existing theory. For example, it is not at all clear why such variables as region, percentage of the population that is nonwhite, age, and government expenditures should be excluded from crime functions, yet this has been the practice in existing studies that have made simultaneous estimates.[5]

When one adds to these difficulties the fact that we are deriving our estimates with a relatively small number of observations (N = 48) and that simultaneous estimation procedures such as two-staged least

[4]The data are based on 2094 cities with a toal population of 77,695,412 (Uniform Crime Reports 1959: Table 13, p. 83). The comparable figures for 1960 and 1961 are 92.3 and 93.1 for the clearance rate; the standard deviations are 5.9 and 4.1.

[5]See Nagin (1975) for a critical discussion.

squares do not provide substantial improvements over direct ordinary least square estimates unless sample size is large (Namboodiri et al. 1975:517), there seems to be no strong reason for expecting that simultaneous estimates will be superior to direct estimates.

Table 3 presents our estimate for two models which contain only Gibbs' repressive factors. In both the linear and the multiplicative forms the estimates of the effects of certainty and severity of punishment are negative and greater than twice their standard errors.[6] The *d*-statistic in Table 3 provides a means of comparing the goodness of fit of the two models in terms of the residual sums of squares.[7] The multiplicative form clearly fits better than the linear form; the residual sum of squares for the transformed linear model is almost twice as large as the multiplicative model, and the *d*-statistic is significant at well beyond the 0.01 level. All of this is consistent with previous analyses of the data and is presented here as a basis of comparison with the estimate derived from the more complete forms of the model.

Estimates from linear and multiplicative forms of a model which includes the six components of the index used by Loftin and Hill along with Gibbs' repressive factors are presented in Table 4. In the linear form of the model the adjusted coefficient of determination[8] increases from 0.20 to 0.83 when the index variables are added to the model; in the multiplicative form the increase is from 0.37 to 0.78. More important, however, for present purposes is the reduction in the magnitude of the estimates of the punishment effects in both forms of the model. In the linear specification the estimated certainty effect falls from -0.070 to -0.019 and severity falls from -0.063 to -0.013. In both cases the estimated regression coefficients are considerably less than

Table 3. Ordinary least squares estimates of models which include only certainty and severity of punishment

A. Linear model $H = \beta_0 + \beta_1 P(I) + \beta_2 T + u$

	$\hat{\beta}$	$\hat{\beta}/S\hat{\beta}$
Certainty	−.07036	−2.19053
Severity	−.062	−2.83778
Constant	−14.93599	

$\overline{R^2} = .20144$

$\Sigma e_1{}^{*2} = 32.99791$

B. Multiplicative model $H = \beta_0 P(I)^{\beta_1} T^{\beta_2} e^u$

	$\hat{\beta}$	$\hat{\beta}/S\hat{\beta}$
ln Certainty	−.85183	−2.95057
ln Severity	−1.26859	−4.66462
Constant	9.41954	

$\overline{R^2} = .36509$

$\Sigma e_2{}^{*2} = 18.91686$

C. Comparison on linear and multiplicative models

d = 13.35339

Table 4. Ordinary least squares estimates of models which include certainty and severity of punishment along with components of the Loftin-Hill Index

A. Linear model $H = \beta_0 + \beta_1 P(I) + \beta_2 T + \beta_3 \text{ IMR} + \ldots \beta_8 \text{ ONEPARENT} + u$

	$\hat{\beta}$	$\hat{\beta}/S\hat{\beta}$
Certainty	−.01906	−1.112
Severity	−.01340	−1.085
IMR	.31587	3.384
LOWED	.84948	2.790
ILLIT	−2.34584	−2.812
LOWINC	−.18781	−1.380
TESTFAIL	−.00557	−.095
ONEPARENT	.49164	2.153
(Constant)	−6.84304	

$\overline{R^2} = .83248$

$\Sigma e_1{}^{*2} = 6.01917$

B. Multiplicative model $H = \beta_0 P(I)^{\beta_1} T^{\beta_2} \text{IMR}^{\beta_3} \ldots \text{ONEPARENT}^{\beta_8} e^u$

	$\hat{\beta}$	$\hat{\beta}/S\hat{\beta}$
Ln Certainty	−.25591	−1.360
Ln Severity	−.53973	−2.725
Ln IMR	1.39258	2.078
Ln LOWED	1.46564	1.932
Ln ILLIT	−1.34457	−2.311
Ln LOWINC	−.16483	−.677
Ln TESTFAIL	.05066	.190
Ln ONEPARENT	1.37484	2.921
(Constant)	−5.02170	

$\overline{R^2} = .78210$

$\Sigma e_2{}^{*2} = 5.64538$

C. Comparison of linear and multiplicative models

d = 1.53868

twice their standard errors. In the multiplicative form the estimated effect of certainty falls from -0.852 to -0.256 and the estimated effect of severity from -1.269 to -0.540; only the severity estimate is more than twice its standard error. Since the two forms of the model lead to different conclusions about the importance of the severity effect, we have examined the goodness of fit for the two forms in some detail. The *d*-statistic indicates that although the multiplicative form has a smaller residual sum of squares (5.645 as opposed to 6.019 for the linear specification), the difference is not significant at the 0.05 level. Moreover, a plot of the residuals against the expected normal order statistics shows that there is one extremely deviant value for the multiplicative specification (the case is Rhode Island) and there is a noticeable

[6]The multiplicative model is estimated by expressing each variable in terms of its natual logarithm. In this form the multiplicative model can be expressed as $\ln H = \ln\beta_0 + \beta_1 \ln P(I) + \beta_2 \ln T + u$. Subsequent multiplicative models are estimated in the same way.

[7]To compute *d* the homicide rate was multiplied by a constant, *C*, which is defined as

$$C = \exp\left(\frac{-\Sigma \ln H}{N}\right)$$

Then both models were estimated using the transformed homicide rate. The residual sums of squares are used to derive *d* as follows:

$$d = \frac{N}{2}\left|\ln\frac{\Sigma e_1^{*2}}{\Sigma e_2^{*2}}\right|$$

The error sum of squares for the linear model is designated Σe_1^{*2} and Σe_2^{*2} for the multiplicative model. See Rao and Miller (1971:107-111) for a discussion of the procedure.

[8]The adjusted coefficient is defined as

$$\bar{R}^2 = 1 - (1 R^2)\frac{(N-1)}{(N-k)}$$

where *N* is the total observations and *k* is the number of parameters estimated, including the constant term.

deviation from a linear relationship near the lower end of the plot. Since there is no compelling theoretical reason for selecting the multiplicative specification and since there is evidence that the linear model fits the data at least as well as if not better than the logarithmic model, we are inclined to select the linear form as the better specification. However, as the next segment of our analysis will demonstrate, the choice between the linear and logarithmic specifications becomes less important in assessing the importance of the punishment variables when the other variables in the homicide model are introduced.

In the analysis described in Table 5 we add the remaining six variables in the homicide model. With the exception of the estimation of the certainty effect in the linear specification, which remains essentially the same, the estimated effects of the punishment variables are reduced still further, and in no case are the estimates of punishment effects greater than twice their standard errors. Moreover, in the linear specification the estimate of the effect of severity of punishment has changed its sign from negative to positive.

Since both forms of the model lead to the same general conclusions about the significance of the punishment variables, the selection between the linear and the multiplicative forms is less important than it appeared to be in the earlier specification (Table 4) where the form of the model made a difference in the significance of the severity estimate. The multiplicative form seems to fit a little better than the linear form but the *d*-statistic (2.467) is consistent with the conclusion that they were drawn from the same population.

It should also be noted that the reduction of the punishment effects to an insignificant size does not depend on entering all 12 of the *etiological* factors. We have already shown that in the linear specification the estimated repressive effects are less than twice their standard errors after the six components of the index are introduced. In a step-wise analysis of the multiplicative specification, percent nonwhite entered on the first step after the punishment variables and the six components of the index had been entered. At that point the estimated severity effect is less than twice its standard error and remains there until all the variables in the model have entered (see Table 6).

Table 5. Ordinary least squares estimates of models which include certainty and severity of punishment and all of the *Etiological* variables

A. Linear Model $H = \beta_0 + \beta_1 P(I) + \beta_2 T + \beta_3 IMR + \ldots \beta_{14} GINI + u$

	$\hat{\beta}$	$\hat{\beta}/s\hat{\beta}$
Certainty	−.02276	−1.556
Severity	.00397	.349
IMR	.18016	1.789
LOWED	.34358	1.151
ILLIT	−1.60651	−1.921
LOWINC	−.05938	−.319
TESTFAIL	.03874	.631
ONEPARENT	.30235	1.495
NONWHITE	.03865	.694
RURAL	−.01734	−.568
AGE	.60985	2.561
HOSPBEDS	−.00203	−1.811
REGION	−1.85923	−2.182
GINI	.01676	.960
(Constant)	−14.93599	

$\overline{R}^2 = .89166$

$\Sigma e_1{*}^2 = 3.30905$

B. Multiplicative Model $H = \beta_0 P(I)^{\beta_1} T^{\beta_2} IMR^{\beta_3} \ldots GINI^{\beta_{14}} e^u$

	$\hat{\beta}$	$\hat{\beta}/s\hat{\beta}$
Ln Certainty	−.18257	−1.126
Ln Severity	−.00721	−.032
Ln IMR	.73758	1.206
Ln LOWED	−.06035	−.077
Ln ILLIT	.48048	−.798
Ln LOWINC	−.05027	−.134
Ln TESTFAIL	.17835	.787
Ln ONEPARENT	1.12196	2.757
Ln NONWHITE	.31237	3.532
Ln RURAL	.17588	.810
Ln AGE	2.31579	2.179
Ln HOSPBEDS	−.38872	−1.404
REGION	−.03707	−.195
Ln GINI	1.20806	.492
(Constant)	−15.37168	

$\overline{R}^2 = .86442$

$\Sigma e_2{*}^2 = 2.98574$

C. Comparison of Linear and Multiplicative Models

d = 2.46752

Note: REGION was coded arbitrarily as a binary variable so it was not expressed as a logarithm in the multiplicative form of the model. It remained in its original form.

Table 6. Ordinary least squares estimates of multiplicative model which includes the certainty and severity of punishment, the components of the Loftin-Hill Index, and percent of population nonwhite

	$\hat{\beta}$	$\hat{\beta}/s\hat{\beta}$
Ln Certainty	−.13560	.857
Ln Severity	−.18916	1.035
Ln IMR	1.32200	2.381
Ln LOWED	.04173	.055
Ln ILLIT	−.54782	1.063
Ln LOWINC	.07524	.361
Ln TESTFAIL	.09962	.453
Ln ONEPARENT	1.24587	3.187
Ln NONWHITE	.33250	4.344
(Constant)	−5.20431	

$\overline{R}^2 = .85067$

The 12 *etiological* variables in the model also generally have small estimated effects; only two estimates are greater than twice their standard error in the linear specification (percent age 20 to 34, and region); for the multiplicative specification three meet this criterion (percent of children living with one parent, percent of population living in rural territory, and percent age 20 to 34). However, this is to be expected given the fact that the 12 *etiological* variables, to a very large degree, reflect a single socioeconomic dimension and thus represent redundant indexes which divide up the socioeconomic variance among themselves

Table 7. Coefficients of determination (R^2) for regression of each independent variable on the set of other independent variables

Variable taken as dependent	Coefficient of determination R^2
IMR	.85667
LOWED	.98837
ILLIT	.98011
LOWINC	.94671
TESTFAIL	.95816
ONEPARENT	.89702
NONWHITE	.91490
RURAL	.79563
AGE	.60534
HOSPBEDS	.49285
REGION	.79129
GINI	.88510
CERTAINTY	.34090
SEVERITY	.48702

resulting in small estimates of independent effects. The *etiological* factors are included to purge the punishment variables of their dependence on *etiological* factors so that their independent effects could be more clearly evaluated; thus the size of individual *etiological* effects is not important for present purposes.

The pattern of high multicollinearity among the 12 *etiological* factors can be seen in Table 7 where the coefficient of determination has been calculated from the regression of each independent variable on the set of other (k-1) independent variables. Note that while the 12 *etiological* variables are very highly intercorrelated with each other (the average R^2 is 0.84), the *repressive* factors are not highly correlated with each other (r^2 = 0.00052) or with the other independent variables (R^2 = 0.341 for certainty and 0.487 for severity). This indicates that the relatively large standard errors of the estimates of the *repressive* variables are not a result of multicollinearity and provides support for the inference that the punishment variables do not have independent effects on the homicide rate.

Conclusions

Our purpose in this paper has been to show that when the effects of sociocultural factors are removed from Gibbs' measures of certainty and severity of punishment, the estimates of their independent effects on homicide rates are reduced to the point that they might reasonably be considered to be nonexistent. Such a demonstration is very important because previous, widely cited analyses of the same data have concluded that they provide evidence for independent deterrent effects. One should not conclude, however, that the data provide very strong evidence for either position. The analysis must remain inconclusive because of serious theoretical and methodological limitations on the research design. Gibbs himself, after a recent, thorough review of deterrence studies, concluded that "... horrendous problems preclude a categorical rejection or acceptance of the deterrence doctrine." (Gibbs 1975:1) We do not attempt to discuss these problems here. Excellent discussions are available elsewhere (Gibbs 1975, Nagin 1975, Greenberg 1975). Our objective has simply been to demonstrate that the application of reasonable statistical estimation procedures to a model that adequately controls for sociocultural variables can provide evidence that is quite different from that provided by previous analyses of Gibbs' data.

Acknowledgments

The original work on this paper began in the Working Group on Deterrence Research at the Workshop in Criminal Justice Statistics which consisted of David Britt, Kinley Larntz, Colin Loftin, Dan Nagin, and David Seidman. At various points other members of the workshop made important contributions to my thinking. The original calculations for the research were done in Washington during the workshop, but all calculations have been repeated with the original Loftin-Hill data rather than the data that were hastily collected during the workshop. I spent two days visiting with Kinley Larntz and Stephen Fienberg at the Department of Applied Statistics of the University of Minnesota. That experience had a major influence on the paper as well as on my general education. Michael Timberlake and Robert Parker, students in the Department of Sociology at Brown University, assisted with some of the data analysis presented in the paper. I am responsible for the draft of the report and any errors that it contains.

References

American Education (1966)
2: no. 9. Washington: U.S. Office of Education.

Antunes, George E., and A. Lee Hunt (1973)
"The impact of certainty and severity of punishment on levels of crime in American states: An extended analysis." *Journal of Criminal Law and Criminology* 64:486–493.

Bean, Frank, and R. Cushing (1971)
"Criminal homicide, punishment, and deterrence: Methodological and substantive reconsiderations." *Social Science Quarterly* 52: 277–289.

Bowers, William J., and Glenn L. Pierce (1975)
"Deterrence, brutalization, or nonsense: A critique of Isaac Ehrlich's research on capital punishment." Unpublished manuscript.

Brearly, H. C. (1932)
Homicide in the United States. Chapel Hill: University of North Carolina Press.

Chiricos, Theodore G., and Gordon P. Waldo (1970)
"Punishment and crime: An examination of some empirical evidence." *Social Problems* 18:200–217.

Coser, Lewis A. (1967)
"Violence and social structure." In *Continuities in the study of social conflict* (Lewis Coser, ed.), New York: Free Press, pp. 55–71.

Dye, Thomas R. (1969)
"Income inequality and American state politics." *American Political Science Review* 63:157–162.

Ehrlich, Isaac (1973)
"Participation in illegitimate activities: A theoretical and empirical investigation." *Journal of Political Economy* 81:521–565.

Ehrlich, Isaac (1975)
"The deterrent effects of capital punishment: A question of life and death." *American Economic Review* 65:397–417.

Gastil, Raymond D. (1971)
"Homicide and a regional culture of violence." *American Sociological Review* 36:412–427.

Gibbs, Jack (1968)
"Crime, punishment and deterrence." *Southwestern Social Science Quarterly* 48:515–530.

Gibbs, Jack (1975)
Crime, punishment, and deterrence. New York: Elsevier.

Greenberg, David F. (1975)
"Theft, rationality and the methodology of deterrence research." Paper presented to the Annual Meetings of the American Sociological Association, August 1975.

Greene, Jon S., and Alvin Renetzky (eds.) (1970)
Standard education almanac. Los Angeles: Academic Media.

Grove, Robert D., and Alice M. Hetzel (1968)
Vital statistics rates in the U.S. 1940–1960. U.S. Department of Health, Education, and Welfare. Washington, D.C.: U.S. Government Printing Office.

Hackney, Sheldon (1969)
"Southern violence." *American Historical Review* 74:906–925.

Henry, Andrew F., and James F. Short (1954)
Suicide and homicide: Some economic, sociological, and psychological aspects of aggression. New York: Free Press.

Loftin, Colin, and Robert H. Hill (1974)
"Regional subculture and homicide: An examination of the Gastil-Hackney thesis." *American Sociological Review* 39:714–724.

Logan, Charles H. (1970)
"On punishment and crime (Chiricos and Waldo, 1970): Some methodological commentary." *Social Problems* 19:280–284.

Logan, Charles H. (1972)
"General deterrence effects of imprisonment." *Social Forces* 51:64–73.

Logan, Charles H. (1975)
"Arrest rates and deterrence." *Social Science Quarterly* 56:376–389.

Nagin, D. (1975)
"An investigation into the association between crime and sanctions." Unpublished Ph.D. dissertation, Carnegie-Mellon University.

Namboodiri, N. K., L. F. Carter, and H. M. Blalock, Jr. (1975)
Applied multivariate analysis and experimental designs. New York: McGraw-Hill.

Orsagh, Thomas (1973)
"Crime, sanctions, and scientific explanation." *Journal of Criminal Law and Criminology* 64:354–361.

Phillips, Llad, and Harold L. Votey, Jr. (1972)
"An economic analysis of the deterrent effect of law enforcement on criminal activity." *Journal of Criminal Law, Criminology, and Police Science* 63:330–342.

Rao, Potluri, and Robert LeRoy Miller (1971)
Applied econometrics. Belmont, Calif.: Wadsworth.

Tittle, Charles R. (1969)
"Crime rates and legal sanctions." *Social Problems* 16:409–423.

Tittle, Charles R. (1964)
"Certainty of arrest and crime rates: A further test of the deterrence hypothesis." *Social Forces* 52:455–462.

Tittle, Charles R. (1975)
"Deterrence or labeling?" *Social Forces* 53:399–410.

Tittle, Charles R., and Charles Logan (1973)
"Sanctions and deviance: Evidence and remaining questions." *Law and Society Review* 7:371–393.

Tullock, Gordon (1969)
"Does punishment deter crime?" *The Public Interest* 36:103–111.

U.S. Department of Justice (1959)
Uniform Crime Reports 1959. Washington, D.C.: U.S. Government Printing Office.

U.S. Bureau of the Census (1962)
Statistical abstract of the United States. Washington, D.C.: U.S. Government Printing Office.

U.S. Bureau of the Census (1964a)
U.S. Census of population: 1960. Vol. 1, part 1. Washington, D.C.: U.S. Government Printing Office.

U.S. Bureau of the Census (1964b)
Statistical abstract of the United States. Washington, D.C.: U.S. Government Printing Office.

Wilkins, Leslie T. (1974)
"Current aspects of penology: Directions for corrections." *Proceedings of the American Philosophical Society* 118:235–247.

Wolfgang, Marvin (1958)
Patterns in criminal homicide. Philadelphia: University of Pennsylvania Press.

Wolfgang, Marvin (1968)
"Homicide." *International Encyclopedia of the Social Sciences.* New York: Macmillan.

Wolfgang, Marvin, and Franco Ferracuti (1967)
The subculture of violence: Toward an integrated theory in criminology. London: Tavistock Publications.

Recent econometric modelling of crime and punishment: Support for the deterrence hypothesis?

STEPHEN S. BRIER
Department of Statistics
University of Iowa

STEPHEN E. FIENBERG
Department of Statistics
and Deparment of Social Science
Carnegie-Mellon University

In this paper we review some recent attempts to develop econometric models for assessing the deterrent effect of punishment on crime, as well as analyses carried out to validate these models. The formulation of the basic econometric model considered here is due to Becker (1968), and the detailed specification of the model, along with much of the empirical work reviewed, has been carried out by Isaac Ehrlich. We find serious flaws with the Becker-Ehrlich model, with the data used in its empirical implementation, and with Ehrlich's conclusions regarding evidence to support the deterrent effect of punishment on crime. Indeed, we can find no reliable empirical support in the existing economics literature either for or against the deterrence hypothesis.

Introduction

The threat to person and property from crime is a central problem of social concern in the United States today. But crime is not a new problem. "Crime in the streets" and "the need for law-and-order" have been slogans used by politicians at least since the early 1960's to exploit popular fears of crime, and the study of crime as a social phenomenon has played a prominent role in American sociology throughout the twentieth century. With the surge in reported crime in the 1960's, considerable public attention was focused on the control of crime through President Lyndon B. Johnson's Commission on Law Enforcement and Administration of Justice and the subsequent implementation of the Commission's recommendations. At the same time economists began to study law enforcement as an economic problem involving the allocation of scarce resources (e.g., see Becker 1968).

The punishment of criminal offenders has played an important role in both the economic and sociological approaches to the control of crime. Although retribution was long considered to be a primary rationale for punishment, this view has given way, at least partially, to the position that punishment of offenders will lead to a subsequent reduction in or prevention of crime. The possible mechanisms for the prevention of crime through punishment, other than retribution, are several:

(1) *Incapacitation.* By removing an offender from society for some period of time, we can prevent that offender from repeating his offense or committing other ones *while* he is not in contact with society.

(2) *Education.* By administering punishment for antisocial acts (such as crimes), government and its agencies achieve a moral and educational effect on individuals in general. Society thus instructs that these crimes are counter to its norms, and by doing so educates individuals and influences their behavior.

(3) *Rehabilitation.* Through the use of education, correctional "treatment," or vocational training during imprisonment, society attempts to "change" offenders so that they will not commit crimes in the future.

(4) *Deterrence.* By enforcing punishment on an offender, society warns the offender, and the community at large, and thus inhibits the offender or others in the community from engaging in criminal activity in the future.

When it is the sanctioned offender who is inhibited from committing crimes by the actual experience of punishment we speak of *specific* deterrence; when individuals other than the sanctioned offender are so inhibited we speak of *general* deterrence.

With regard to the deterrent effect of punishment we need to make a further distinction, suggested by Zimring and Hawkins (1973), between *absolute* and *marginal* deterrence:

> The problem of absolute deterrence relates to the question, does this particular criminal sanction deter? The problem of marginal deterrence relates to such questions as, would a more severe penalty attached to this criminal prohibition more effectively deter? In the capital punishment debate the issue is not that of absolute deterrence—whether the death penalty is a deterrent. It is that of marginal deterrence—whether it is a more effective deterrent than the alternative sanction of long imprisonment.[1]

Most people can accept the notion of the absolute deterrence effect of many forms of punishment. For example, the removal of all sanctions for a crime such as robbery, most would agree, would lead to an increase in the rate at which robberies are committed; others might argue that the reason for the increase is not the removal or threat of punishment but rather the removal of the educational and moral value the punishment provided for society as a whole. The issue which is explored in this paper and which has been the subject of recent concern is that of marginal deterrence, i.e., we are interested in the potential effects of shifts in the level of sanctions on subsequent criminal behavior.

Measuring the effects of these different mechanisms through which punishment might prevent crime is a far more difficult task than it may at first seem. As we just noted, it is often difficult to separate the educational and general deterrent components of the effects of punishment. Similarly, it is difficult to separate out the effects of incapacitation from those of deterrence. The difficulty is in fact one of explaining

[1] Zimring and Hawkins (1973:14).

exactly what the mechanism of incapacitation actually entails.

A simple-minded argument to explain the effect of incapacitation might go as follows. Suppose we imprison a 20-year-old male for a period of 2 years for committing the crime of burglary. Then, if that offender had been committing burglaries at the rate of, say, 12 per year, the simple fact of his removal from society for 2 years will prevent the commmission of 24 burglaries. Then, in any modelling we do, we would wish to remove these 24 burglaries from the overall effect of the 2 years of punishment on the future burglary rate.

Even if we knew the rate of offending for each imprisoned offender, the calculation in this argument may completely distort the actual effects on the overall crime rate of incapacitating various individuals. In the case of the 20-year-old burglar, we need to know more about the nature and circumcumstances of his offenses. If he were a member of a group of burglars, all of whom work together, then his imprisonment may not have any effect at all on the number of burglaries committed by the group. Indeed, if he was arrested and convicted because he was the worst burglar in the group, the other members of the group might well recruit a superior replacement or simply increase their criminal activities, and thus the overall number of burglaries the group commits could easily increase. Cook (1977) discusses this replacement effect in an economic context in terms of its relationship to the quality of crime opportunities. Reiss (1980) has argued persuasively that the interpenetrating social networks of offenders and their victims need to be better understood before we can make any serious attempt to separate out the effects of incapacitation from those of deterrence. At the same time no empirical model for measuring the deterrent effect of punishment can be complete unless it also includes the related incapacitation effect in some form. (See also the discussion of incapacitation in the report of the National Academy of Sciences Panel on Research on Deterrent and Incapacitative Effects (Blumstein et al. 1978).)

Let us explore the mechanism of deterrence somewhat further. The effects of deterrence can be thought of only in the context of those who are likely to commit a crime and who are influenced by the threat of punishment for that crime if apprehended and convicted. The nature and extent of punishment can have no effect on those who do not need to be deterred from criminal activities and on those who are not or cannot be deterred by the threat of punishment. Thus, to measure the deterrent effect of punishment, we need to attempt to get at those on the margin of criminal activity. The size of the group of individuals on the margin limits the potential deterrent effects of any form of punishment. Since economic theory so often deals with the marginal effects of various policy changes, the appeal of crime and punishment as an application of economic modelling is great. What remains to be seen is whether such economic models are consistent with empirical observations.

In this paper we review some recent attempts to develop econometric models for assessing the deterrent effect of punishment on crime, as well as analyses carried out to validate these models. The first section discusses the possible types of empirical investigations that might be used to study deterrence. Next we briefly describe the Becker-Ehrlich econometric model for crime and punishment. Then we describe some of the general problems in the empirical implementation of this model. The next two sections outline the published empirical tests of the model along with some of our own reanalyses. One deals with cross-sectional data for 1960 and 1970, while the other deals primarily with longitudinal data for homicide and includes a special look at the deterrent effects of capital punishment.

Finally our own conclusions regarding the empirical evidence on deterrence are presented. They can be summarized briefly as follows:

(1) The Becker-Ehrlich model has glaring shortcomings and, when examined critically, does not lead to the claimed testable hypotheses regarding the effect of punishment of crime.

(2) The use of the Becker-Ehrlich model for aggregate data requires extensive justification that has never been given.

(3) The crime and imprisonment data used to empirically examine the Becker-Ehrlich model are so untrustworthy as to render any serious analysis meaningless.

(4) The empirical implementations of the Becker-Ehrlich model are badly flawed and have extremely grave statistical shortcomings and most published conclusions from them are not to be trusted.

(5) Even if one accepts the Becker-Ehrlich model, and Ehrlich's choice of data to implement it (which we do not), Ehrlich's affirmative conclusions regarding the deterrent effect of punishment on crime in general, and of capital punishment on murder in particular, do not stand up to careful statistical scrutiny.

Thus our conclusions are in agreement with those of the National Academy of Sciences Panel on Research on Deterrent and Incapacitative Effects: There is no empirical evidence to warrant an affirmative conclusion regarding the deterrent effect of punishment in general, and the available correlational or observational studies on homicide provide no useful evidence on the deterrent effect of capital punishment. The only available evidence that give support for the notion of general deterrence comes from "natural experiments" such as Ross's (1973) study of the British Road Safety Act (see the discussion in Cook 1977). Moreover, given the limitations of aggregate data on crime and incarceration, we believe that much more attention in the future should be focused on studies of individual criminal behavior. We do not expect any empirical research, at least in the near future, to provide definitive evidence on the deterrent effects of capital punishment.

Possible vehicles for studying deterrence

As with many other social phenomena, the relationship between punishment and crime can, in principle, be explored either in an experimental setting or by an observational study.

In a randomized controlled field trial, experimental units—individuals, collections of individuals, political or legal jurisdictions, etc.—are randomly assigned to treatment groups and are then carefully followed to assess the actual effects of the treatment. The randomization allows the experimenter to avoid the dangers of self-selection, and it allows for the control of variables not included directly into the design of the field trial. The "controlled" nature of such field trials implies that the choice of treatment for an experimental unit is that of the investigator and that at least two treatments (or levels of treatment) are being compared (see Gilbert, Light, and Mosteller 1975) for further discussion). The randomized controlled field trial is the most demanding of all research strategies for investigating social innovations, but the increased reliability gained from a randomized trial can often far outstrip the costs. For a field trial involving deterrence the treatments would involve different levels of punishment and/or the threat of punishment. Unfortunately, there have been few examples of randomized controlled field trials dealing with deterrence. Zimring and Hawkins (1973) note that "[i]t is difficult to conceive of an acceptable experiment in which, after random assignment, the severity of sanctions threatened for a violation of a particular criminal law was varied between the two groups." While we do not completely agree with their statement, we do recognize the many legal and possibly moral roadblocks to careful experimentation on deterrence. Because of this difficulty in mounting experiments, most investigators resort to alternative research strategies, two of which are highly prominent.

In the first approach the researcher attempts to assess the effect of a change in the level of sanctions by comparing reported crime rates before and after the change in a given jurisdiction. To reduce the potential biases and errors of such an approach, investigators often compare the change in rates with those in other jurisdictions where no changes in the level of sanctions took place. For a recent discussion of such "natural" experiments with regard to the deterrent effect of capital punishment, see Baldus and Cole (1975); with regard to other deterrent effects, see Cook (1977) and Zimring (1978).

The other approach to the measurement of deterrent effects of punishment is the gathering of aggregate data on crime and punishment as well as on various social and economic variables. The researcher then studies the variations in crime that occur either among jurisdictions or over time. This is the approach adopted by Ehrlich (1973, 1975a, 1977b) in his attempt to develop an econometric model for the effect of varying sanctions, and the inherent difficulties in this approach are the primary subject of this paper.

The Becker-Ehrlich model for crime and punishment

Becker (1968) introduces his attempt to model crime and its optimal control by noting that "'crime' is an economically important activity, or 'industry', not withstanding the almost total neglect by economists." His model involves basically five behavioral relationships which he claims underlie the costs of crime—the relations between:

(1) The number of offenses and their cost.

(2) The number of offenses and the punishments meted out.

(3) The number of offenses, arrests, and convictions and the public expenditures on police and courts.

(4) The number of convictions and the cost of imprisonments or alternative punishments.

(5) The number of offenses and the private expenditures on protection and apprehension.

We do not discuss (5) any further here since it plays no role in Ehrlich's empirical implementation of Becker's model nor in the bulk of the related literature we consider.

Becker's basic premise is that criminals maximize their expected gains (according to some utility function) from illicit activity. A person commits an offense if the expected utility he will receive exceeds the utility he would receive by engaging in other activities. Thus the criminal's decision is based on benefits and costs of both a monetary and psychic nature. According to Ehrlich (1973) an individual allocates a fixed amount of time among legal and illegal income-generating activities. The effects of this time allocation are introduced only implicitly by Ehrlich through the effects of time allocation on wealth. From the basic model Ehrlich goes on to derive some behavioral implications, such as: An increase in the probability of apprehension and punishment, with no change in other variables, reduces the incentive to participate in illegitimate activities.

Block and Heineke (1975) have examined Ehrlich's model with great care and have shown that if the allocation of time is introduced explicitly into the utility analysis, the behavioral implications of the model derived by Ehrlich do not necessarily hold. Moreover, they note that, contrary to the assumptions of Becker and Ehrlich, it is not necessarily true that monetary equivalents to labor and penalty attributes of an offense exist. As a result, propositions regarding the deterrent effect of punishment become empirical questions rather than theoretical consequences subject to empirical validation.

Implicit in the utility analysis of Becker, Ehrlich, and others is the assumption that criminals or potential criminals are rational decisionmakers in that they make their decisions according to a list of axioms. The appropriateness of such an economic model of crime is clearly open to question. For example, Avio and Clark (1976) state that: "It would be difficult to argue that perpetrators of violent crimes behave according to the usual set(s) of axioms. . . . Murders involving some form of premeditation and motivated by economic gain might be consistent with the economic model. Most murders, however, occur in the home, involve members of the same family, and seem unpremeditated."

The net result of this econometric modelling is a functional relationship between the number of offenses committed by individual j, O_j, and—

(a) his probability of conviction, p_j,

(b) his punishment given conviction, f_j,

(c) his rate of return (benefits) if he successfully commits the crime, w_{ij} (i stands for illegal returns),

(d) the rate of returns from alternative legal activities, w_{lj} (l stands for legal returns),

(e) the probability of legal unemployment, u_{lj}, and

(f) a vector of other variables, $\mathbf{v}_j$:

$$O_j = O_j(p_j, f_j, w_{ij}, w_{lj}, u_{lj}, v_j). \tag{1}$$

The function (1) is the one discussed by Ehrlich (1973), and is simply an elaboration of Becker's supply-of-offenses function. To arrive at (1), one must adopt many untested and possibly untestable assumptions regarding criminal behavior and its determinants.

There is little or no discussion by Becker or Ehrlich regarding those "economic" variables not included in (1). Nor do they provide support for the forms of the variables appropriate for inclusion for the empirical validation of the model. This is a serious matter. No amount of utility theory, systems of partial differential equations, Kuhn-Tucker first-order optimality conditions, and analogies to the supply and demand for bread and butter (see Ehrlich and Gibbons 1977) can make up for the logical leaps that lead to the specification of (1).

The first step adopted by Ehrlich and others in making the model of expression (1) suitable for empirical examination is the aggregation of data across individuals. Thus Ehrlich uses the aggregate function

$$O^* = O^*(p^*, f^*, w_i^*, w_l^*, u_l^*, v^*) \tag{2}$$

where O^* is the aggregate number of offenses in a particular jurisdiction, p^* is the aggregate probability of conviction, and so on. The justification for such aggregation typically rests on the assumption that either the parameters used to specify the relation (1) are constant across individuals or the parameters are stochastic, coming from some *common* distribution. Actually, to justify the aggregation, one needs further to specify a specific functional form. For example, suppose log O_i is linearly related to the logarithms of the variables on the right-hand side of (1). Then, if the coefficients are the same across individuals, the aggregate number of offenses, O^*, is related to the geometric means of the individual values for the other variables. Thus the justification of functional form must ultimately be established at the individual rather than the aggregate level. Since aggregation also takes place over time, we need to assume some form of constancy or, at a minimum, stochastic stationarity of the parameters in the functional specification (see e.g., Kuh and Welsch 1976).

One of the few empirical examinations of a related functional specification at the individual level was carried out by Witte (1977). Her analyses show the importance of the specification of individual sociodemographic variables, such as race and age, and are supportive of arguments indicating the inappropriateness of the aggregate data used by most investigators in this area.

Alternatives to the Becker-Ehrlich model exist. These are based on concepts such as the saturation of the resources of the criminal justice system (Cook 1977), and "homeostasis," the apparent stability of imprisonment, and effective prison capacity (Blumstein and Cohen 1973; Nagin 1977). While we do not directly discuss the empirical examination of these alternatives, we note that they cast doubt on Ehrlich's claims regarding marginal deterrent effects of punishment.

Moving toward empirical examination of the Becker-Ehrlich model

To move from the aggregate supply-of-offenses model of expression (2) to an empirical study of deterrence, one needs to

(a) Specify in detail the "other" variables to be included as part of the vector v^*.

(b) Describe how the variables are to be measured.

(c) Specify the actual functional form of the relationship.

Ehrlich (1973, 1975a) provides *a priori* specifications for (a) and (c) which we call into question in our reanalysis of his data. Thus we defer our discussion of these matters to later sections of this paper. In this section we discuss (b).

The actual rate of offending in any community for any specific crime during a specific time period is not known. What we have available are data on offenses reported to the police. As a substitute for O^* (the actual offense rate), Ehrlich (1973, 1975a) uses Q/N, where Q is the number of offenses reported to the police *and* recorded by them for a jurisdiction, and N is an estimate of the population size for that jurisdiction. It is well known that not all offenses committed are reported to the police. Estimates of the ratio of Q to Q^*, the true number of offenses, vary from as low as 10 percent for rape to close to 100 percent for murder. In the case of automobile theft there have been occasional reports that Q exceeds Q^* because of nonthefts reported to the police for insurance purposes.

Unfortunately, the ratio of Q/Q^* can vary dramatically from jurisdiction to jurisdiction. For example, Skogan (1976) notes that in the 26 city victimization surveys conducted under the auspices of the National Crime Survey Q^{**}/Q^* (where Q^{**} is the number of crimes reported to the police) for robbery varied from 52 percent to 76 percent.[2] The estimates of the ratio of the number of crimes appearing on police records (i.e., Q) to the number actually reported (i.e., Q^{**}) for these 26 cities, however, varied from 19 percent to 100 percent!

The quantity Q is also used by Ehrlich to get a measure of p^*, the aggregate subjective probability of punishment (i.e., apprehension and imprisonment). He uses C/Q as an estimate of p^* where Q is the number of recorded crimes in a given period of time and C is the number of offenders imprisoned during the same time period for the same jurisdiction. There are three problems here.

First, C and Q involve different units. C deals with offenders, and Q deals with offenses. There is not a one-to-one correspondence between the two.

Second, for a fixed period of time and a given jurisdiction, it is impossible to determine which of the offenses that are included in Q ultimately lead to apprehension and imprisonment. For example, juveniles are usually handled by separate juvenile justice systems and are rarely sent to prison. Individual offenders are not tracked over time through the criminal justice system, and the aggregate figures for those imprisoned that are available for jurisdictions such as states include individuals whose crimes may have been committed in other states. These aggregate figures also include offenders whose crimes took place possibly several years prior to incarceration. Thus the use of contemporaneous values of C and Q by state can lead to sizeable discrepancies between C/Q and p^*, which can be shown to vary from state to state in occasionally very strange ways. We discuss this problem further in the next section.

Third, Q now appears on both the left- and right-hand sides of the operational version of equation (2). As is noted in the report of the National Academy of Sciences (Blumstein et al. 1978), variation in the error in measuring Q^*, the true number of crimes committed, can induce a spurious negative relation between the offense rate and the punishment rate, when no such relationship actually exists. Klein, Forst, and Filatov (1978) demonstate how such errors can induce a similar bias in the estimated relation between murder and execution rates.

The source of the data used by Ehrlich and others to analyze the deterrent effects of punishment in the United States is the *Uniform Crime Reports* (UCR) produced by the Federal Bureau of Investigation (FBI). UCR data are collected as part of a *voluntary* reporting system involving state and local law enforcement agencies and are based on crimes recorded by these agencies. In addition to the problems of reporting and recording offenses noted earlier, UCR rates may seriously distort the level of offenses because multiple crimes with possibly multiple victims often are recorded as single offenses, and conversely single criminal events often involve multiple offenders, etc.

Since the beginning of the UCR program in 1933, both the number and the percentages of law enforcement agencies reporting to the FBI have increased dramatically. Moreover, reporting practices have also evolved, and officials generally agree that a larger proportion of offenses made known to the police now get reported to the FBI than was the case in the past. Finally, although the proportion of units reporting has increased over time, specific police departments that are once included in the UCR data base may not be included at later points in time.

[2]These rates are subject to substantial sampling error, and some of the variability can be attributed to this source.

Table 1. Variables used in Ehrlich's analysis[a]

$\left(\frac{Q}{N}\right)$	:	Crime rate (the number of offenses known per capita)
$\left(\frac{Q}{N}\right)_{j,t-1}$	:	Crime rate lagged one year
$\left(\frac{C}{Q}\right)_j$	:	Estimated probability of apprehension and imprisonment (the number of offenders imprisoned per offenses known)
T_j^*	:	Average time served by offenders in state prisons
W	:	Median income of families
X	:	Percentage of families below one-half of median income
NW	:	Percentage of non-whites in the population
A_{14}	:	Percentage of all males in the age group 14-24
V_{14}, V_{35}	:	Unemployment rate of civilian urban males aged 14-24 and 35-39, respectively
L_{14}	:	Labor-force participation rate for civilian urban males ages 14-24
Ed	:	Mean number of years of schooling of population 25 years old and over
SMSA	:	Percentage of population in standard metropolitan statistical areas
$\frac{E}{N}$, $\left(\frac{E}{N}\right)_{t-1}$	:	Per capita expenditure on police in fiscal 1960, 1959, respectively
M	:	Number of males per 100 females
D	:	Dummy variable distinguishing northern from southern states (south = 1)

[a] A subscript j denotes that the variable is indexed by specific crime categories.

The UCR data used by Ehrlich for his cross-sectional analyses (discussed in the next section) are based on the UCR for the years of interest. The data used in his national longitudinal analysis of murder (discussed later) are not the reported rates for each point in time. Rather, they consist of FBI estimates of what the reported crime rates would have been had the units included in the UCR system been the same as in 1972. Unfortunately there is no published description of the procedure used by the FBI to reestimate the rates. We have no way to assess the appropriateness of the FBI's reestimation procedure, but it is reasonable to conclude that it is likely to add further biases and increased variability to an already poor measure of crime.

We do not believe that UCR data collected prior to 1960 merit serious attention since they are almost completely unreliable; they are subject to enormous errors and biases. Any substantive conclusions one might draw from the analysis of these data are not to be trusted. Even the FBI no longer wishes to report data prior to 1960 because of doubts regarding their validity. We also have serious reservations regarding the quality and validity of the 1960 UCR data used by Ehrlich and others. Nonetheless, we discuss the 1960 data and national longitudinal UCR data for murder from 1933 through 1969 in the following two sections.

One alternative to the use of UCR data is the use of data from the special city surveys associated with the National Crime Survey (see Penick and Owens 1976). Cook (1977) and Wilson and Boland (1976) have used NCS city data to examine the effects of sanctions on the crimes of burglary and robbery, respectively. Cook (1977) in some simple preliminary analyses shows that the "usual" negative correlation between the UCR reported burglary rates and the corresponding UCR clearance rates is greatly diminished when city data from the NCS are used in their place. The Wilson and Boland (1976) study, which does find a significant negative effect using NCS data for robberies, unfortunately suffers from many of the methodological flaws discussed in the remaining sections of this paper. Given the serious technical problems with the NCS city data (see Penick and Owens 1976), we do not see the use of victimization data as a way of correcting the problems associated with UCR data.

The analyses of U.S. cross-sectional crime data

In this section we discuss some recent analyses that have attempted to empirically examine the Becker-Ehrlich econometric model of crime and punishment using cross-sectional data for the United States. We have attempted to replicate some of these results, and we report on our findings in this regard along with some additional analyses.

Ehrlich's 1960 data

Ehrlich (1973) analyzed data from 47 states. The states omitted from the analyses were Alaska, Hawaii, and New Jersey (see Vandaele 1978). New Jersey was omitted because certain key variables could not be obtained. The listing and definitions of all variables used in Ehrlich's study are given in Table 1.

To implement the aggregate model of expression (2), Ehrlich chose to use a Cobb-Douglas production function of the form

$$\frac{Q}{N} = a\left(\frac{C}{Q}\right)^{b_1} T^{*b_2} W^{b_3} X^{b_4} V_{14}^{b_5} \times L_{14}^{b_6} (NW)^{b_7} A_{14}^{b_8} \exp(e), \quad (3)$$

where Q/N is the operational attempt to estimate the offense rate, O^*.
C/N is the operational attempt to estimate the probability of punishment, p^*.
T^*, the average time served by prisoners in state prisons, is the empirical measure of f^*, the punishment given conviction.
W, the median family income, and X, the percentage of families below the median income, are used as replacements for the differential monetary returns of crime relative to legal alternatives, i.e., w_i^* and w_l^*.
V_{14}, the male urban unemployment rate for ages 14–24, and L_{14}, the labor force participation rate for this same group, are used in lieu of u_l^*, the probability of legal unemployment.
NW, the percentage of nonwhites, and A_{14}, the percentage of males in the 14–24 age group, are the "other variables," v^*.
a, and b_i for $i = 1, 2, \ldots, 8$, are parameters to be estimated.
e is a random error term.

The relationship among the variables in (3) is multiplicative, and we must take logarithms of both sides of the equation to produce linearity (we use natural logarithms, denoted by ln):

$$\ln\left(\frac{Q}{N}\right) = b_0 + b_1 \ln\left(\frac{C}{Q}\right) + b_2 \ln T^* + b_3 \ln W + b_4 \ln X + b_5 \ln V_{14} + b_6 \ln L_{14} + b_7 \ln NW + b_8 \ln A_{14} + e. \quad (4)$$

(Here we take $b_0 = \ln a$.) As a result of a variety of statistical analyses, Ehrlich concluded that V_{14}, L_{14}, and A_{14} had virtually no effect on the rest of the estimated equation and so he dropped them to yield the model:

Table 2. Estimated coefficients of selected variables in equation [5]. Estimates obtained by two-stage least squares

Offense	Intercept	$(C/Q)_j$	T_j^*	W	X	NW
Robbery	−11.030 (−1.804)[a]	−1.303 (−7.011)	−0.372 (−1.395)	1.689 (1.969)	1.279 (1.660)	0.334 (4.024)
Burglary	−2.121 (−0.582)	−0.724 (−6.003)	−1.127 (−4.799)	1.384 (2.839)	2.000 (4.689)	0.250 (4.579)
Larceny	−10.660 (−2.195)	−0.371 (−2.482)	−0.602 (−1.937)	2.229 (3.465)	1.792 (2.992)	0.142 (2.019)
Auto theft	−14.960 (−4.162)	−0.407 (−4.173)	−0.246 (−1.682)	2.608 (5.194)	2.057 (4.268)	0.102 (1.842)
Property crimes	−6.279 (−1.937)	−0.796 (−6.140)	−0.915 (−4.297)	1.883 (4.246)	2.132 (5.356)	0.243 (4.805)
Murder	0.316 (0.085)	−0.852 (−2.492)	−0.087 (−0.645)	0.175 (0.334)	1.109 (1.984)	0.534 (8.356)
Rape	−0.599 (−0.120)	−0.896 (−6.080)	−0.399 (−2.005)	0.409 (0.605)	0.459 (0.743)	0.072 (0.922)
Assault	−7.567 (−1.280)	−0.724 (−3.701)	−0.979 (−2.301)	1.650 (2.018)	1.707 (2.111)	0.465 (3.655)
Crimes against the person	1.635 (0.380)	−0.803 (−6.603)	−0.495 (−3.407)	0.328 (0.570)	0.587 (1.098)	0.376 (4.833)
All offenses	−1.388 (−0.368)	−0.991 (−5.898)	−1.123 (−4.483)	1.292 (2.609)	1.775 (4.183)	0.265 (5.069)

[a]Numbers in parentheses are t-ratios (i.e. the estimated coefficients divided by their estimated standard errors).

$$\ln\left(\frac{Q}{N}\right) = b_0 + b_1\ln\left(\frac{C}{Q}\right) + b_2\ln T^* + b_3\ln W + b_4\ln X + b_7\ln NW + e. \tag{5}$$

Ehrlich uses a simultaneous equations approach to estimate the coefficients in equation (5) in which the crime rate (Q/N), the probability of imprisonment (C/Q), and the amount of police expenditures per capita (E/N) are simultaneously determined or *endogenous* variables. Ehrlich does not present a system of structural equations for the entire system but only gives the one determining the crime rate, i.e., equation (5). In this system, however, it is the variables W, X, and NW (the only socioeconomic variables included) that "identify" equation (5) and allow for the estimation of the key coefficients of the punishment variables, b_1 and b_2. Fisher and Nagin (1978) have noted that using socioeconomic variables to identify simultaneous relationships can be hazardous. We address this point below, but emphasize again that Ehrlich does not present his entire system so it is impossible for anyone to completely judge his specification (see the related criticisms by Hoenack et al. 1978).

Ehrlich estimates the parameters in (5) for each of the seven basic crime types and for various combinations of crime. Estimates are obtained by two-stage least squares (using a weighting scheme) and also by the method of "seemingly unrelated regressions" due to Zellner (1962). In all cases the coefficient of most interest, b_1, has a negative estimate and is judged to be significantly different from 0 (based upon its t-ratio being larger than 2). We give Ehrlich's two-stage least squares estimates of the coefficients in Table 2.

Ehrlich's final conclusion, based on the analyses just outlined, is that these 1960 data provide strong support for the theory that sanctions deter crime. We take issue with this conclusion. We present detailed comments on these analyses, but first we mention a few key points relating to model formulation and the statistical methods used:

(a) Ehrlich assumed *a priori* the specification of equation (3) thus leading him to estimate a relationship that was linear in the logarithmic scale, i.e., equations (4) or (5). He made no attempt to check the appropriateness of this assumption, or to assess the overall goodness of fit of his model. Many alternatives are available which are consistent with the original econometric formulation.

(b) As noted above the use of the variables, W, X, and NW to identify the crime rate equation, (5), is highly suspect.

(c) No justification is given for the use of W and X in place of the variables measuring the differential monetary returns of crime.

(d) When fallible measures are used for key variables in a regression or simultaneous equation model, it does not suffice to substitute them directly into the model, as Ehrlich did with equation (3). The problems of dealing with models where there are errors in the variables are well-known (e.g., see Sprent 1969 or Zellner 1971).

(e) There is a special statistical technology especially suitable to problems where there are multiple indicators available for given unobservable variables (see Joreskog 1970 for a general formulation, and Bielby, Hauser, and Featherman 1977 for an illustrative application). Ehrlich simply ignores this matter.

(f) The equations for all of the crimes include the monetary variables W and X, even those equations for violent crimes such as murder, assault, and rape. We see no justification for this. (Parenthetically, we note that Ehrlich classifies robbery as a property crime, whereas the FBI classifies it as a violent crime.)

(g) The equations for each crime type are analyzed separately and not linked. In our view a more realistic model would relate crime rates to one another because of the known tendencies of career criminals to substitute one crime type for another (e.g., see Petersilia, Greenwood, and Lavin 1978).

Vandaele (1978) reanalyzed Ehrlich's 1960 data, incorporating a number of different model specifications as well as an attempt to identify possible outliers, i.e., states whose data do not fit the pattern of the remainder. His analyses lead him to essentially the same conclusions as Ehrlich. We merely outline here some of the highlights of Vandaele's work. He supports Ehrlich's use of weighted least squares but notes that weighting the variables does not have any real effect on the conclusions. His analyses are also done in the logarithmic scale, although he does obtain one set of estimates for a model in which Q/N is in the log scale and all other variables are untransformed. The results of this analysis are not very different from the others. Two different models for the supply-of-offenses equation that Vandaele fits include (a) the model of equation (5), where the reduced form contains only the 6 variables in this model together with $(E/N)_{59}$, and (b):

$$\ln\left(\frac{Q}{N}\right) = b_0 + b_1\ln P + b_2\ln T^* + b_3\ln W + b_4\ln X + b_5\ln NW + b_6\ln V_{14} + b_7\ln L_{14} + b_8\ln A_{14} + b_9\ln SMSA + b_{10}\ln M + b_{11}\ln Ed + b_{12}DUMMY + b_{13}\ln N + e, \tag{6}$$

Table 3. Estimated coefficients of ln (C/Q)j and lnTj (from Vandaele (1978))

Offense	Equation [5][a] ln(CQ)j	lnTj	Equation [5][b] ln(CQ)j	lnTj	Equation [5][c] ln(CQ)j	lnTj
Murder	−.492 (−.35)[d]	−.124 (−.69)	−2.944 (−1.68)	.129 (.35)	2.178 (.11)	−.482 (−.11)
Rape	−.771 (−1.39)	−.316 (−.78)	−1.347 (−4.92)	−.699 (−2.30)	−2.979 (−.43)	−1.240 (−.38)
Assault	−3.882 (−.35)	−7.216 (−.37)	−.968 (−3.37)	−1.412 (−2.36)	3.851 (.22)	5.754 (.23)
Robbery	−4.223 (−.57)	−1.336 (−.50)	−1.584 (−6.49)	−.465 (−1.46)	−1.109 (−2.19)	−.357 (−1.24)
Burglary	−.445 (−2.73)	−.793 (−2.99)	−.884 (−6.05)	−1.317 (−4.91)	−.416 (−2.81)	−.547 (−2.43)
Larceny	−1.441 (−1.63)	−2.127 (−1.93)	−1.554 (−1.65)	−2.287 (−1.52)	−1.231 (−1.81)	−1.637 (−1.58)
Auto theft	−.616 (−2.58)	−.341 (−1.66)	−.880 (−3.65)	−.460 (−1.76)	−.650 (−3.22)	−.246 (−1.25)
All offenses	−1.021 (−3.84)	−1.156 (−3.4)	−1.249 (−5.43)	−1.407 (−4.30)	−1.043 (−3.70)	−.824 (−2.57)

[a]Model identified by the exclusion of ln $(\frac{E}{N})_{59}$.

[b]Model identified by the exclusion of ln $(\frac{E}{N})_{59}$ and ln $(\frac{Q}{N})_{59}$.

[c]Model identified by the exclusion of ln $(\frac{E}{N})_{59}$.

[d]Numbers in parenthesis are t-ratios (i.e., the estimated coefficients divided by their estimated standard errors).

where the reduced form includes only the variables (6) together with $(E/N)_{59}$. The estimates of the coefficients of $\ln P$ and $\ln T$ for these two models are given in Table 3, along with estimates for the model in which the crime rate equation is identical to (5), but where $(Q/N)_{59}$ is used as an additional reduced form variable. Vandaele's overall conclusion is that the inferences about deterrence are not sensitive to changes in the specification of the model. A careful examination of Table 3 raises doubts about this conclusion which we will return to later. Finally, Vandaele notes that some states (Utah, Vermont, and Wisconsin) have recorded values for some variables that are inconsistent, e.g., the estimate of C/Q, the probability of conviction for assault in Vermont, is given as 1.56. These are not mistakes in recording but result from the way in which the variables such as C/Q are defined (see Table 1). As we noted earlier the problems here include the lag between offenses and imprisonments as well as cross-overs between states. After redoing the analysis with certain of these states omitted, Vandaele claimed that the results are not substantially changed. He concludes that, although the question of what is the proper model specification needs to be studied further, the fact that the estimates do not depend on the specification lends support to the theory that sanctions are an effective deterrent.

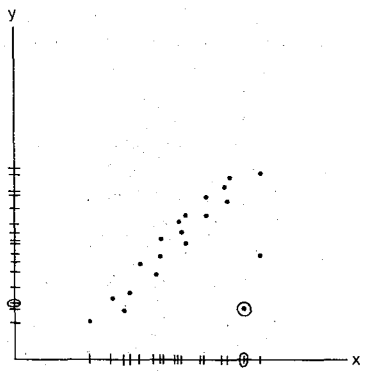
Using the data reported in Vandaele's paper we have reestimated the parameters for the equations determining murder and burglary. In both cases we were able to duplicate Vandaele's estimates up to round-off error. For these two crimes our findings agree with those of Vandaele in that the burglary estimates are insensitive to changes in specification while those for murder are not. We have done the analysis, in both cases, using a model that is linear in the original scale and found that the results are consistent with those for the logarithmic scale. The residual analysis (see, e.g., Daniel and Wood 1972) indicates that either of the models fits the data adequately. We also agree with Vandaele's findings that deletion of the states with values of C/Q greater than 1 does not change the results, but we do not agree that this is not a problem.

Vandaele has stated that his conclusions are not sensitive to changes in the specification of the model. This point is a crucial one since there is not nearly enough theoretical knowledge of the system to determine which *a priori* assumptions are correct. Indeed many are likely to be incorrect. Thus different estimates for the key coefficents corresponding to different specifications would force us to do more thinking about the correct form. A careful examination of Table 3 shows that, especially for the violent crimes of murder, rape, and assault, the estimates do change for the four specifications described above. Note also that there was some instability in the estimates for murder and assault. Thus we do not agree with Vandaele's claim. We see evidence of sensitivity of the conclusion to changes in the specification of the model, and given the comments of Fisher and Nagin (1978) concerning identification, this finding casts doubts on any inferences drawn.

Continuing our analysis in a purely analytic framework, we note that multicollinearity is a serious problem in this data set. The correlation matrix indicates very high correlations among the socioeconomic variables. Our computations were done using a CDC 6600 computer and, although we did not run into any problems in calculating the inverses of matrices, it is well known that even a nearly singular design matrix can lead to unstable estimates of parameters. Another problem is that of "outliers" or what would be more appropriately called influential observations (see R. D. Cook 1977). These are points that are a large distance away from the "x-space" spanned by the other points. Vandaele just looked at univariate plots to determine if any outliers were present, but the use of univariate plots may be very misleading. Figure 1 shows a point which is clearly far from the others but the two univariate plots along the axes do not show this at all. An adaptation of Cook's techniques suitable for use in simultaneous equations problems would have been of great help in Vandaele's and our reanalyses.

The most serious problem with Ehrlich's data and Vandaele's reanalyses of them, however, lies with the data themselves. We are trying to estimate effects of certain variables on others but we cannot observe what we desire. For instance, NW is actually the percentage of blacks in the population, not the percentage of nonwhites. If this variable is attempting to measure minorities then it may be very misleading. We will see below in a reanalysis of 1970 data that changes in the way some of the variables are measured can yield very different results. Our biggest stumbling block is with Ehrlich's measure of the probability of conviction, one of the keys to the analysis. Since this measure is determined by the ratio of the number of convictions to the number of offenses, the "probability" may be larger than 1. While Vandaele minimizes this by showing that states with probabilities larger than 1 do not make a difference, he misses the point raised by the obviously

Figure 1.
Why one-dimensional graphs do not always indicate influential observations



wrong estimated probabilities. This variable is clearly not measuring (or at least not very well) what it is supposed to, and there is no way of determining whether a value of C/Q equal to 0.1 or to 0.2 is an accurate reflection of the probability of conviction in that state! We can all agree that an estimated probability of 1.6 is wrong because we know that probabilities cannot exceed one in value. When we observe estimated probabilities of conviction of 0.9 or 0.5 or 0.3, however, what are we to do? Why should we trust the estimates of magnitude 0.9 or 0.1 any more than we trust those of 1.1 or 1.6? The problem may be that there are transitions between states which preclude using this measurement of the probability. At any rate, it seems impossible to justify seriously any inferences made using this measure of sanction, especially since Ehrlich uses $\ln(C/Q)$ in his equation rather than C/Q. When C/Q is close to zero a small error in estimate is greatly magnified by taking logarithms.

Forst's 1970 data

Forst (1976) presents an analysis of 1970 cross-sectional data that parallels that of Ehrlich. This analysis is based on data from all 50 states plus Washington, D.C. A description of the variables used is given in Table 4. Many of these variables are similar to Ehrlich's but there are some important differences.

There are four variables that are measured differently in the two data sets: Forst measures the average length of prison sentence from the Statistical Abstract (1972) while Ehrlich uses National Prisoner Statistics; he uses *YDSPR* to measure income dispersion while Ehrlich uses the proportion of families earning below the median family income; he measures the proportion of persons between 18 and 20 years of age while Ehrlich uses the proportion between 14 and 24; he measures population density from Census data while Ehrlich uses *SMSA*, the proportion living in standard metropolitan statistical areas. In addition, Forst uses four variables for which there are no corresponding measures in Ehrlich's analysis; namely *MIGR*, *BRHO*, *QJ*, and *AVTMP*. We will see below that the first two of these prove to be quite important.

Forst models the criminal justice system using a five-equation structure:

$$\begin{aligned} CR &= f_1(PJ, AVSENT, QJ, MIGR, URB, BRHO, MFY, YDSPR, UMPL, TEEN, MALE, NWITE, AUTEMP) \\ PJ &= f_2(POL\$, CR, SOUTH, URB) \qquad (7) \\ QJ &= f_3(Y/POP, SOUTH) \\ POL\$ &= f_4(Y/POP, CR_{t-1}) \\ COR\$ &= f_5(Y/POP, CR_{t-1}) \end{aligned}$$

where f_i represents an affine function of the included variables. This formulation allows for transformations of variables such as *PJ*, which is a bounded dependent variable in the second equation. Two important distinctions between this model and Ehrlich's are that Forst considers police expenditure to be an exogenous variable (determined in this case by Y/POP and CR_{t-1}) and, more importantly, Forst uses considerably more variables to determine the crime rate. We also note that Forst analyzes only the aggregate crime rate, not the rates for individual crimes.

Once again we will focus only upon the estimation of the crime rate, the first equation in (7). Forst presents his analysis with the variables measure in their original scale, i.e., the additive relationship as opposed to the multiplicative one assumed by Ehrlich. He states that this is the more appropriate model because of the higher R^2 (a comment which is incorrect statistically). Another point of difference is that Ehrlich assumed that the variance of the errors decreased with N, state population, and hence he performed weighted least squares with $\sqrt{N}$ as weights. Forst did not, empirically, find a need for such a weighting and hence did not use one. At any rate we will see later that the weighting does not have much of an effect on the conclusions.

Forst's estimated coefficients are given in Table 5. These estimates imply conclusions very different from those of Ehrlich. Neither of the deterrence variables has significant coefficients and, in fact, the coefficient of *AVSENT* is positive. Note further that *MIGR*, *URB*, and *BRHO* appear to have strong effects on the crime rate. These three variables were not included in Ehrlich's analysis. Forst also replicated Ehrlich's analysis, as closely as possible, using the 1970 data. He does not present specific results in his paper but notes that the estimated elasticities of the two deterrence variables are substantially smaller than Ehrlich's. It is difficult to judge how different things are since he provides no standard errors. It must be pointed out again that Forst provides no assessment of the fit of his model, aside from reporting R^2 values. More will be said about this analysis as we present our analysis of this data set.

The data set that we used in the analysis of the 1970 data was identical to that of Forst with the exception of three variables, Y/POP, *YDSPR*, and *UMPL*. We took Y/POP to be "personal income" and these data were obtained from the Statistical Ab-

Table 4. Variables used in Forst's analysis

CR	: Number of FBI index crimes per 100,000 residents
PJ	: Estimated probability of apprehension and imprisonment
AVSENT	: Average time served by offenders
QJ	: Expenditures on correction system per prisoner
POL$	: Expenditure on police per state resident
COR$	: Expenditure on correction system per state resident
MIGR	: Population migration rate (population growth divided by number of residents)
URB	: Proportion of residents living in places defined as "urban" by the Census Bureau
BRHO	: Proportion of households that are not husband-wife households
MFY	: Median family income
Y/POP	: Income per capita
YDSPR	: Income dispersion (difference between median family income and national poverty level, weighted by proportion of families below poverty level)
UMPL	: Proportion of the adult population that is unemployed or not in the labor force
TEEN	: Proportion of residents between ages 18 and 20
MALE	: Number of males divided by the number of females
NWITE	: Proportion of residents who are non-white
AVTMP	: Average temperature (Fahrenheit)
South	: Dummy variable distinguishing northern and southern states (south = 1)

Table 5. Estimates of the crime-rate equation in [5]

	Forst analysis		Our reanalysis		
Variable	Elasticity	t-Ratio	Elasticity	t-Ratio	Coefficient[a]
PJ	−.02	−.14	.00	.01	31×10^{1}
AVSENT	.01	.10	−.00	−.00	-46×10^{-2}
QJ	−.07	−.64	−.06	−.60	-16×10^{-3}
MIGR	a	3.42	−.00	3.70	33×10^{2}
URB	.71	2.17	.74	2.40	26×10^{2}
BRHO	.96	2.13	.97	2.30	78×10^{2}
MFY	.60	1.51	.57	1.50	15×10^{-2}
YDSPR	.40	1.93	.33	1.80	13×10^{-1}
UMPL	.11	.21	.20	1.90	98×10^{2}
TEEN	.63	1.45	.69	1.70	30×10^{3}
MALE	.65	.65	.43	.20	21×10^{2}
NWITE	−.07	−1.51	−.07	−1.70	-15×10^{2}
AVTEMP	.11	.37	.21	.70	9.3×10^{1}

[a]Not computed by Forst.

stract (1975). In computing *YDSPR* we took the national poverty level to be $3,601 as given in the Statistical Abstract (1975). The other discrepancy was for *UMPL*, which was supplied by Forst, but over average value differs from that given by Forst.

Using the same model as Forst, we derived the estimated coefficients for the crime rate equation given in Table 5. Our estimates are similar to Forst's with only that for the coefficient of *UMPL* being substantially different. Note that our estimate of the *PJ* coefficient is in fact positive although there is no strong evidence, in either case, of the actual coefficient being different from 0. Our analysis of the residuals from the fitted equation indicated that Michigan might be a possible outlier (it had a standardized residual of 2.6). Although testing this point as an outlier using a Bonferroni *t*-test (see Miller 1966) did not lead to rejection, we reestimated the parameters with this state omitted and found that they did not change substantially. Scatter plots of the variables indicated that the District of Coumbia (due to very high values of *NWITE, BRHO*, and *POL$*) and Alaska (with a very high percentage of males) were well away from the center of the array of independent variables. However, eliminating either or both of these points does not change the results either. One final point of note is that there was a tendency for the variance of the residuals to increase with *URB*. However, no heteroscedasticity was observed when the residuals were plotted against *N*.

In Table 6 we give estimates for the coefficients when the model has all variables in logarithms and when only *CR* is in logarithms. Again the seemingly important variables remain the same. In both of these models there are no outlier problems nor is nonconstant variance apparent. Hence there is evidence that the logarithmic scale is the better one to work in, although the choice of scale doesn't affect the findings in a meaningful way.

As mentioned above, Forst replicated Ehrlich's model as closely as possible and obtained results for 1970 data that were not consistent with Ehrlich's. We have also done this but our results do not agree with either Forst or Ehrlich. The estimates we obtained are given in Table 6. The only significant variables in our analysis are *MFY* and *YDSPR*. Evidently the slight differences between Forst's variables and ours are giving very different results. The residual plots for this model, however, showed a definite need for fitting the additional variables, *URB* and *MIGR*, so we would not want to use the model, as it stood, for making inferences.

Despite these discrepancies, our conclusion is the same as that of Forst: Using variables measured similar to those in his study, we find no evidence that, in 1970, sanctions were an effective deterrent to crime.

The explanation for the discrepancy between Ehrlich's results for 1960 and Forst's ones for 1970 may be simply that behavior patterns of criminals changed in these 10 years. While certainly a possibility, this cannot reasonably be inferred from the data at hand. In comparing the two analyses, it must be remembered that the 1970 data include only an overall crime rate and there has been no individual analysis for different crimes. Of course, we can compare the two results in terms of the overall crime rate but this may not be very fruitful as there are surely different structural relationships for different crime types which are masked by such an aggregation.

Other analyses of cross-sectional data for murder

Given the widespread interest in the potential deterrent effect of capital punishment on the commission of homicide, several investigators have done special analyses of cross-section data by state specifically for this crime. These include Passell (1975) who examines data for 1950 and 1960, Forst (1977) who analyzes data for 1960 and 1970 (actually he used differences), Ehrlich (1977b) who uses data for 1940 and 1950 (his analysis of murder for 1960 was described earlier in this section), and Loftin (1980) who examined a quite different data set for 1960 which had been explored earlier by several sociologists. Because the longitudinal analyses described in the next section deal only with homicide, we defer a discussion of these cross-sectional studies until the end of that section.

Table 6. Estimated coefficients for various models fit to Forst data

Variable	Model (1)[a]		Model (2)		Model (3)
$\hat{PJ}$	11×10^{-2}	(.42)[b]	2.2	(.22)	−.21
AVSENT	3.7×10^{-2}	(.13)	-1.9×10^{-2}	(−.30)	−.16
QJ	-3.9×10^{-2}	(−.20)	-1.0×10^{-2}	(−.77)	–
MIGR	80×10^{-2}	(3.33)	1.3	(3.17)	–
URB	81×10^{-2}	(2.46)	1.4	(2.75)	–
BRHO	57×10^{-2}	(1.12)	2.1	(1.31)	–
MFY	1.2	(2.67)	97×10^{-6}	(2.11)	1.7
YDSPR	62×10^{-2}	(2.95)	73×10^{-5}	(2.09)	1.0
UMPL	31×10^{-2}	(2.82)	4.9	(2.04)	–
TEEN	91×10^{-2}	(2.17)	20	(2.41)	–
MALE	−2.0	(−.91)	−2.5	(−.61)	–
NWITE	-6.9×10^{-2}	(−1.97)	−.79	(−1.98)	.04
AVTEMP	33×10^{-2}	(.92)	45×10^{-5}	(.76)	–

[a]The three models used are:

(1) All variables, except SOUTH, transformed to natural logarithms.
(2) Only CR is transformed to natural logarithms.
(3) This corresponds to Ehrlich's model using the 5 corresponding variables available in this data set.

[b]Numbers in parentheses are t-ratios.

Deterrence, capital punishment, and murder

Ehrlich's longitudinal data for murder

One of the most controversial studies of deterrence is that of Ehrlich (1975a). Based upon a time series of aggregated national data for the years 1933-1969, Ehrlich claims to have found strong evidence that the death penalty has a deterrent effect upon potential murderers. In this section we review Ehrlich's analysis as well as prominent criticisms of it. We then present the results of our reanalysis of essentially the same data set.

Ehrlich's model is one in which the murder rate, probability of apprehension, and probability of conviction given apprehension are endogenous variables having simultaneous effects on each other over time, while a number of socioeconomic variables are considered to be exogenous variables along with the probability of execution given conviction. The murder supply function, by an elaboration of the earlier arguments, now replaces p^* and f^* of equation (2) by P_a (the probability of arrest), $P_{c|a}$ (the probability of conviction given arrest), and $P_{e|c}$ (the probability of execution given conviction). The empirical implementation of this function, using the variables described in Table 7, takes the form:

$$\left(\frac{Q}{N}\right) = C P_a^{\beta_1} P_{c|a}^{\beta_2} P_{e|a}^{\beta_3} U^{\beta_4} L^{\beta_5} \times Y_p^{\beta_6} A^{\beta_7} \exp(\beta_8 T) \exp(u). \tag{8}$$

Note that T (time) is entered into this equation in a manner different from all other variables. Time is used here as a surrogate for the improvement of medical technology over time and, as a variable, it plays a crucial role in the analysis, as we shall see below.

Since the observations all involve aggregate national data measured annually for 1933–1969, Ehrlich assumes that the errors are subject to first-order serial correlation, i.e.,

$$u_t = \rho u_{t-1} + e_t, \tag{9}$$

where ρ is the serial correlation and the e_t are independent random errors. The equation whose coefficients he thus sets out to estimate is

Table 7. Variables used in the time-series analyses

$\frac{Q}{N}$	:	Murder rate (per 1000 civilian population)
P_a	:	Probability of arrest (percent of murders cleared)
$P_{c\|a}$	:	Proportion of those charged that were convicted of murder
PXQ_1	:	Number of executions for murder in the year t + 1 divided by the number of convictions in year t
PXQ_2	:	Number of executions for murder in the year t divided by the number of convictions in year t
L	:	Proportion of the civilian population in the labor force
U	:	Proportion of the civilian labor force that is unemployed
A	:	Proportion of population in the age group 14-24
Y_p	:	Friedman's estimate of (real) permanent income per capita in dollars
T	:	Time (years)
NW	:	Proportion of non-whites
N	:	Civilian population (in 1000's)
XGOV	:	Per capita (real) expenditures (excluding national defense) of all governments in millions of dollars
$XPOL_{-1}$	:	Per capita (real) expenditures on police in dollars lagged one year
C	:	Violent crime rate (offenses of rape, robbery and aggravated assault)

$$\Delta \ln\left(\frac{Q}{N}\right) = \beta_0 + \beta_1 \Delta \ln P_a + \beta_2 \Delta \ln P_{c|a} + \beta_3 \Delta \ln P_{e|c} + \beta_4 \Delta \ln U + \beta_5 \Delta \ln L + \beta_6 \Delta \ln Y_p + \beta_8 \Delta T + e, \tag{10}$$

where, for a generic variable Z, the value of ΔZ at time t is

$$\Delta Z_t = Z_t - \rho Z_{t-1}, \tag{11}$$

and $\beta_0 = \ln C$.

In a simultaneous equation framework, one must be concerned with whether or not the structural equations are identified. This particular equation is identified by omitting a number of socioeconomic variables from this equation. The variables that were omitted must have a direct effect on some of the other endogenous variables so that equation (10) can be identified, i.e., for the parameters to be distinguishable. Fisher and Nagin (1978) point out the difficulty inherent in using socioeconomic variables to identify the parameters in a structural equation. It is impossible to determine whether Ehrlich has validly identified his structural equations because he only presents the one equation given in (10).

Ehrlich presents estimates of the parameters in equation (10) using six different measures of $P_{e|c}$, the probability of execution given conviction. Two of these are defined in Table 7. In all but one of the six cases $P_{e|c}$ is viewed as an exogenous variable. He finds that for all six measures the estimated coefficient of the $P_{e|c}$ variable is negative. Also, for four of the six, the estimated coefficient is significantly different from zero (having a t ratio smaller than −2). He also repeats the analysis after excluding some years from the beginning and some from the end of the series. These modifications do not change his results appreciably. Using two of the estimates, Ehrlich derives an estimate of the average number of murders that would be prevented by one additional execution per year. He estimates this to be between 7 and 8.

Ehrlich concludes that his analyses show the deterrent effect of capital punishment. He claims that he has analyzed the data in scales other than the logarithmic one without the conclusions changing substantially, and he also states that his results are unaffected by the time period of analysis. We show below that there is considerable doubt about these two claims.

Criticism of Ehrlich's conclusions can be divided into two broad categories: the relevance and accuracy of the data sources, and the methodology used. We address these in turn. The reliability of the data sources that Ehrlich used has been questioned before. We noted earlier the severe shortcomings of the UCR crime rates. They are not accurate measures of the variables that they claim to represent. Furthermore, Bowers and Pierce (1975) point out that the UCR arrest and clearance rates are especially suspect, primarily in the earlier years. What is particularly troublesome is that recording practices have changed so much over time. Moreover, while the UCR murder rates have been reestimated for the earlier years, the arrest and clearance rates have not been so adjusted. When one is making inferences based upon a time series, such dramatic changes in the coverage of data collection inevitably are confounded with any real effects that are present.

A second data-related problem is conceptual as well. The variable $P_{e|c}$ relates to crimes of murder subject to punishment by execution, actually a small proportion of

Table 8. Comparison of estimated coefficients in equation [10]

	$\hat{\rho}$	Constant	P_a	$P_{c\|a}$	PXQ_1	U	L	Y_p	A	T
	Ehrlich's estimates									
a)	.257	−3.176	−1.553	−.455	−.039	.067	−1.336	1.481	.630	−.047
		(−.78)	(−1.99)	(−3.58)	(−1.59)	(2.00)	(−1.36)	(4.23)	(2.10)	(−4.60)
	Our estimates[a]									
b)	.257	−2.02	−.56	−.25	−.038	.01	−.91	.65	.71	−.02
		(−2.61)	(−.55)	(−1.57)	(−1.24)	(.26)	(−.71)	(2.64)	(2.28)	(−2.6)

[a]Our estimate of ρ was actually $\hat{\rho}$ = .55 but we present estimates using the same value as Ehrlich obtained for comparison purposes. Our estimates were not sensitive + changes in ρ except for ρ near 1.

all murders and one which varies from state to state. Yet the UCR data used by Ehrlich for the number of murders, the number of arrests, and the number of convictions refer to all murders and nonnegligent homicides rather than to only capital crimes. If both capital and noncapital murder rates have production functions of the form (8), then the overall murder rate cannot have a production function of this form (see the related discussion in Hoenack and Weiler 1977). Ehrlich's discussion of this problem is noninformative and sidesteps the issues.

The data used for $P_{c|a}$ present further problems. In particular, the data for $P_{c|a}$ and *XPOL* were not available for odd years from 1933–1951. Ehrlich estimated these values by a regression technique that he does not describe. Among the many problems inherent in this type of procedure is the loss in real degrees of freedom due to the missing data. The effective degrees of freedom for estimation is an important issue here and more will be said about it later.

A fourth problem that Ehrlich faced was that there were no executions after 1967. He arbitrarily defined the number of executions in those years to be one in order to be able to take logarithms. This is indicative of a basic flaw in the production function model of equation (8). It is simply inappropriate for application to a social structure allowing zero executions, since it predicts an essentially infinite murder rate for such situations if the coefficient β_3 is negative (i.e., the sign associated with a deterrent effect of capital punishment).

An overriding issue related to the data is the choice of aggregate data for making inferences. Baldus and Cole (1975) include an excellent discussion of the dangers of making causal inferences using nationwide crime data. They note that crime patterns have differed over time from state to state as has the use of the death penalty. If murder rates increased in states that used the death penalty often, but rates decreased in states not using the death penalty, the use of nationwide data might completely obscure this fact.

Assuming that the variables used in the study are faithful measures of what we really want to observe (which we believe to be untrue), there are still a number of methodological questions to be answered. Ehrlich's model is linear in the logarithmic scale (although he leaves *T* untransformed) and a number of authors have shown that the conclusions reached are dependent upon the scale of measurement. Passell and Taylor (1977) and Bowers and Pierce (1975), using data sets that are effectively identical to Ehrlich's, find that the coefficient of $P_{e|c}$ is not significantly different from zero when estimated in a model that is linear in the original scale of measurement. Finally, we note that all of the variables actually used by Ehrlich are fallible measures of the variables of interest, and that the models of real interest should thus involve errors-in-the-variables, and multiple-indicator structures. The difficulties here are the same as those described in the last section.

Another methodological question is whether the relationship between the actual variables used is changing over time. Passell and Taylor and Bowers and Pierce indicate that the data for years after 1963 exert heavy influence on the estimated coefficients. In their analyses, both sets of authors show that the effect of $P_{e|c}$ is not significant when these later years are deleted from the data set (see the related discussion in Klein, Forst, and Filatov (1978)). Passell and Taylor perform an F-test to determine if the regression function is the same for the years before 1963 as for the years after 1963 and conclude that it is not. (Unfortunately, as Ehrlich (1977a) notes, the properties of their test statistic are not known.) Further, Bowers and Pierce argue that doing the analysis in the logarithmic scale gives more weight to these later years. In his rebuttal to Bowers and Pierce, Ehrlich (1975b) claims that there is no justification for arbitrarily deleting data points and that doing so loses precious degrees of freedom. While there is some validity to Ehrlich's response, the deletion of these points is not arbitrary if a fundamental change in American society took place around 1963, as has been argued by many criminologists and sociologists. Moreover, as we found in our reanalyses described below, these years stand out as discrepant in various forms of residual analysis.

Klein, Forst, and Filatov (1978) consider a number of additional exploratory variables that Ehrlich might have used. One is the average length of time served by convicted murderers—it very well may be that this is the important variable for explaining the increase in murders. Unfortunately, these data are not readily available. Another variable that they do use is an overall index of violent crime. The justification for its inclusion is that murder may be increasing as a by-product of an overall increase in the level of lawlessness. Reanalyzing Ehrlich's data, they find that with this extra variable in the structural equation the coefficient of $P_{e|c}$ is no longer significantly different from zero.

Hoenack and Weiler (1977) reanalyzed the Bowers-Pierce data, using a fully specified simultaneous system of equations, a theoretical justification for which is given in Hoenack, Kudrle, and Sjoquist (1978). Their model interprets Ehrlich's murder supply function as the society's response to murder behavior, not vice versa. For their specification, the coefficient of the execution variable is positive although never more than one standard deviation from zero. A key feature of their specification is the separation of the variable *A*, the proportion of the population between the ages of 14 and 24, into two parts: one for the proportion of juveniles (ages 14–18) and one for young adults (ages 19–24). We also note that their alternative to (8) does not require the inclusion of *T* (its estimated coefficient is essentially zero when included); this point is especially interesting given the crucial role played by *T* in Ehrlich's analysis and specification, which we discuss in detail shortly.

These criticisms raise a number of questions. The estimated coefficients differ depending on the scale in which the analysis is done, but which is the proper scale? The estimates are affected by the later years (1963–1969) but what distinguishes these later years from the earlier ones? A further criticism deals with the basic formulation of the model as a simultaneous set of equations and the identification of the key structural equation for the supply of murders. We use a somewhat different approach in our analysis and try to answer some of these and other questions.

In our analyses we have used data furnished by Bowers and Pierce and used in their analyses. Ehrlich consistently refused to make his data available to us and to others for reanalysis. The only versions of the variable $P_c|_c$ that we had values for were PXQ_1 and PXQ_2. The bulk of our analysis focused on using either PXQ_1 or its lagged values, $PXQ_{1(-1)}$, as a measure of probability of execution given conviction.

Our first goal was to see if we could reproduce Ehrlich's results using this data set. In Table 8 we give our estimates, along with Ehrlich's, of the coefficients in the murder rate equation. Note that the signs of all coefficients agree but that there are some differences. The likely cause of the different estimates is in Ehrlich's estimation procedure based on the method of Fair (1970). We have programmed a method suggested by Fair ourselves to get estimates, while Ehrlich used a packaged routine. We suspect the problem is in how the modified first differences are created. Fair actually suggests a number of different methods, some of which are asymptotically more efficient than others. The method that we adopted uses $(\hat{Y}_1 - p\hat{Y}_{1=1})$ as the values for the endogenous variables in the second stage regression. Thus in our analysis the effective years are from 1934–1969. Ehrlich appears to have used a somewhat different differencing operation which allowed an analysis only for the period 1935–1969. At any rate, the conclusions based upon our estimates are not substantially different from Ehrlich's. A major discrepancy occurs in the estimates of $\hat{\rho}$. We found $\hat{\rho} = 0.55$ while Ehrlich reports $\hat{\rho} = 0.257$. Since our estimates were not very sensitive to changes in $\hat{\rho}$ (except near 1) we used Ehrlich's value to get the remaining estimates.

We take issue with Ehrlich's arguments for the endogenous variables Q/N, P_a, and $P_c|_a$ having simultaneous effects on one another. It makes far more sense to us to think of a criminal's subjective assessment of punishment rates as affecting his current behavior, but that the murder rate affects the punishment variables after some delay of time. The delay might be 6 months, 1 year or even 2; however, since the data only allow for delays which are multiples of a year, we have arbitrarily fixed on a delayed effect of murder rates on punishment of 1 year. This assumption means that our system of equations, unlike Ehrlich's, is recursive not simultaneous, and thus we do not have Ehrlich's identification problem nor the problem of making inferences about structural parameters. Our model differs in this way from those used by Ehrlich, Bowers and Pierce, and Passell and Taylor, and we use it primarily to make a series of methodological points.

Our model for reanalysis is thus of the form:

$$\left(\frac{Q}{N}\right) = a_0 + a_1 P_a + a_2 P_c|_a + a_3 PXQ_1 + a_4 L + a_5 U + a_6 A + a_7 Y_p + a_8 T + a_9 C_1 + e_1, \tag{12}$$

$$P_a = b_0 + b_1 \left(\frac{Q}{N}\right)_{(-1)} + b_2 T + b_3 XGOV + b_4 XPOL_{(-1)} + b_5 N + e_2, \tag{13}$$

$$P_c|_a = c_0 + c_1 \left(\frac{Q}{N}\right)_{(-1)} + c_2 PXQ_{1(-1)} + c_3 P_{a(-1)} + c_4 T + c_5 \mathrm{NW} + c_6 XGOV + c_7 A + e_3, \tag{14}$$

where the errors, e_i, are assumed to be independent with mean 0 and variance σ_i^2 (i.e., independent for different points within equations). A subscript of (−1) indicates that the variable was lagged by 1 year. By assuming that the errors are independent for different points within equations, we have set $\rho = 0$. This zero value for the serial correlation is completely consistent with our findings in the replication of Ehrlich's results, greatly simplifies the estimation, and allows for detailed residual analyses based on standard methods for multiple linear regression. Those residual plots that we have examined indicate that this assumption may not quite be met, but when we have reestimated the parameters assuming different values for the first-order serial correlation coefficient we have found that the estimates do not change substantially. The residual plots indicated that the correlation might be of greater than first order but this possibility was not explored. If such higher order serial correlation exists, it affects Ehrlich's analyses as well as our own.

Figure 2.
Residual plot for the crime rate equation in [12]

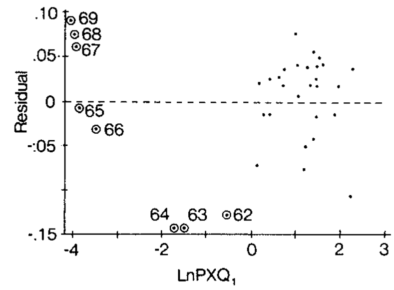


Table 9 gives estimates of the coefficients in the murder rate equation for different transformations. Note that the first two sets of estimates are for an equation that does not include C, the violent crime level index. We considered three different measures of $P_c|_c$: PXQ_1, $PXQ_{1(-1)}$, and PXQ_2. The only significant coefficients (whose t-values, which are given in parentheses, are in excess of 2) are for PXQ_1 when variables are measured in the original scale or when only Q/N is measured in the logarithmic scale. These equations are the only ones to contain significant coefficients for the other punishment variables, namely $P_c|_a$.

Bowers and Pierce (1975) conclude that Ehrlich's results are heavily dependent on the analysis being done in the logarithmic scale. They state that due to the very low number of executions after 1962, taking logarithms of PXQ_1 emphasizes the effect of these later years thus yielding a negative coefficient which is significantly different from zero. Our conclusion, after studying the residual plot shown in Figure 2 as well as other graphs, is in conflict with theirs. Taking the logarithms of PXQ_1 moves the values of PXQ_1 in later years far from the center and tends to flatten out the slope. These residual plots do, however, support the conclusions of Bowers and Pierce and of Passell and Taylor that the years after 1962 do not fit the pattern of the previous ones.

Other residual plots that we examined indicated a second problem which has not been noted in previous analyses. Large residual values corresponding to 1934 point to it as a possible outlier. Looking at the original data, we find that for 1934 there is a very large value of PXQ_1 and a small value of Q/N. We suspected that this point might have a strong effect on the estimates so we reestimated the coefficients in both the original and the logarithmic scale.

Table 9. Estimated coefficients for equation [12] under various transformations

Transformation[a]	P_a	$P_{c\|a}$	PXQ_1	L	U	A	Y_p	T	C
1	−1.60	−.089	.027	−.80	−.06	.91	−.05	.09	
	(−1.65)	(−.64)	(1.08)	(−.57)	(−1.71)	(2.94)	(−.20)	(.82)	
2	−.75	−.15	−.043	−.42	.012	.63	.55	−.023	
	(−.94)	(−1.25)	(−1.43)	(−.32)	(.03)	(2.17)	(2.5)	(−2.56)	
3	-9.7×10^{-5}	-39×10^{-5}	-26×10^{-4}	−.23	-5.1×10^{-4}	.015	8.8×10^{-6}	-18×10^{-4}	9×10^{-8}
	(−.31)	(−3.25)	(−4.03)	(−3.19)	(−2.55)	(.30)	(.88)	(−7.5)	(9.0)
4	-1.1×10^{-3}	-6.2×10^{-3}	−.042	−3.6	-8.7×10^{-3}	.622	1.9×10^{-4}	−.033	15.7×10^{-7}
	(−.21)	(−2.95)	(−3.82)	(−3.00)	(−2.64)	(.68)	(1.12)	(−7.86)	(6.28)
2	−.070	−.017	−.015	−.930	−.031	.620	.320	−.026	.330
	(−.08)	(−.12)	(−.44)	(−.72)	(−.78)	(2.21)	(1.28)	(−3.25)	(1.74)
1	−1.23	.037	.071	−1.3	−.12	.94	−.44	.16	.34
	(−1.26)	(.23)	(1.87)	(−.93)	(−2.4)	(3.13)	(−1.22)	(1.33)	(1.48)
	P_a	$P_{c\|a}$	$PXQ_{1(1-1)}$	L	U	A	Y_p	T	C
1	−.76	3.8×10^{-3}	−.035	−.92	−.03	.43	.44	−.32	−.02
	(−.82)	(.03)	(−.92)	(−.66)	(−.60)	(1.59)	(1.16)	(−2.13)	(−.10)
3	-7.6×10^{-5}	-16×10^{-5}	-73×10^{-5}	−.18	-29×10^{-5}	.039	1.3×10^{-5}	-17×10^{-4}	9×10^{-8}
	(−.32)	(−1.88)	(−1.74)	(−3.0)	(−1.61)	(.93)	(1.59)	(−8.5)	(9.00)
	P_a	$P_{c\|a}$	PXQ_2	L	U	A	Y_p	T	C
1	−1.2	−.12	−.027	−1.5	−.06	.58	−.007	.002	−.10
	(−1.20)	(−.71)	(−.68)	(−1.00)	(−1.00)	(1.87)	(−.02)	(.02)	(−.40)
3	17×10^{-5}	-19×10^{-5}	-99×10^{-5}	−.25	-65×10^{-5}	.063	4.9×10^{-6}	-15×10^{-4}	9×10^{-8}
	(.43)	(−1.27)	(−1.14)	(−2.63)	(−2.6)	(.97)	(.38)	(−5.00)	(4.5)
	P_a	$P_{c\|a}$	PXQ_1	L	U	A	Y_p	T	C
1[b]	−.99	.14	.063	−.64	−.084	.81	−.029	−.13	.39
	(−1.16)	(1.00)	(1.91)	(−.49)	(−1.83)	(3.00)	(−.09)	(−.93)	(1.95)
3[b]	-1.0×10^{-4}	-1.8×10^{-4}	-7.6×10^{-4}	−.20	-4.0×10^{-4}	.035	1.1×10^{-5}	−.002	9×10^{-8}
	(−.38)	(−1.5)	(−1.04)	(−3.33)	(−2.35)	(.78)	(1.31)	(−7.62)	(9.00)

[a]Transformations are designated as follows:

1. All variables in logarithms,
2. All variables except T in logarithms,
3. All variables untransformed,
4. $\frac{Q}{N}$ in logarithms, all others untransformed.

[b]The data for 1934 were deleted in estimating the coefficients.

These estimates, given in Table 9, show that things do change considerably once 1934 is deleted. The estimated coefficient in the original scale is still negative but not significant while the estimate in the logarithmic scale is now positive. The residual plots after deletion of this point do not indicate any other outlier problems. We note that in Ehrlich's analyses, a significant result is obtained when using $PXQ_{1(-1)}$ but not with PXQ_1. With the lagged variable, $PXQ_{1(-1)}$, the value for 1934 is used while for PXQ_1 that year's data are excluded from the analysis. This seems to support our finding regarding the suspect nature of the 1934 data.

It is not surprising that one point can exert such a strong effect on the estimated coefficients. Even in our recursive model there are only 25 degrees of freedom available for estimating the coefficients in the model. In Ehrlich's model the problem is much worse. Although Ehrlich, as well as other authors, have routinely computed the degrees of freedom associated with their estimates, it is not clear what the effective degrees of freedom actually are in a two-stage estimation problem. In Ehrlich's first stage regression, there are 18 independent variables in addition to the 9 independent variables in the second stage regression. As far as we know, there has been no work done on determining the appropriate degrees of freedom for this type of problem but we would think there are considerably fewer than 25. The difficulty in computing degrees of freedom is compounded by Ehrlich's estimation of the missing values of $P_{c|a}$ and *XPOL* using the remaining data.

Table 9 contains only the results for equation (12) of our three-equation model. We have analyzed the other equations as well but do not report on them here since they do not affect the deterrence hypothesis.

Earlier we noted the arbitrary choice by Ehrlich of the use of logarithms for all variables but *T*, time. In Table 9 we show some equations where ln *T* was used in place of *T* in a fully logarithmic specification. The changes in sign for the coefficient of *PXQ* that go with this change in specification are suggestive that the choice of scale for the variable *T* may have a strong influence on the coefficient. Now the arbitariness of Ehrlich's choice hits home. Why choose *T* or log *T* instead of T^{-1}, or $\ln(T-1900)$, or even $\ln(T-1776)$? In other reanalyses we have discovered that suitable transformations of *T* can dramatically change the size and sign of the coefficient of PXQ_1.

In summary, we find that these data do not support Ehrlich's conclusion that there is a deterrent effect created by an increase in the probability of execution. First, we find the model formulation suspect, and inappropriate for application in situations with essentially zero execution rates. Second, we question the choice of data used to measure key variables in the model. Third, we have noted that the analysis is sensitive to the specification of the model. Using a recursive model we have obtained results that are different from Ehrlich's. The question of which model is more appropriate is not an easy question to answer, but we think that there is as much *a priori* support for our model as for Ehrlich's. Our residual analysis does not indicate any lack of fit other than the two problems discussed above; in contrast Ehrlich does not examine the goodness of fit of his model. Others, such as Bowers and Pierce (1975) and Hoenack and Weiler (1977), have also formulated alternative model specifications which when used in analyses make the deterrent effect of capital punishment disappear. In all, there are far too many flaws in Ehrlich's model, his data, and his analyses for him to claim that a real deterrent effect of any sort has been found using this form of longitudinal data.

Cross-section analyses of murder rates

To buttress the arguments in his paper on the analysis of the longitudinal data on murder, Ehrlich (1977a) has also analyzed the cross-sectional variations of murder and execution in 1940 and 1950. The basic regression model used in his analyses resembles equations (4), (5), and (10).

$$\ln\left(\frac{Q}{N}\right) = b_0 + b_1 \ln T^* + b_2 \ln\left(\frac{Q}{C}\right) + b_3 \ln PXQ + b_4 \ln NW + b_5 \ln X + b_6 \ln W + b_7 \ln A + b_8 \ln U + e. \tag{15}$$

Ehrlich uses ordinary least squares to estimate the coefficients in equation (15) and uses data first for states with positive executions, and then for all states. He also attempts to compare the results of a fully linear specification with (15), using an approach suggested by the method of Box and Cox (1964). His comparison strongly favors the log-log specification of (15). Ehrlich finds the estimated coefficients of T^*, Q/C, and PXQ (measured in several different ways) to be negative and significant at at least the 0.05 level *both* for the linear and log-log specifications. His conclusion is that these data and the analyses of them corroborate his earlier analysis of the longitudinal data.

As in Ehrlich (1975a) these conclusions seem at first sight convincing, until we note that the data used have even greater shortcomings than those noted earlier and that almost all of the other analytical problems mentioned earlier remain. Moreover, Ehrlich's results run contrary to those of other investigators who have examined cross-sectional data. For example, Passell (1975) used 1950 and 1960 data to estimate the coefficients in the model,

$$\left(\frac{Q}{N}\right) = b_0 + b_1 T^* + b_2 \left(\frac{C}{Q}\right) + b_3 PXQ + b_4 A + b_5 I + b_6 M + e, \tag{16}$$

where I is the percentage of the family population below an arbitrary cash income poverty line, and M is the ratio of net nonwhite migrants in the previous 10 years to the total population. Using both ordinary and two-stage least squares he found positive (but not significant) estimated coefficients for PQX, and negative (and significant) estimated coefficients for C/Q and T^*. Passell's specifiction, unfortunately, has little more to recommend it than does Ehrlich's, but his results do illustrate the importance of the specification on the results and the inferences one is likely to draw from the analysis.

Forst (1977) examined data on the change in the crime and punishment measures that occurred between 1960 and 1970 for all 50 states. In regressions based on a subset of 32 states for which complete data were available, he found the execution variable to have a positive coefficient. Although his results appear to be in agreement with Passell, they are almost certainly dominated by the fact that PXQ for 1970 for all states was zero! Forst models the change in the homicide rate, $\Delta(Q/N)$, as a function of the change in execution rates, $\Delta(PXQ)$, and changes in other variables. Since $PXQ = 0$ for all states in 1970, Forst thus models $\Delta(Q/N)$ as a function of PXQ for 1960 and the changes in other variables. We do not understand the logic behind this specification.

Finally we note the analyses carried out by Loftin (1980) for 1960 cross-sectional data using a markedly different set of variables aside from C/Q and T^*, motivated primarily by sociological rather than economic considerations. Loftin finds little to support the inclusion of the punishment variables in a regression equation with either a linear or a log-log specification.

We have concluded that these cross-sectional analyses offer no support to the conclusion that there is a deterrent effect of capital punishment, and the conflicting results for the other punishment variables cast serious doubt on any attempt to infer deterrent effects.

Conclusions

Becker's (1968) paper has stimulated many economists and others to use modern statistical methods for the analysis of regression and simultaneous equations models to search for evidence in support of the deterrence hypothesis. Following Ehrlich's (1973a, 1975a) pioneering attempts to implement Becker's theoretical model, the flood of papers and manuscripts on the analysis of crime and punishment data has been almost overwhelming.

What has this work contributed to our knowledge of the deterrent effects of punishment on crime? We have concluded that little or no progress has been made during the past 10 years in our understanding of the potential deterrent effects of punishment on crime. Indeed much of the controversy that erupted over Ehrlich's work has served to divert the efforts of serious scholars of crime from more productive pursuits to a battle with Ehrlich and his supporters. The battle has raged before the United States Supreme Court, which heard arguments based on Ehrlich (1975a) and Passell and Taylor (1977) in the case of *Fowler v. North Carolina*. It has filled the pages of many different journals. Some journals devote entire issues to the topic (e.g., see *Journal of Behavioral Economics*, Vol. 6, Numbers 1 and 2, Summer/Winter, 1977). It has been investigated by a panel established by the National Academy of Sciences. And in the end we seem to be no further ahead of where we were 10 years or more ago.

In this paper, we have reviewed a large proportion of the empirical attempts to model the economics of crime and punishment, including the work of Ehrlich (1973a, 1975a) based on the model suggested by Becker (1968). We have concluded that:

(a) The Becker-Ehrlich model has glaring shortcomings and, when examined critically, does not lead to the claimed testable hypotheses regarding the effect of punishment of crime.

(b) The use of the Becker-Ehrlich model for aggregate data requires extensive justification that has never been given.

(c) The crime and imprisonment data used to empirically examine the Becker-Ehrlich model are so untrustworthy as to render any serious analysis meaningless.

(d) The empirical implementations of the Becker-Ehrlich model are badly flawed and have extremely grave statistical shortcomings, and most published conclusions from them are not to be trusted.

(e) Even if one accepts the Becker-Ehrlich model, and Ehrlich's choice of data to implement it (which we do not), Ehrlich's affirmative conclusions regarding the deterrent effect of punishment on crime in general, and of capital punishment on murder in particular, do not stand up to careful statistical scrutiny.

Ehrlich's critique of the National Academy of Sciences Panel's report (Ehrlich and Marks 1978) is essentially an attempt at self-justification. He argues that only his version of the econometric model of deterrence is relevant, that any work inconsistent with it can be dismissed out of hand, that only analyses done by Ehrlich (or those confirming his findings) are correct or appropriate, and that all of the work of others whose conclusions run counter to Ehrlich's is incorrect, distorted, technically flawed, theoretically eclectical, or irrelevant. This is hardly a dispassionate perspective, and we find little in the critique which counters the crucial points made in the preceding sections and in the Panel's report itself.

We find no reliable empirical support in the existing econometrics literature either for or against the deterrence hypothesis. Moreover, we believe that little will come from further attempts to model the effects of punishment on crime using the type of data we have described in this paper.

Acknowledgments

The authors are indebted to several of the participants of the SSRC-LEAA Workshop whose ideas and suggestions inevitably have found their way into this article, although perhaps in a distorted way. In particular, thanks are due to Albert D. Biderman, Alfred Blumstein, William Fairley, Kinley Larntz, Colin Loftin, Daniel Nagin, Albert J. Reiss Jr., David Seidman, and Leslie Wilkins. We also wish to thank the many others who have shared their thoughts and work with us, including William Bowers, Phillip Cook, Franklin Fisher, Brian Forst, Stephen Hoenack, Paul Meier, Frederick Mosteller, Glenn Pierce, Christopher Sims, Walter Vandaele, Ann Witte, and William Weiler. William Bowers, Brian Forst, and Walter Vandaele all provided their data for our reanalyses; Isaac Ehrlich consistently refused to make his data available.

References

Avio, K., and C. S. Clark (1976)
Property crime in Canada: An econometric study. Toronto: Ontario Economic Council.

Baldus, D. C., and J. W. L. Cole (1975)
"A comparison of the work of Thorsten Sellin and Isaac Ehrlich on the deterrent effect of capital punishment." *Yale Law Journal* 85:170–186.

Becker, G. S. (1968)
"Crime and punishment: An economic approach." *Journal of Political Economy* 78: 169–217.

Bielby, W. T., R. N. Hauser, and D. L. Featherman (1977)
"Response errors in nonblack males in models of the stratification process." *Journal of the American Statistical Association* 72:723–735.

Block, M. K., and J. M. Heineke (1975)
"A labor theoretic analysis of the criminal choice." *American Economic Review* 65:314–325.

Blumstein, A., and J. Cohen (1973)
"A theory of the stability of punishment." *Journal of Criminal Law and Criminology* 64:198–207.

Blumstein, A., J. Cohen, and D. Nagin (eds.) (1978)
Deterrence and incapacitation: Estimating the effects of criminal sanctions on crime rates. Washington, D.C.: National Academy of Sciences.

Bowers, W. J., and G. L. Pierce (1975)
"The illusion of deterrence: A critique of Isaac Ehrlich's research on capital punishment." *Yale Law Journal* 85:187–208.

Box, G. P., and D. Cox (1964)
"An analysis of transformations." *Journal of the Royal Statistical Society* (B)26:211–243.

Cook, P. (1977)
"Punishment and crime: A critique of current findings concerning the preventive effects of punishment." *Law and Contemporary Problems* 41:164–204.

Cook, R. D. (1977)
"Detection of influential observations in linear regression." *Technometrics* 19:15–18.

Daniel, C., and F. S. Wood (1971)
Fitting equations to data. New York: John Wiley and Sons.

Draper, N., and H. Smith (1966)
Applied regression analysis. New York: John Wiley and Sons.

Ehrlich, I. (1973)
"Participation in illegitimate activities: A theoretical and empirical investigation." *Journal of Political Economy* 81:521–565.

Ehrlich, I. (1975a)
"The deterrent effect of capital punishment: A question of life and death." *American Economic Review* 65:397–417.

Ehrlich, I. (1975b)
"Deterrence: Evidence and inference." *Yale Law Journal* 85:209–227.

Ehrlich, I. (1977a)
"The deterrent effect of capital punishment: Reply." *American Economic Review* 67:452–458.

Ehrlich, I. (1977b)
"Capital punishment and deterrence: Some further thoughts and additional evidence." *Journal of Political Economy* 85:741–788.

Ehrlich, I., and J. C. Gibbons (1977)
"On the measurement of the deterrent effect of capital punishment and the theory of deterrence." *Journal of Legal Studies* 6:35–50.

Ehrlich, I., and R. Marks (1978)
"Fear of deterrence." *Journal of Legal Studies* 7:293–316.

Fair, R. C. (1970)
"The estimation of simultaneous equation models with lagged endogenous variables and the first order serially correlated errors." *Econometrica* 38:507–516.

Fisher, F. M., and D. Nagin (1978)
"On the feasibility of identifying the crime function in a simultaneous model of crime rates and sanction levels." In A. Blumstein et al. (eds.), *Deterrence and incapacitation: Estimating the effects of criminal sanctions on crime rates*, Washington, D.C.: National Academy of Sciences.

Forst, B. E. (1976)
"Participation in illegitimate activities: Further empirical findings." *Policy Analysis* 2: 477–492.

Forst, B. E. (1977)
"The deterrent effect of capital punishment: A cross-state analysis of the 1960's." *Minnesota Law Review* 61:743–767.

Gilbert, J. P., R. J. Light, and F. Mosteller (1975)
"Assessing social innovation: An empirical base for policy." In A. R. Lunsdaine and C. A. Bennett (eds.), *Evaluation and experiment: Some critical issues in assessing social programs*, New York: Academic Press.

Hoenack, S. A., T. T. Kudrle, and D. L. Sjoquist (1978)
"The deterrent effect of capital punishment: A question of identification." *Policy Analysis* 4:491-527.

Hoenack, S. A., and W. C. Weiler (1977)
"A structural model of murder behavior and the criminal justice system." Unpublished manuscript.

Joreskog, K. G. (1970)
"A general method for analysis of covariance structures." *Biometrika* 57:239-251.

Klein, L. R., B. E. Forst, and V. Filatov (1978)
"The deterrent effect of capital punishment: An assessment of the estimates." In A. Blumstein et al. (eds.), *Deterrence and incapacitation: Estimating the effect of criminal sanctions on crime rates*, Washington, D.C.: National Academy of Sciences.

Kuh, E., and R. E. Welsch (1976)
"The variances of regression coefficient estimates using aggregate data." *Econometrica* 44:353-363.

Loftin, C. (1980)
"Alternative estimates of the impact of certainty and severity of punishment on levels of homicide in American states." This volume.

Miller, R. G., Jr. (1966)
Simultaneous statistical inference. New York: McGraw-Hill.

Nagin, D. (1977)
"Crime rates, sanction levels, and constraints on prison population." Unpublished manuscript.

Passell, P. (1975)
"The deterrent effect of the death penalty: A statistical test." *Stanford Law Review* 28:61-80.

Passell, P., and J. B. Taylor (1977)
"The deterrent effect of capital punishment: Another view." *American Economic Review* 67:445-451.

Penick, B. K. E., and M. E. B. Owens (eds.) (1976)
Surveying crime. Washington, D.C.: National Academy of Sciences.

Petersilia, J., P. W. Greenwood, and M. Lavin (1978)
Criminal careers of habitual felons. Washington, D.C.: National Institute of Law Enforcement and Criminal Justice, Law Enforcement Assistance Administration.

Reiss, A., Jr. (1980)
"Understanding changes in crime rates." This volume.

Ross, H. L. (1973)
"Law, science, and accidents: The British Safety Act of 1967." *Journal of Legal Studies* 2:1-78.

Skogan, W. G. (1976)
"Crime and crime rates." In W. G. Skogan (ed.), *Sample surveys of the victims of crime*, Cambridge, Mass.: Ballinger.

Sprent, P. (1969)
Models in regression and related topics. London: Methuen.

U.S. Bureau of the Census (1972)
Statistical abstract of the United States (93rd edition). Washington, D.C.: U.S. Government Printing Office.

U.S. Bureau of the Census (1975)
Statistical abstract of the United States (96th edition). Washington, D.C.: U.S. Government Printing Office.

Vandaele, W. (1978)
"Participation in illegitimate activities: I. Ehrlich revisited." In A. Blumstein et al. (eds.), *Deterrence and incapacitation: Estimating the effects of criminal sanctions on crime rates*, Washington, D.C.: National Academy of Sciences.

Wilson, J. Q., and B. Boland (1976)
"Crime." In W. Gorham and N. Glazer (eds.), *The urban predicament*, Washington, D.C.: The Urban Institute.

Witte, A. D. (1977)
"Estimating the economic model of crime with individual data." Working Paper 77-6. Chapel Hill, N.C.: Department of Economics, University of North Carolina.

Zellner A. (1962)
"An efficient method of estimating seemingly unrelated regressions and tests for a regression bias." *Journal of the American Statistical Association* 57:348-368.

Zellner, A. (1971)
An introduction to Bayesian inference in econometrics. New York: John Wiley and Sons.

Zimring, F. E. (1978)
"Policy experiments in general deterrence: 1970-1975." In A. Blumstein et al. (eds.), *Deterrence and incapacitation: Estimating the effects of criminal sanctions on crime rates*, Washington, D.C.: National Academy of Sciences.

Zimring, F. E., and G. Hawkins (1973)
Deterrence: The legal threat in crime control. Chicago: University of Chicago Press.

A prolegomenon for a macro model for criminal justice planning: JUSSIM III

ALFRED BLUMSTEIN
Urban Systems Institute
Carnegie-Mellon University

GARY C. KOCH
Department of Biostatistics
University of North Carolina

JUSSIM I is a deterministic, steady-state, nonqueueing flow simulation of a multistage system like the criminal justice system. The flow is processed through stages, and the process is characterized by an input into the first stage and by stage-to-stage branching ratios. At each stage, resources are applied to the units of flow, resulting in linear costs and resource consumption. JUSSIM II expands on JUSSIM I by incorporating the feedback flow of recidivists through subsequent arrests. In this paper, an expanded "JUSSIM III" is outlined. JUSSIM III incorporates a crime-generation process and a victim-generation process from the characteristics of offenders and victims. Offenses are generated by first-time offenders coming from the general population and recidivists released by the criminal justice system. Similarly, the population produces first-time victims from a potential victim population, and "victim recidivism" provides an opportunity for further victimization, which could be influenced by the victim's self-protective actions. The offense and victimization processes interact in a "victimization events" stage. The paper considers the various sources of data which could be used to generate the parameters of such a model and the statistical approaches to identifying the important parameters and for developing estimates for them.

I. Background and need for the model

The principal computerized model in current use in the criminal justice planning process is the JUSSIM I model (see Belkin et al. 1971) which examines the "downstream" flow through the criminal justice system beginning with crimes and following the handling of such crimes as they become associated with suspects, defendants, convicted offenders, and prisoners.

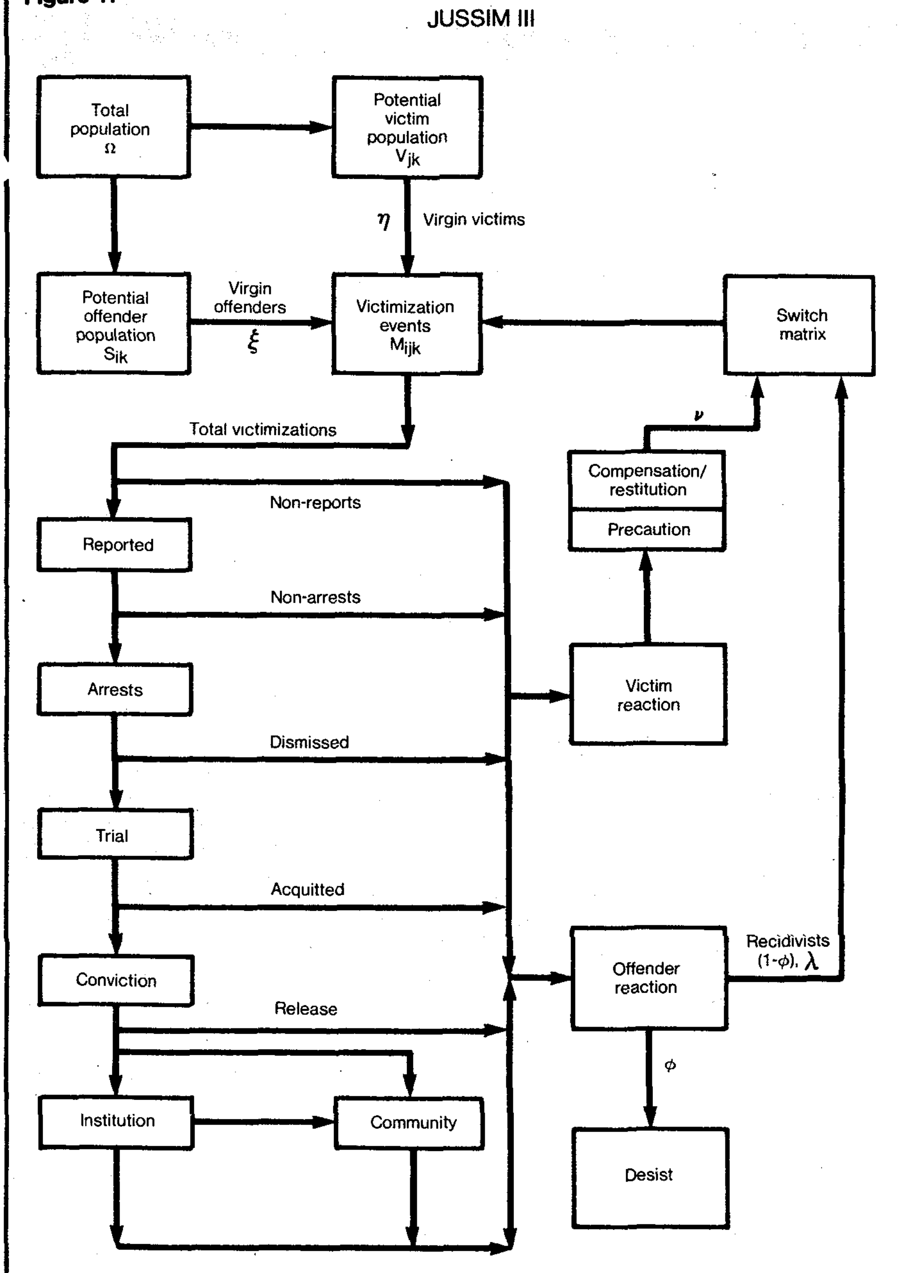
JUSSIM II (see Belkin et al. 1974) incorporates recidivism into that model through the addition of the feedback features associated with re-arrests. These features include probability of re-arrest at various points of departure from the criminal justice system, the time lags until recidivism, a crime-type-switch process reflecting the transformation from a previous crime type to a subsequent crime type, and a distinction between virgin arrestees (those arrested for the first time) and recidivist arrestees.

The JUSSIM I model has seen fairly widespread implementation (as discussed in Cohen et al. 1973 and Blumstein 1975), but very little actual use has yet been made of the JUSSIM II model. This is true largely because of the limited availability of data on recidivism, and to a lesser degree, because the intelligent use of this more complex and advanced model demands greater technical sophistication. JUSSIM II is, however, now ready for implementation in a number of jurisdictions, and so it is important to begin consideration of the future generation of such planning models. In this paper, we expand the view of the existing JUSSIM models to introduce the following additional considerations:

(1) Explicit concern for the demographic and socioeconomic characteristics of the population generating criminality.
(2) Explicit concern for the demographic and socioeconomic characteristics of the population generating victims.
(3) Explicit concern for the effect of the criminal justice system on crime reduction through incapacitation and deterrence.
(4) The effects on crime of environmental factors such as the state of the economy or social conditions.

Even though the information on most of these relationships is still extremely limited, there are nevertheless important reasons for pursuing the development of a model that incorporates them. Such a model forces an identification of the relevant variables in those relationships, and so iterates with the data collection process to assure that the information to be collected is consistent with the formulation of that model. Alternatively, as the data arrive from the

Figure 1.



various data collection programs, they provide a stimulus for reshaping the formulation of the model. This interaction between the data and the model represents an important contribution of any such model formulation.

The need for developing models is particularly important today in view of the major data collection efforts being undertaken. These include:

(1) The Offender-Based Transaction Statistics (OBTS) system, which tracks individual crimes and arrestees in their processing through the criminal justice system.

(2) The National Crime Panel's victimization survey, which collects a wide variety of personal and crime-experience data from a sample of households and businesses.

(3) The increasingly computerized criminal-history records (or "rap sheets") such as the FBI's Computerized Criminal History (CCH) file which records for each individual the sequence of his arrests and any relevant dispositional information that subsequently becomes available for each arrest.

In addition, a variety of computerized systems are being developed for handling the specific functions of various subsystems of the CJS. For example, the PROMIS system, designed to aid prosecutors, maintains detailed information on defendants as they are handled through the prosecutory processes.

All of these data systems were devised and are used for reasons other than the formulation and use of a planning model, and most of them are relatively unresponsive to the information demands of such a model. Nevertheless, the data they do provide can often be transformed into appropriate inputs for a planning model. In addition, the planning model can indicate variables central to planning but for which no appropriate data collection system has yet been organized. Those variables, once identified, might then be appended to existing collection efforts or, in some cases, might warrant an entirely new data collection effort.

Even though the principal value of such a model at this time is in shaping data collection and manipulation and in the formulation of functional relationships, it could ultimately become an important policy instrument when it is developed and provided with appropriate and valid data. That role, however, will continue to be limited by the validity of the assumed relationships incorporated into the model. As the forms of those relationships are revised, further modifications of the models will become necessary.

II. Basic structure of the model

A. Flow diagram. The basic structure of the model is shown in Figure 1, which depicts the flow of virgin offenders from a "criminal population" and of virgin victims from a "victim population" (both of which, of course, are drawn from the same total population) into a "victimization events" stage where the offender and the victim interact with each other or with other recidivist victims and/or offenders in the generation of victimization events. Subsequent to that victimization, the event is either reported or unreported and, if reported, an offender may or may not be arrested. If arrest occurs, then the suspect is handled through the criminal justice system; he may be dropped out at any one of a number of stages of successive penetration through the system; and then he may either recidivate or desist from future criminal activity.

The victimization process has a similar "recidivism" loop in which a victim may subsequently engage in various forms of preventive action reducing his vulnerability to victimization, and he may or may not subsequently become a victim again. In both the offender and the victimization recidivism processes, there is a "crime-switch" process reflecting the possibility that a subsequent appearance involves a different crime type than the previous appearance.

This structure deals with all the functional aspects explicitly incorporated into the JUSSIM I and JUSSIM II models, particularly the considerations of flow, resource consumption, and CJS workloads associated with JUSSIM I and the rehabilitation/recidivism aspects of JUSSIM II. In addition, however, it introduces a number of new considerations. First, it goes back from the crime or arrest stage to identify the relevant criminal population and the relevant victim population in terms of their demographic and socioeconomic characteristics. In addition, it incorporates the effects of the various "environmental factors" (e.g., economic conditions, social conditions) as considerations in their criminality or victimization vulnerability. Otherwise, the model attempts to reflect the deterrent effects of actions within the criminal justice system on the criminality of the relevant potential criminal population as well on the recidivism characteristics. The following discussion of each of these aspects of the model indicates how these representations would be accomplished.

B. Identification of parameter groups. In this section, the major groups of variables are identified, and their individual structures, their relationships to each other, and the effects of exogenous variables are explored.

(1) Potential offender population (s_{ik}). The potential offender population is created by exhaustively partitioning the total population (Ω) into subpopulations indexed by a subscript $i = 1, 2, \ldots, N_O$. The partitioning is based on demographic variables such as age, race, sex, and marital status, as well as socioeconomic variables such as education, income, and employment status. The partitioning is intended to divide the total population into subpopulations which have reasonably homogeneous crime-propensity rates (ξ_{ik}) which can then be measured for each crime type, k.

(2) Victim population (V_{jk}). Similarly, the total population is partitioned into victim subpopulations which are indexed by subscripts $j = 1, 2, \ldots, N_V$. This partition is likely to be different from the partition into which the offender population was structured, for it represents groups with internally homogeneous *victimization* risks. This might reflect their "value" as victims of crime as well as their vulnerability to potential offenders. The structure might include business or commercial enterprises as well as individuals. Here also, a victimization rate (η_{jk}) would reflect the rate at which a member of the j^{th} victim subpopulation becomes a victim of crime type k.

(3) Victimization events (M_{ijk}). The central feature of the model is the stage labeled "victimization events." This stage involves the convergence of the offenders who come initially from the "potential offenders" population and the victims who come initially from the "potential victims" population. Both of these "potential" populations first generate first-time or "virgin" offenders and virgin victims, many of whom reappear subsequently as recidivist offenders and recidivist victims. A victimization event involves the joint interaction of an offender of type i with a victim of type j in a crime of type k, and so the victimization events are characterized by rate parameters M_{ijk}, which would involve virgin or recidivist offenders and virgin or recidivist victims (i.e., the four combinations of potential victimization activity).

The M_{ijk} entries could be absolute quantitative rates reflecting the respective rates of victimization, or they could be more general functions reflecting the way those rates vary with other exogenous variables. These could include contemporary economic conditions, factors in the social environment, police deterrent activity like preventive patrol, or CJS sanction variables. These functions could be as large and as elaborate as could validly be built from the available data and statistical evidence.

(4) CJS resource consumption. As is seen in Figure 1, the flow of reported crimes and arrestees through the criminal justice system is structured similarly to JUSSIM I, and the JUSSIM I structure could be used for examining the resource use associated with this model. The principal extension would involve the incorporation of the victim and offender attributes.

(5) Rehabilitation (Φ_{ik}, λ_{ik}). The offender feedback loop used here is very similar to that associated with JUSSIM II; the feedback here, however, is to the victimization event rather than to re-arrest as in JUSSIM II. The feedback is characterized by parameters of recidivism probability or (the complement) desistance probability (Φ_{ik}). These could be introduced as average values, or, more richly, as functional forms of the exogenous environment in a way similar to that discussed previously for victimization events. Effort would be directed towards identifying the functional relationships of recidivism probability ($1-\Phi_{ik}$) and the mean time until recidivism (or its reciprocal, the mean recidivist crime rate (λ_{ik}) for the offender subpopulation groups. More elaborate functional structures could be introduced to account for the contemporary socioeconomic environment, the criminal justice system's deterrent activity, the nature of the rehabilitation treatment offered, and other factors that might reflect the individual's prior criminal history and his last treatment by the CJS.

(6) Victim vulnerability (ν_{jk}). Just as the model displays a feedback process for criminal offenders, it also includes a feedback loop for victims, with their associated "recidivism" rates, ν_{ik}. This rate can be affected by factors related to the compensation of victims or to restitution by the offender or by various types of protective action taken after a victimization event in order to reduce further victimization.

(7) Offender-victim-crime-switch process (T). In each cycle through the recidivism process, it is possible for each dimension of the previous path to be transformed.

First, there is the potential of straightforward transformation of crime type (a burglar switches to larceny or a victim of robbery has his automobile stolen the next time). There can also be transformation of both the offender's and the victim's subgroup, either through internal processes, such as aging, or by exogenous effects, such as change in socioeconomic status through change in economic and employment conditions. A large switch matrix, T, is included to reflect these transformations.

(8) Incapacitation effects. Changes in incapacitation strategy in the criminal justice system would be reflected through branching ratios associated with incarceration (e.g., reflecting the increased use of prison) or through the effect of longer sentences on the longer observed time between criminal events, with the associated reduction in the individual crime rate (λ_{ik}) for persons routed through prison.

(9) General-deterrence effects. The general-deterrent effect of sanctions used in the criminal justice system (e.g., apprehension probability, conviction probability, probability of imprisonment and sentence) would be reflected through the functional relationships among the sanction levels and the virgin criminality (ξ), the desistance probability (ϕ), and the recidivist crime rate (λ_{ik}). As these relationships develop from the deterrence literature, they could be incorporated directly into estimating these parameters.

(10) Environmental factors. It is well known that a wide variety of factors in the socioeconomic environment are highly correlated with crime rates, both cross-sectionally and longitudinally. It would be extremely desirable to incorporate the effect of these factors into the victimization-event functional relationships to reflect the issues surrounding the question of the "causes of crime." These factors would be reflected in a manner similar to the deterrence variables, i.e., through the desistance and crime rates associated with the recidivist population. If they could be adequately identified, these would provide important policy bases for crime-control actions that go beyond the confines of the criminal justice system. As with the other relationships in the model, the problems relate to the difficulty of estimating these relationships, particularly in an identified causal form.

III. Structure of the model parameters

The basic variables in the model include the following:

(1) Criminal population, their characteristics, and their criminality.

(2) Recidivist population and their recidivism probabilities and crime rate.

(3) Victimization population and their characteristics.

(4) Victimization event rates reflected in the M matrix.

(5) Offender-victim-crime switch process.

(6) Branching ratios of the criminal justice system.

(7) Resource costs and workloads as used in the JUSSIM I model.

In this section, we examine these various groups of variables and the structure associated with estimating these variables from available data. In many cases, we try to indicate our current best judgment of the appropriate data elements in estimating these variables, but these will of course depend on the degree to which data on the indicated variables are available and the degree to which they do indeed serve to estimate the relevant variables, since they could well be augmented by other elements or replaced by other more efficient estimating variables. Thus, the specifics of the variables identified serve more as illustrations than as the ultimate definitive set that will eventually be used.

A. Offender population. The offender population is characterized by a set of attributes reflecting the demographic and socioeconomic characteristics along which they are best partitioned to generate subgroups with homogeneous offending rates. The presumed relevant attributes include the following:

(1) Demographic variables—age, race, sex, marital status.

(2) Socioeconomic status variables—education level and income level.

The demographic variables might include binary partitions of race and sex, a three-way split for marital status (never married, currently married, and previously married but not currently), and some type of grouping for the age variable. For the socioeconomic status variables, an initial estimate might include for education (less than 8 years, 8 to 12 years, 12 years, and more than 12 years) and for income (less than $5,000 per year, $5,000–$10,000 per year, $10,000–$20,000 per year, and more than $20,000 per year). In addition, there would be a variable reflecting the individual's employability or skill level.

This potential criminal population would give rise to an input of virgin criminals V_{ik}, reflecting the number of offenders of the i^{th} potential offender group committing crimes of type k.

B. Offender recidivist population. The recidivist population is similarly characterized by a pair of variables, ϕ_{ik} and λ_{ik}, where ϕ_{ik} is the probability of desistance of a person of type i who last committed a crime of type k (where desistance in this case implies complete cessation of crime committing behavior). In addition, those who *do* continue to commit crimes do so at a rate λ_{ik}, reflecting the reciprocal of the time until the next crime by a person of type i who has just previously committed a crime of type k.

The attributes associated with the recidivist population could include the same ones characterizing the criminal population, augmented by the individual's prior criminal record (the number of prior arrests, convictions, sentences) and by the treatment last given him by the criminal justice system (failure to arrest, arrest but no charge, acquittal at trial, conviction but released into community supervision, or institutionalization). In any of these, any level of detail appropriate to alternative specific treatment programs could be incorporated into such a model.

C. Victim population. The structure of the victimization population would be developed in a manner similar to that of the offender population. The dominant demographic, social, and economic variables would include age, race, sex, marital status, education level, and income. It may well turn out, however, that an examination of the detailed data will suggest a different structuring of the individual variables for the victimization population than for the offender population, even though the basic variables would be the same.

In addition, the victimization population would also be characterized by some measure of exposure. This measure would include a combination of considerations of assets or vulnerability to victimization as well as considerations of self-protective actions taken to reduce the risk of victimization.

D. Rate of victimization events. The rate of victimization events is indicated by the entries in the victimization matrix, M_{ijk}, the rate at which a victim of class j is victimized by an offender of class i for a crime of class k. The structure of the matrix is dictated by the structure of the offender and victim populations. The entries would be the rate of aggregate victimizations by aggregate offenders (combining the virgins and recidivists of both offender and victim groups). The entries in the matrix would be either a simple scalar value for the rate of victimization events or, more generally, a function of current environmental factors, the deterrent effects of current criminal justice sanctions, and the effect of such police practices as preventive patrol.

Examples of the relevant environmental factors include:

(1) Population density as measured by people per acre in the district being studied, such as a census tract.

(2) A measure of the housing condition in the district, as measured, for example, by the persons per room or the measured state of dilapidation of the housing in the area.

(3) A measure of the migratory mobility of the population in the area as measured by the percent of population resident in the area for less than 2 years.

(4) A measure of the state of unemployment in the area as measured by the percent employed, the percent seeking employment but unemployed, and the percent eligible for employment but no longer seeking employment.

(5) A measure of the state of family disintegration in the region as reflected, say, in the percent of one-parent families.

(6) A measure of the differences across regions, as measured by a variety of regional indicator variables.

The deterrence variables would be reflected through additional functional relationships on the M matrix. These relationships would include variables such as the probability of arrest given a crime, the probability of conviction given an arrest, the probability of imprisonment given conviction, and the mean time served for those imprisoned.

The police crime prevention activities might include the intensity of preventive patrol (as measured, for example, by the mean patrol passage time for a random point in the district), by the percentage of unmarked patrol activity, as well as by the rate of various forms of police crime-prevention activity, such as the use of family crisis intervention units.

To the extent that data do not permit determination of the M_{ijk} functions in terms of environmental factors and deterrence factors, then these factors should be brought into the relationship for determination of the virgin crime-propensity from the potential criminal population (ξ), and for the determination of the recidivism parameters, λ and ϕ, associated with the recidivist offenders. It is probable, however, that the relationship will be more easily determined for the victimization rates than for the offenders, because most criminal records do not make an adequate distinction between virgin and recidivist offenders. Therefore, the relationship may well be more easily applied directly through the M_{ijk} matrix than partitioned into separate relationships for virgins and recidivists, with the recidivists' relationship partitioned between the desist probability (ϕ) and the associated crime rate, λ.

E. Victim recidivism. Recidivism is built in for victims just as it is for offenders. The parameters for the victimization recidivism process, identified as victimization at a rate ν_{jk} would be a parameter very similar to λ_{ik} for the offender recidivists. The victimization-recidivism function would be associated with each victimization population group j, and would be designed to incorporate consideration of the victims' prior history, just as offenders' prior history is taken into account in estimating the recidivism parameters. In addition, ν_{jk} would take into account the victims' reactions to their prior victimization experiences and to the compensation or restitution environment. For example, if compensation arrangements tend to reduce the incentives for self-protective action, then an environment that provides victim compensation would be expected to stimulate a higher victimization-recidivism risk. If prior victimization history stimulates self-protective action such as target-hardening or escape, then the victimization-recidivism rate would be reduced correspondingly.

F. Offender-victim-crime switch. JUSSIM II calls for a crime-switch matrix reflecting the transformation from a prior crime type to a subsequent one by recidivists; JUSSIM III will require a more elaborate switch process. First, it must incorporate the offender's crime-type switch as in JUSSIM II. It will also require transformations of both the offender class and the victim class on subsequent recidivism. Thus, an offender would transform from type i to i', a victim from type j to j', and the crime type from type k to k'. For example, an offender might age, change his education level (perhaps because of an educational component of his treatment program) increase his earnings, or change his marital status between successive incidents. The switch matrix would therefore reflect those rates of state transition. For victims, the transition might reflect similar changes in demographic structure or changes in exposure reflecting various actions taken for self protection. The matrix would be a square matrix with rank equal to the product of the number of offender types by the number of victim types by the number of crime types on each of its dimensions.

G. Branching ratios in the criminal justice system. For the flow through the criminal justice system, JUSSIM III requires branching ratios similar to those associated with JUSSIM I and JUSSIM II. These branching ratios, however, must now reflect the much richer "characteristic" or "crime-type" structure of the JUSSIM III model. Thus, the branching ratios would depend in general on the characteristics of the offender, the characteristics of the victim, and the crime type. This provides the opportunity for reflecting the differential treatment given to offenders or defendants based on their different demographic or socioeconomic attributes. It also provides the opportunity for reflecting the influence of a victim's characteristics on the treatment accorded his attackers (e.g., reflecting a difference in the courts' responses to intergroup events rather than to intragroup ones).

H. Resources, costs, and workloads. The JUSSIM III model would include data on the resource costs and workloads just as they are incorporated in the basic JUSSIM I format. If the parameters are independent of the offender or victim class, their values would simply be replicated for that segment of the workload vector. If the workloads are sensitive to either class, then that relationship would be reflected in richer detail within the workload vector. The resource costs would depend only on the resource and would be independent of any of the crime type, victim, or offender characteristics, as in the previous versions of JUSSIM.

IV. Statistical considerations

The statistical issues underlying the development of the JUSSIM III model pertain to the formulation of valid methodological strategies for estimating the functional relationships which characterize the respective model components. For this purpose, attention must be directed to the following four basic problems:

(1) The specification of the structure of the data array required for the estimation of a particular functional relationship and its corresponding parameters.

(2) The specification of methods of statistical analysis for parameter estimation from appropriate data arrays.

(3) The specification of available data sources and their possible limitations as a basis for parameter estimation.

(4) The specification of the potential need for new data sources in terms of alternative sampling and measurement procedures.

The remainder of this discussion is concerned with the implications of these considerations for the JUSSIM III model.

A. Specification of data array structure for parameter estimation. Since the JUSSIM III model is concerned with the rates of occurrence of discrete events (victimizations, arrests, trials, etc.), the data arrays required to estimate the various functional relationships in its structure are multidimensional contingency tables which link the frequencies of such events to the characteristics of the corresponding victims, offenders, and environmental subpopulations. In this framework, the dependent variables would pertain to the outcome (or response) status for exposure units at risk for such phenomena as:

(i) Whether a person of a particular type *experiences any victimizations* during a particular time period (the most recent day, week, month, year).

(ii) Whether a person of a particular type *commits any victimizations* during a particular time period.

(iii) Whether a victimization event is *reported* to police by either the involved persons or witnesses.

(iv) Whether a reported victimization event ultimately leads to an *arrest.*

(v) Whether an arrested person is ultimately sent to *prison* or experiences *other sanctions.*

(vi) Whether an arrested person commits a *subsequent victimization* (and/or is *re-arrested*) or not during a particular time period after release from the criminal justice system (either prior to trial, at the end of trial, or at the end of prison sentence or other sanction).

(vii) Whether a victimized person *experiences a subsequent victimization* or not during a particular time period after the previous victimization.

The independent variables would reflect the specific nature of such exposure units as:

(a) The demographic, socioeconomic, and vulnerability characteristics of potential victims.

(b) The demographic, socioeconomic, and opportunity characteristics of potential offenders.

(c) The crime type and the environmental characteristics for reported victimizations together with any available information for the involved victim and offender as in (a) and (b).

(d) The demographic, socioeconomic, and previous criminal record characteristics of an arrested person together with corresponding information for the involved victim as in (a) and the crime type and environment as in (c).

(e) The demographic, socioeconomic, and criminal record characteristics of a person released from the criminal justice system (either prior to trial, at the end of trial, etc.) and their opportunity characteristics for committing subsequent victimizations.

(f) The demographic, socioeconomic, previous victimization history, and vulnerability characteristics of a victimized person.

Thus, the parameters in the JUSSIM III model can be estimated through the analysis of such contingency tables as:

(1) (i) vs. (a) to determine the structure $\{V_{jk}\}$ of the victim population and the virgin victimization rates $\{\eta_{jk}\}$.

(2) (ii) vs. (b) to determine the structure $\{S_{ik}\}$ of the offender population and virgin offender rates $\{\xi_{ik}\}$.

(3) (iii) vs. (a) and (iii) vs. (b) to determine the reporting rate structure for victimizations from the victim and offender perspectives.

(4) (iv) vs. (c) to determine the arrest rate structure for reported victimizations.

(5) (v) vs. (d) to determine the prison rate structure for arrests.

(6) (vi) vs. (e) to determine the recidivism rate structure $\{\lambda_{ik}\}$ and the desist rate structure $\{\phi_{ik}\}$.

(7) (vii) vs. (f) to determine the victimization recurrence rate structure $\{\nu_{jk}\}$.

Similarly, if simultaneous information for both the offender and victim can be obtained for individual victimization events and consecutive pairs of victimization events, appropriate contingency tables can also be formulated to determine the victimization rate structure $M\{_{ijk}\}$ and the crime switch structure $\{T_{ijk,i'j'k',i''j''k''}\}$ respectively. More realistically, however, such data would be very difficult, if not impossible, to obtain. For this reason, a more practical estimation strategy may be to apply synthetic estimation ("raking") procedures (as described in Appendix A) to adjust analogous arrest data like (d) to the corresponding separate marginal structures for victims and offenders like the $\{V_{jk}\}$ and $\{S_{ik}\}$. However, as indicated in the next section, such methods should be used cautiously because they presume that the joint victim-offender crime association structure for victimizations is the same as for arrests.

In summary, various types of multidimensional contingency tables can be constructed as a basis for estimating the parameters in the JUSSIM III model. Some of these tables may involve rather crude data while others may require highly sophisticated data. In either situation, the critical issue is the formulation of such estimation problems in terms of contingency tables which can then be manipulated via the methods of analysis in Appendix A to determine the functional relationships which characterize the model components. In any event, a considerable effort in data collection, analysis and assumption testing can be anticipated before valid and credible relationships are developed.

B. Specification of available data sources. The JUSSIM III model described in this paper is extremely elaborate in order to account for the diversity of victim types, offender types, and crime types. For this reason, obtaining data for the formal construction of the model will require extensive use of available data sources as well as the development of new data sources in the future. In this section, various existing data sources are discussed in terms of their applicability to the JUSSIM III model.

The FBI's *Uniform Crime Reports* (UCR) present basic data for police reported crime rates in different parts of the United States. The corresponding arrest data provide some information concerning offender demographic status as well as the environmental situations in which the respective crimes have occurred. On the other hand, the principal limitation of these data is that the arrest process can involve a biased selection of the offender population, some offenders being more arrest prone than others. Thus, estimates obtained from the UCR data for offender characteristics need to be contrasted with corresponding results from other sources of information like victimization surveys and offender self-report studies in order to evaluate the general magnitude of this bias. Such comparisons also provide a basis for the adjustment of competing estimators from alternative data sources to a common framework by synthetic estimation (raking) procedures as described in Appendix A. Other comparisons of UCR crime data with victimization data and offender self-report data can be used to estimate the extent to which various crime types are reported or not reported to the police.

The victimization surveys provide the primary sources of information for estimating functional relationships in the victimization process. In addition, victim recidivism information is potentially obtainable from the longitudinal aspects of the national survey, which obtains information repeatedly from panel members. The panel data could be used to construct estimates for the intervals between victimizations or the rate of victimization and the nature of the crime switch process for victims. All of this analysis, however, should be undertaken with caution because of the errors with which individuals recall victimization events, especially sequences of such events.

Where available, the Offender Based Transaction Statistics (OBTS) are particularly well suited as data sources for the branching ratios throughout the criminal justice system. Typically, these can be estimated for each of the offender demographic groups since the OBTS incorporate offender characteristics as part of the regular record. If the OBTS contain some limited victimization information, such data might provide preliminary estimates for the interactions in the joint offender-victim victimization rate structure.

Another relevant information source is the criminal history record or "rap sheet." These records typically include data on the age, race, sex, and possibly other demographic characteristics of the offenders, as well as the sequence of offenses for which they were arrested, the dates of arrest, and, when reported, the disposition of each arrest. Thus they provide a basis for the estimation of the recidivism behavior of offenders. The rap sheet data and the OBTS data are largely complementary. They also provide some opportunities for consistency checks, but they are not likely to be available for all offenders in all jurisdictions for a long time to come. Special adjustment procedures will be required to use such data in jurisdictions other than the few for which good data are available.

Other aspects of criminal history, like nonreported offenses or offenses for which no arrests were made, can be analyzed with offender self-reports. Such self-reports are unquestionably suspect *a priori*, but further study is required to determine the degree to which there is distortion in the self-reporting process (both suppression and elaboration) and to identify methods for adjusting for such distortions. In this context, the most important use of self-report information would be the estimation of association parameters for offender characteristics which could then be synthesized with analogous data from rap sheets and from victimization surveys.

In summary, these available data sources provide a basis for estimating in at least a preliminary way many of the parameters of the JUSSIM III model, even in the face of their respective limitations and even though none of them is available at the victim-offender-crime level of detail for victimization events. Moreover, a wide variety of special-purpose studies (e.g., the Kansas City Police Patrol Experiment 1974, the Wolfgang, Figlio, and Sellin 1972 birth cohort study, etc.) represent additional sources of information for estimating various parameters after appropriate adjustments have been made for their special character or the particular jurisdiction where they took place. Thus, although suitable caution must be exercised, existing data sources can be used to form preliminary estimates for various functional relationships in the JUSSIM III model, and those estimates will stimulate the collection of better data and the improved specification of relationships, eventually resulting in a reasonable and useful formulation.

C. Specification of new data sources. A primary motivation underlying the formulation of the JUSSIM III model is its potential usefulness as a planning instrument for criminal justice system policy decisions. However, this type of application is appropriate only if the estimated functional relationships in the model are based on valid data. Thus, the inherent weaknesses of the existing sources of information, like those described previously in Section IV.*B*, imply the need for new data collection systems having less selection bias in their sampling procedures and more accurate measurements. From a statistical point of view, each of the respective types of contingency tables (1)–(7) described in Section IV.*B*, together with the corresponding dependent variables (i)–(vii) and independent variables (a)–(f), should be considered in this light, each requiring detailed specifications to obtain data with the required validity. Thus, a realistic discussion of the development of the new information sources required warrants a separate research effort beyond the scope of this paper. Nevertheless, some brief remarks about the general requirements for such data collection efforts can be given to motivate those future investigations.

First of all, attention should be directed toward improving the quality of the data pertaining to the processing of arrests through the criminal justice system. A national sample of jurisdictions will be needed in order to obtain complete follow-up information on the ultimate disposition (e.g., dismissal before trial, acquittal after trial, probation after conviction, prison sentence of specific duration) for an appropriately defined parallel set of stratified subsamples of different types of arrests (e.g., all homicides, 30% of all other felonies, 10% of all misdemeanors, etc.). Moreover, such data should be collected in a manner which is not biased by the selected police departments and courts, does not interfere with their everyday operation, and encourages their cooperation and participation.

Table 1. Predicted values for prison rates from linear model contingency table analysis

Offense[1]	Prior arrests	Arrest Promptness[2]	Income[3]	Total Defendants	Defendants sentenced to prison	Predicted prison rate	Est. S.E.
NRB	One or more	S	L	29	15	.537	.050
NRB	One or more	S	HU	15	4	.304	.025
NRB	One or more	A	L	34	12	.304	.025
NRB	One or more	A	HU	31	11	.304	.025
NRB	None	S	L	12	7	.537	.050
NRB	None	S	HU	11	3	.304	.025
NRB	None	A	L	18	6	.304	.025
NRB	None	A	HU	14	1	.072	.013
RB	One or more	S	L	30	10	.304	.025
RB	One or more	S	HU	5	1	.304	.025
RB	One or more	A	L	51	15	.304	.025
RB	One or more	A	HU	36	4	.072	.013
RB	None	S	L	10	2	.304	.025
RB	None	S	HU	5	1	.304	.025
RB	None	A	L	18	1	.072	.013
RB	None	A	HU	20	1	.072	.013
OFML	One or more	S	L	66	15	.193	.032
OFML	One or more	S	HU	43	5	.072	.013
OFML	One or more	A	L	82	14	.193	.032
OFML	One or more	A	HU	56	3	.072	.013
OFML	None	S	L	26	2	.072	.013
OFML	None	S	HU	72	6	.072	.013
OFML	None	A	L	58	5	.072	.013
OFML	None	A	HU	56	3	.072	.013

[1] NRB = nonres. burg.; RB = res. burg.;
OFML = larceny (felony and misdemeanor)
[2] S = arrest same day as offense; A = later day
[3] L = low; HU = high or unclassified

Residual goodness of fit Q (D.F.=21) = 6.01
Total variation Q (D.F.=23) = 87.71
Percent unexplained variation = 7%

Secondly, a national sample similar to that described for arrests will be needed for information on the subsequent experience of convicted offenders who are released from the criminal justice system via probation, parole, or other means. This information source would also involve a multistage selection beginning with institutions, such as courts, full probation, halfway houses, and prisons, and ultimately focusing on released offenders. Various types of data would then be collected for the individuals in the sample by both active (interview) and passive (reports of subsequent arrests by the police or FBI) methods for a particular fixed future time period of perhaps as long as 10 years.

The formulation of new information sources about the arrest process, although perhaps prohibitively expensive in practice, is straightforward in principle. The questions pertaining to improving the quality of victimization data are considerably more difficult. The critical issue here is that victimizations are publicly perceived at an aggregate level (for instance, a 1-month time period for all persons who live in a specific census tract) as relatively prevalent events, but are experienced as very rare events by specific individuals for such isolated time periods as 1 week or 1 month. Thus, data collection systems like the victimization surveys which focus on the recent experience of selected individuals can provide only a limited amount of information about the reported victimizations themselves (ignoring temporarily the corresponding measurement error problems) because of the relatively small number who are involved even in very large samples. They do, however, provide useful information about different types of victimization rates in the exposed population at risk. Thus, another data source is needed for obtaining more extensive coverage of victimization events. One system which could be developed for this purpose would be a national system of victimization reporting centers whose locations were based on an appropriate sampling plan. The explicit mission of these centers would be to provide free counseling and legal advice in a confidential manner (i.e., without specific notification to the police or other official agencies) to both offenders and victims; and their implicit mission would be to obtain information about the corresponding victimization events and possibly previous events. Obviously, the principal disadvantage of this data source would be the bias associated with its reliance on volunteer reporting. Nevertheless, if the reporting rates for victimizations were high, then the effects of this bias might be reduced by using matching procedures to link such information to that obtained from police records and victimization surveys and applying multiple-record-system statistical estimation procedures like those described in Bishop, Fienberg, and Holland (1975, Chapter 6), Koch, El-Khorazaty, and Lewis (1976a), and Marks, Seltzer, and Krotki (1974).

Finally, other types of new information systems may be of interest with respect to certain specific aspects of victimization events. In this regard, a national sample of long-term cohort studies on virgin victimizations would provide useful data similar to that provided by Wolfgang, Figlio, and

Sellin (1972) on virgin offenses. Moreover, a national follow-up sample for suitably identified victims who are linked to arrests or other methods of reporting might obtain information about their subsequent victimization experience.

In summary, the formulation of a JUSSIM III model helps to identify several new sources of data for the measurement of crime and the activities of the criminal justice system. A broad range of possibilities exists for the design of such future information systems, but their specific development requires a substantially more comprehensive investigation. The JUSSIM III model sketched here represents an important instrument for generating an agenda of future research in the development of criminal justice statistics for use in the planning process.

Appendix A: Specification of methods of statistical analysis

The principal objective underlying statistical analyses for contingency tables such as (1)–(7) in Section IV.*A* pertains to the characterization of the relationship between dependent outcome variables such as (i)–(vii) and corresponding sets of independent variables such as (a)–(f). Attention is directed at identifying which of the independent variables account for the variation in the respective dependent variables together with the extent of the interaction among them. One approach for dealing with these questions is given in Clarke and Koch (1975), who considered the relationship between the probability of prison sentence and independent variables corresponding to the defendent's age, race, sex, income, employment, type of offense charged, prior arrest record, and arrest promptness (a proxy for strength of evidence against the defendant) for a sample of certain types of burglary and larceny arrests occurring in Mecklenburg County (Charlotte), North Carolina, during 1971. Their analysis consisted of two phases:

Phase I: A screening of the variables to select those responsible for the greatest amount of variation in prison rates among the subpopulations defined by different combinations of the independent variables.

Phase II: The fitting of a parsimonious model involving the independent variables selected as important during Phase I.

Phase I was conducted in the same spirit as stepwise multiple regression, but in a context appropriate for the categorical nature of the discrete variables under investigation. In this regard, certain Pearson χ^2-statistics divided by their respective degrees of freedom were used like the "F to enter" statistic in multiple regression as measures of relative importance for suitably eligible combinations of variables in a multivariate relationship. The first variable selected was the one having the largest (χ^2/d.f.) with respect to its first-order (two-way) relationship to prison outcome. Additional variables were selected by applying a similar selection rule using (χ^2/d.f.) computed for three-way, four-way, etc. contingency tables involving all of the previously selected independent variables successively with each of the eligible remaining variables vs. prison outcome. Phase I also included a procedure for terminating the selection process when the remaining variables were not statistically important. Two different types of statistics were used for this purpose:

(A) The Pearson χ^2-statistics for the partial association (from two-way contingency tables) of a specific eligible variable vs. prison outcome *summed* over all possible combinations of variables that have already been selected.

(B) A χ^2-statistic developed by Cochran (1954) and Mantel and Haenszel (1959), which combines information with respect to the effect of a specific eligible variable on prison outcome over all combinations of previously selected variables.

The statistic (A) reflects both the main effects of a specific variable and its interactions with previously selected variables. After the first few steps of the selection process, this statistic tends to lose its usefulness for two reasons. First, its degrees of freedom increase rapidly causing selection to become overly stringent. Second, the data become thinned to the extent that many of the cell frequencies in the multidimensional contingency table become smaller than 5 so that this criterion begins to lose its validity as a chi-square statistic. At this point, statistic (B) becomes useful because it combines information across all combinations of previously selected variables and is thus more resistant to the thinning problem. This statistic is highly sensitive with respect to detecting weak but consistent relationships for variables which have not yet been selected; however, it has the disadvantage that it reflects the "average" effects of a variable as opposed to its "total contribution," which includes interactions with other variables. For the most part, this difficulty should not often pose a major problem because statistic (A), which is used in the earlier stages, does pertain to the "total contribution" of a variable. In the later stages of the selection process where statistic (B) is used, the "average effects" of variables are the ones of primary interest because the interactions with other variables are likely to be of minor importance because the relationships to which they apply are generally weaker. In summary, statistic (A) is used to decide whether to terminate selection after the first two or three variables have been selected, and statistic (B) is used thereafter. In each case, the basis for terminating selection is failure to meet a significance level of $\alpha = 0.10$ or $\alpha = 0.05$, as in forward stepwise regression. Finally, in some situations, certain variables are known *a priori* to be of importance. These can be included either at the beginning of Phase I or at the end, depending on their potential causal role with regard to the dependent variable.

For the burglary and larceny arrest data considered by Clarke and Koch (1975), Phase I led to the selection of type of offense, income, prior arrest record, and arrest promptness, respectively. These four variables were then cross-tabulated with prison status to produce a five-dimensional contingency table. Phase II then involved the fitting of a model to the prison rates for all combinations of the selected independent variables. For this purpose, the weighted least squares methods described in Grizzle, Starmer, and Koch (1969) were used to formulate an effective additive linear model for characterizing the variation among prison rates by systematically removing unimportant sources of variation such as higher order interaction effects. Such effects are not retained unless they are significant at an appropriate level like $\alpha = 0.05$. In addition, the model is not considered to be satisfactory until a residual goodness of fit statistic Q becomes small (i.e., not significant at $\alpha = 0.25$), for otherwise not all of the important sources of variation have been identified. A final criterion for model effectiveness is a measure of unexplained variation, analogous to $(1 - R^2)$ in multiple regression, which is defined as the ratio of the goodness of fit statistic for the model to an analogous statistic for total variation among all the prison rates (i.e., the goodness of fit statistic for a model which implied that all the rates were equal).

Analyzing prison rate data in this way leads to an efficient description or smoothing of the data as shown in Table 1. These results indicate that the functional relationship between prison status and the independent variables which were included in the analysis can be summarized in terms of four distinct predicted values or clusters as follows:

Cluster 1. High prison rate = 53.7 percent *if* nonresidential burglary and low income and arrested same day as offense.

Cluster 2. Moderate prison rate = 30.4 percent *if* nonresidential burglary and (a) low income but not arrested same day, or (b) arrested same day but not low income, or (c) one or more arrests but not low income and not arrested same day, etc.

Cluster 3. Low prison rate = 19.3 percent *if* larceny and one or more arrests and low income.

Cluster 4. Very low prison rate = 7.2 percent *if* not in Clusters 1, 2, or 3.

These "predicted values" are different only if the corresponding observed values are significantly different; moreover, they represent better estimates (i.e., with smaller standard errors) of the prison outcome probabilities than the actual observed rates for the respective subpopulations corresponding to the cross-classified independent variables, because they are based on the entire set of data rather than on its component subsets. Thus, a clearer and more simplified framework is obtained for interpreting the relative effects of the important independent variables than would be possible from casual inspection of the full multidimensional contingency table.

An alternative framework for Phase II analysis is based on the use of log-linear models. These types of functional relationships are discussed extensively in Bishop, Fienberg, and Holland (1975), Goodman (1970, 1971), and Ku and Kullback (1974). They can often be applied to contingency tables such as the previously considered prison data by the Deming-Stephan Iterative Proportional Fitting procedure in conjunction with an appropriate set of marginal subtables. For this reason, the resulting estimated parameters have reasonably robust and stable statistical properties even for sparse contingency tables (with many small frequencies which are less than 5). Their major disadvantage is the possibly unclear interpretation of the log-linear model parameters, particularly when interaction is present in the relationships among the dependent and independent variables. Further discussion of the relative merits of linear vs. log-linear models is given in Bhapkar and Koch (1968a, 1968b). This subject is also considered in Koch et al. (1976b) in the context of a general methodological strategy for estimating log-linear model parameters, which is based on the simultaneous use of weighted least squares and Iterative Proportional Fitting or other related computational algorithms.

Another statistical method of potential use with respect to the JUSSIM III model is synthetic estimation (or "raking") for contingency tables. This analytical procedure permits an observed table corresponding to a sample from a specific population to be adjusted to yield estimators for other target populations, e.g., (1) local or regional subdivisions of a sampled national population; (2) other local or national populations which partially overlap a sampled local population. The assumptions required for the validity of this approach are:

(1) Certain marginal distributions for the target population, referred to as the "allocation structure," are known on the basis of census or other data.

(2) The higher order interactions across the subsets in (1), referred to as the "association structure," are the same for the target population as for the sampled population.

Given these considerations, the estimated table for the target population can be obtained by using the Deming-Stephan Iterative Proportional Fitting algorithm which preserves "association structure" as it successively adjusts the observed table to the components of the "allocation structure." Additional details pertaining to the statistical properties of this procedure are given in Bishop, Fienberg, and Holland (1975), Causey (1972), Ireland and Kullback (1968), and Freeman and Koch (1976). As indicated by Seidman (1975) in a more general context, such synthetic estimation methods should be used cautiously because of the potentially severe bias they may suffer when assumptions (1) or (2) are not justified or are incorrectly applied.

Finally, Koch et al. (1975) and Brock et al. (1975) discuss weighted least squares methods of analysis for contingency tables based on complex sample surveys such as those used to measure victimization. These procedures are directly analogous to those described in reference to the prison rate from Clarke and Koch (1975).

In summary, a broad range of statistical methods are available to estimate functional relationships for the components of the JUSSIM III model. Some of these are based on weighted least squares computations while others are based on the Deming-Stephan Iterative Proportional Fitting procedure. Thus, for any specific application, the critical issue is to use that method which is most appropriate to the estimation problem under consideration.

References

Belkin, Jacob, A. Blumstein, and W. Glass (1971)
"JUSSIM, an interactive computer program for analysis of criminal justice systems." Urban Systems Institute Report, School of Urban and Public Affairs, Carnegie-Mellon University, July 1971.

Belkin, Jacob, A. Blumstein, and W. Glass (1974)
"JUSSIM II, an interactive feedback model for criminal justice planning. Urban Systems Institute Report, School of Urban and Public Affairs, Carnegie-Mellon University, January 1974.

Bhapkar, V. P. and G. G. Koch (1968a)
"Hypotheses of 'no interaction' in multidimensional contingency tables." *Technometrics* 10:107–123.

Bhapkar, V. P., and G. G. Koch (1968b)
"On the hypotheses of 'no interaction' in contingency tables." *Biometrics* 24:567–594.

Bishop, Y. M. M., S. E. Fienberg, and P. W. Holland (1975)
Discrete multivariate analysis: Theory and practice. Cambridge, Mass.: M.I.T. Press.

Blumstein, Alfred (1975)
"A model to aid in planning for the total criminal justice system." In *Quantitative tools for criminal justice planning,* Washington, D.C.: U.S. Department of Justice, Law Enforcement Assistance Administration.

Brock, D. B., D. H. Freeman, Jr., J. L. Freeman, and G. G. Koch (1975)
"An application of categorical data analysis to the National Health Interview Survey." *Proceedings of the Social Statistics Section of the American Statistical Association.*

Causey, B. C. (1972)
"Sensitivity of raked contingency table totals to changes in problem conditions." *The Annals of Mathematical Statistics* 43:656–658.

Clarke, S. H., and G. G. Koch (1975)
"Who goes to prison? The likelihood of receiving an active sentence." *Popular Government* 41:25–37.

Cohen, Jacqueline, et al. (1973)
"Implementation of the JUSSIM model in a criminal justice planning agency." *Journal of Research in Crime and Delinquency,* June 1973.

Cochran, W. G. (1954)
"Some methods of strengthening the common X^2 test." *Biometrics* 10:417–451.

Freeman, D. H., Jr., and G. G. Koch (1976)
"The asymptotic covariance structure of estimated parameters from marginal adjustment (raking) of contingency tables." *1976 Proceedings of the Social Statistics Section of the American Statistical Association* 330–335.

Goodman, L. A. (1970)
"The multivariate analysis of qualitative data: Interactions among multiple classifications." *Journal of the American Statistical Association* 65:226–256.

Goodman, L. A. (1971)
"The analysis of multidimensional contingency tables; stepwise procedures and direct estimation methods for building models for multiple classifications." *Technometrics* 13: 33–61.

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

Ireland, C. T., and S. Kullback (1968)
"Minimum discrimination information estimation." *Biometrics* 24:707–713.

Kelling, G. L., T. Pate, D. Dieckman, and C. E. Brown (1974)
The Kansas City Preventive Patrol Experiment: A technical report. Washington, D.C.: Police Foundation.

Koch, G. G., D. H. Freeman, Jr., and J. L. Freeman (1975)
"Strategies in the multivariate analysis of data from complex surveys." *International Statistical Review* 43:59–78.

Koch, G. G., M. N. El-Khorazaty, and A. L. Lewis (1976a)
"The asymptotic covariance structure of log-linear model estimated parameters for the multiple recapture census." *Communications in Statistics* A-5:1425–1445.

Koch, G. G., D. H. Freeman, Jr., P. B. Imrey, and H. D. Tolley (1976b)
"The asymptotic covariance structure of estimated parameters from contingency table log-linear models." *Proceeding of 9th International Biometric Conference*, Cambridge, Mass., August 1976.

Ku, H. H., and S. Kullback (1974)
"Log-linear models in contingency table analysis." *The American Statistician* 28:115–122.

Mantel, N., and W. Haenszel (1959)
"Statistical aspects of the analysis of data from retrospective studies of disease." *Journal of the National Cancer Institute* 22:719–748.

Mark, E. S., W. Seltzer, and K. J. Krotki (1974)
Population growth estimation. New York: The Population Council.

Seidman, D. (1975)
"Simulation of public opinion: A caveat." *Public Opinion Quarterly* 39:331–342.

Uniform Crime Reports (1973)
Washington, D.C.: Federal Bureau of Investigation, U.S. Department of Justice.

Wolfgang, M. E., R. M. Figlio, and T. Sellin (1972)
Delinquency in a birth cohort. Chicago and London: University of Chicago Press.

Redesigning the Kansas City Preventive Patrol Experiment

STEPHEN E. FIENBERG
Department of Statistics
and Department of Social Science
Carnegie-Mellon University

KINLEY LARNTZ
Department of Applied Statistics
University of Minnesota

ALBERT J. REISS, JR.
Department of Sociology
Yale University

The most striking result of the Kansas City Preventive Patrol Experiment (KCPPE) was that no significant differences were found among the three types of beats ("reactive," "control," and "proactive") in reported crime, rates of victimization, level of citizen fear, or level of citizen satisfaction with the police. In this paper, we focus on several related shortcomings of the design and implementation of the KCPPE. These include (1) the orientation toward proving rather than disproving the null hypothesis of no differences among treatment effects, (2) the confusion of manpower differences and treatment (patrol strategy) differences, and (3) the lack of control of other variables affecting the outcome of the experiment and the lack of proper randomization in assigning treatments to beats. Following this discussion of the KCPPE, we outline the design of a class of randomized controlled field studies for patrol-related experiments that are adaptable to most large cities and that could have yielded considerably more information than the actual design used in Kansas City.

I. The Kansas City experiment

From October 1, 1972, through September 30, 1973, the Kansas City, Missouri, Police Department conducted an experiment designed to assess the impact of routine preventive patrol on the incidence of crime and the public's fear of crime. The experiment involved 3 variations in the level of preventive patrol used in 15 of the 25 beats comprising the Kansas City South Patrol Division. The 15 beats were blocked into 5 groups of 3 matched beats each, on the basis of reported crime counts, calls for police service, and demographic data. Within each triplet one beat, the "control," was to be patrolled by a single police car in the usual fashion whenever it was not answering calls. In a second beat in each triplet, routine preventive patrol was eliminated and officers were to respond only to calls for service. These were referred to as the "reactive" beats. Finally, in the third beat in each triplet, routine preventive patrol was increased to 2 or 3 times the usual level. These were referred to as the "proactive" beats. Complete details of the experimental design and its implementation can be found in Kelling et al. (1974).

The striking result of this Kansas City Preventive Patrol Experiment (KCPPE) was that there were no "significant" differences among the three types of beats in reported crime, rates of victimization, level of citizen fear, and level of citizen satisfaction with the police. There was simply no apparent effect that could be related to differences in patrol manpower allocation, deployment, and operation. Now this result appears to fly in the face of conventional wisdom that increases in police patrol should reduce the levels of street crimes such as robbery and automobile theft. Since one of the purposes of the experiment was "to determine whether the Kansas City Police Department could safely divert resources ordinarily allocated to routine preventive patrol to other, possibly more productive patrol strategies," the results would appear to have important policy implications regarding the use of various manpower allocation schemes. It is therefore incumbent upon us to reexamine the experimental setup and the results of the KCPPE with great care in order to determine:

(a) what explanatory model lay behind the KCPPE.

(b) what was measured and what was not measured.

(c) what the results really were.

(d) what aspects of the experimental design and its execution may have led directly to these results.

(e) whether the foregoing facts may in any way compromise the conclusions reached by Kelling and his coworkers and others commenting on the experiment.

(f) if the conclusions are compromised, whether there is a different experiment that an appropriate police department might conduct which would get around many of the problems associated with the KCPPE.

A major difficulty with the KCPPE and any consideration of it are the terms "prevention" and "preventive patrol." Basically, the problem is that "prevention" is a black box, and exactly what police officers do with time spent on what is called "preventive patrol" is difficult to determine. The KCPPE did not actually manipulate this elusive concept of preventive patrol. Rather, the experiment attempted to manipulate the amount of manpower and its deployment. We return to this issue in the following sections.

II. Experimenting with police manpower

Most social experiments, especially those involving the police, are not really experiments in the statistical sense, but are in fact innovations for which there are no direct means available for separating out or controlling the effects of a change in public policy from concomitant changes in related variables. For example, "Operation 25," a 4-month "experiment" conducted by the New York City Police Department in 1954, involved the more than doubling of police strength in Manhattan's 25th Precinct. Although reported street crime went down markedly over this 4-month period, no direct information was available for the true crime rates. Alternate explanations of the observed decrease were available, and no effort was made to assess the possible deleterious effects of the increased manpower on adjacent areas or to determine what the reported crime would have been had the changes in manpower not been instituted. Far better than "Operation 25" was the 1966 experiment in New York's 20th Precinct (see Press 1971, 1972), although it too had its problems. Other examples of innovations in police practice are readily available. In 1953, the Kansas City Police Department carried out a complete 24-hour changeover from two-man to one-man patrol cars, and the 24 patrol areas were reorganized into 41 substantially smaller beats (see Brannon 1956). No adequate evaluation of such a change was possible.

In light (or perhaps we should say shadow) of these earlier attempts to evaluate the effects of changes in police patrol practices, the KCPPE would appear to provide a model worthy of careful study. Patrick Murphy, President of the Police Foundation which supported the KCPPE, has been quoted as saying that the preventive patrol project "ranks among the very few major social experiments ever to be completed and was unique in that never before had there been an attempt to determine through such extensive scientific evaluation the value of visible police patrol." While the KCPPE does not quite measure up to the well-evaluated randomized controlled field trials of large-scale social action programs described by Gilbert, Light, and Mosteller (1975), it represents a major step forward in the design and strategy of police research. It is only because of the use of sophisticated statistical techniques and their extensive documentation in the Technical Report (Kelling et al. 1974) that we are able to assess very carefully the various aspects of design and analysis. The KCPPE has not only shown that a police department can successfully experiment with patrol strategies without disastrous consequences, but it has also provided a starting point for future experimentation in this area.

III. Shortcomings of the KCPPE

Our discussion of the design and analysis of the KCPPE here focuses on several interrelated issues:

(1) proving rather than disproving the null hypothesis.

(2) the difficulty in actually assessing the treatment differentiation.

(3) the absence of adequate measurement of treatment implementation.

(4) the confusing of manpower differences and treatment (patrol strategy) differences.

(5) the possible contamination of data in experimental beats by the patrolling strategies in adjacent beats (whether they be experimental or not).

(6) the lack of proper randomization in assigning treatments to beats.

A. *Proving the null hypothesis*

A primary purpose of carefully designed statistical experiments is to control for as many sources of variation as possible in order to maximize the probability of detecting differences in the effects of different treatments when these differences do in fact exist. Because the KCPPE was unable to detect any differences in outcome attributable to the differences in the level of preventive patrol, we must therefore ask whether this result is due to the true lack of treatment effect differences or simply the weakness of the experimental setup when in fact differences do exist. The framework of hypothesis testing adopted by the authors in their report is geared toward rejecting the null hypothesis (in this case that of no difference) in favor of some alternative, by giving heavy weight to the null hypothesis unless it is demonstrably false. In technical terms we fix the level of significance (the probability of falsely rejecting the null hypothesis) at some low value (e.g., $\alpha = 0.05$) and the probability of failing to detect the alternative hypothesis when it is true is determined. Unless the latter is also small, we end up weighting the experiment in favor of the null hypothesis. The danger of "proving" the null hypothesis by means of a weak experiment haunts much of the literature on social experimentation.

Because the no-difference outcome of the KCPPE has such important social and policy implications, it is important to ask whether the design chosen sufficiently controls for crucial sources of variability to allow for the detection of the effects of true treatment differences when they exist.

The following example, adapted from Kempthorne (1952), may help clarify the issues just raised. A soft drink manufacturer has developed a new procedure for producing its best-selling cola. The question to be resolved is whether or not people can detect a difference in taste between the colas produced by the new and old procedures. The manufacturer's statistician sets up the following experiment. A panel of $n = 6$ tasters is presented with 3 glasses of cola, 2 produced by the old process and 1 by the new. Each taster is asked to pick the one cola whose taste differs from the other two. The tasters work independently of one another, and the colas are presented in a random order to each taster. The null hypothesis, H_0, specified by the statistician is that "no differences are detectable," and when H_0 is true the probability of detecting the cola produced by the new process is $p = 1/3$.

There are 7 possible results of this experiment, indexed by the number of tasters who correctly picked the cola from the new process. Under H_0, the probabilities of observing these results are:

No. correct	*Probability*
6	1/729 = 0.001
5	12/729 = 0.016
4	60/729 = 0.082
3	160/729 = 0.219
2	240/729 = 0.329
1	192/729 = 0.263
0	64/729 = 0.088

If all 6 tasters correctly pick the cola from the new process then we would appear to have strong evidence that H_0 is false. Even if only 5 tasters are correct the evidence is strong, but if only 4 tasters are correct, we might have observed a result so extreme with probability $0.001 + 0.016 + 0.082 = 0.099 \cong 0.1$. This is not a very rare event and, even if a small difference between the colas exists, we would likely not take the result of 4 out of 6 as being indicative of such a difference. Thus we take 5 or 6 correct as evidence against H_0 and in favor of some alternative hypothesis.

In order to evaluate how good an experiment this really is, we need to ask what would happen if some alternative to H_0 is really true. Suppose, for example that $p =$

1/2, and the cola from the new process is more likely to be picked than is indicated by H_0. Then the probability of getting 5 or 6 correct is now $(1/64) + (6/64) = 0.109$. *Thus even though there is a difference* we would fail to detect it about 8 out of 9 times using our criterion of 5 or 6 correct. The probability 0.109 is referred to as the *power* of our test for the alternative $p = 1/2$, and in the present case it would appear that the experiment as planned will simply not be powerful enough to detect an alternative as big as $p = 2/3$ (for which the power is only 0.35). The experiment appears to be stacked in favor of the null hypothesis, and unless we are careful we might erroneously conclude that $p = 1/3$ even when it really is as large as $p = 2/3$. Note the importance here of explicitly stating what effects we would like to be able to detect if they are present.

Basically, there are three ways to make an experiment more sensitive to detecting reasonable alternatives to H_0:

(1) increase the sample size.

(2) reorganize the structure of the experiment.

(3) refine the experimental technique.

In the cola experiment we could employ all three of these methods, but in the KCPPE where there are severe limitations on how big we can make the sample size, we clearly must know how to reorganize and refine the experiment to avoid the trap of proving the null hypothesis. We address these issues in Section IV of this paper.

B. *Differences in treatments*

Directly related to the issue of "proving the null hypothesis" is the choice of treatment for the experiment. If there is little or no difference among treatments, then we can hardly expect to have much luck in discovering differences among the resulting effects. The treatment standard chosen as the "control" in the KCPPE was the current patrol strategy and level of patrol used in the individual beats involved in the experiment. Yet we don't know whether this "control" level of patrol differs markedly across beats, nor whether it is high, moderate, or even low relative to the levels of patrol that are feasible for a given beat area or for police patrol beats more generally.

Kelling et al. (1974) state the aim of the KCPPE as follows:

> In its ultimate design the project would be a rigorous and systematic attempt to test the outcomes of different patrol strategies and ultimately could lead to cost-benefit analyses of varied strategies to determine the most efficient methods of undertaking patrol. It was likewise felt that the preventive patrol experiment would help to maintain a climate of innovation and self-evaluation, not only on the part of the department as a whole but also among individual officers. The task force and the Police Foundation realized that since the effectiveness of routine preventive patrol was not self-evident and because the capacity to deal with crime is a central police function, the preventive patrol experiment would fill a real professional need heretofore not addressed by other police agencies.

The notion of preventive patrol in this description is really a "black box," and the problem for the experimenter is that it is hard to manipulate such a "black box." As Riecken and Boruch (1974) note:

> A specified treatment allows a much more powerful experimental test than does a "black box." Furthermore, specifying treatments helps to achieve comparability between different experimental sites... when an experimental treatment is repeated (replicated) it should be kept the same or deliberately and systematically altered, rather than being allowed to vary haphazardly.

What was actually manipulated in the KCPPE was some combination of changes in manpower, the deployment of manpower, and visibility. Each of these variables contributes in a partially undefined way to the level of activity in each experimental unit, and then the level of activity somehow affects the "black box" of prevention. Some of the experimental areas were to continue with the same manpower levels, the same deployment, and the same visibility as before. This was intended to keep constant the level of activity in these "control" areas. In a second group of areas the manpower was doubled (very roughly), but the deployment of manpower was to be no different from before. The resulting increase in activity leads to the label of "proactive" for these areas. In the third group of experimental areas, the level of manpower was left the same, but deployment was manipulated. Some of the activities associated with the "controls" were eliminated from these areas, and thus we label the activity as "reactive." These manipulations of treatment variables leave us with the three patrol strategies of the KCPPE.

The "reactive" and "proactive" patrol strategies appear on the surface to differ markedly from the control strategy, but is this in fact the case? Routine preventive patrol is carried out during an officer's noncommitted time, but only 60 percent of observed time in the experimental area turned out to actually be noncommitted, and only a fraction of the noncommitted time is used for activities that might be labelled preventive patrol. The uses of noncommitted time by officers do not seem to vary very much for officers assigned to the beats with the three different levels of preventive patrol. Since calls for service and other routine uses of committed time were treated in a similar fashion in all experimental beats, the differences among the levels of treatment are not nearly as great as they appear to be from the initial experimental description. Moreover, officers on reactive beats used their preventive patrol time either on the perimeter of their own beats or in patrolling proactive beats to compensate for manpower shortages. The amount of time spent on the beat perimeters in reactive areas is not given, but it should have been carefully assessed due to the shape and relatively small size of several of the beats. Finally, officers often cross beat boundaries in response to calls for service. Since there are more officers on patrol in the proactive beats, they are more likely to be free to cross into reactive or control beats in response to service calls than the other way around. Finally, as Davis and Knowles (1975) note, specialized units (helicopters, K-9 units, etc.) were deployed independently of the preventive patrol experiment at a level consistent with activity in the preceding year.

There is a set of incident types defined by Kelling et al. as part of routine preventive patrol: traffic violations, building checks, car checks, and pedestrian checks. These routine patrol incidents are presented by treatment group in Table III-5A of Kelling, et al. (1974, p. 47). We have recategorized these data to assess more easily treatment differences. If there were no differences in the levels of preventive patrol across beat type we would expect each of the three sets of treatment areas to have roughly 33.33 percent of the routine preventive patrol incidents. The 1972 data (collected from January through September) are reasonably compatible with this expectation, with percentages varying from a low of 28.97

percent to a high of 38.69 percent but with most near 33 percent. The 1973 data (also January through September), in which we would expect the proactive beats to have a considerably larger share of the incidents and the reactive beats a considerably smaller share, present a somewhat different picture. The proactive beats in 1973 had 43 percent of the car checks, 53 percent of the pedestrian checks, and close to 33 percent of both traffic violations and building checks as opposed to 38, 39, 35, and 34 percent, respectively, in these categories in 1972. On the other hand, the reactive beats in 1973 had 26 percent of the car checks, 24 percent of the pedestrian checks, 28 percent of the traffic violations, and 28 percent of the building checks. These should be compared with figures from 1972 of 33, 32, 32, and 37 percent, respectively. The control beats were much more stable for 1972 and 1973 except that pedestrian checks dropped from 29 to 24 percent and car checks increased from 29 to 37 percent. There do appear to be differences in the relative proportions of routine preventive patrol incidents across treatment areas, but the differences are not all that big. What is even more troublesome is that the number of patrol-initiated events that took place in the "reactive" areas actually *increased* from 1972 to 1973. Although the proportion of car checks out of the total number that took place in the reactive areas decreased from 33 to 26 percent, car checks still comprised 45 percent of all patrol-initiated activities in 1973 data, whereas they comprised less than 36 percent of the activities for the comparable period of time in 1972. Thus there is some question as to what the experimental conditions were and how they were maintained. Larson (1976) discusses this point in greater detail.

The figures in the preceding paragraph may tend to overstate the treatment differences. To see this it is important to note that, unlike many other cities where manpower is scheduled over the day according to anticipated police activities, Kansas City attempts to equalize its manpower over 24-hour periods by allocating the same number of officers to each of the three 8-hour watches. Thus most preventive patrol activity occurs in the midnight to 8 a.m. watch, and the least in the peak workload watch from 4 p.m. to midnight. When these activities are examined on a day by beat by watch basis we find that we are faced with a low productivity situation, where big treatment changes may be required to increase or decrease productivity. Since the average number of patrol-initiated activities is about one per day per beat per watch, how much of an effect can we expect from the supposed elimination of preventive patrol from the reactive areas and the doubling of manpower, which likely produces far less than two per day per beat per watch? In fact, the biggest source of variation in patrol-initiated activities may well be among watches within beat, rather than among treatment areas.

C. *Treatment implementation: What was measured?*

Kelling et al. (1974) seem to have abstracted from the literature on patrol strategies the assumption that visibility is the main factor underlying patrol variations and that visibility impact on the crime rate might be mediated by changes in "fear of apprehension on the part of criminals" and "police response time" (see page xvii). The extent to which differences in treatment led to differences in actual visibility is thus crucial. The KCPPE *used no direct measurement* of visibility anywhere in the study and, while some attempt was made to measure response time, the data were scanty and not very reliable. Kelling et al. (1974) note that there was no consequential difference in response time to calls for the three areas based on the KCPPE measurements, but Larson (1976), using a simple "back-of-the-envelope" operations research model, has shown that, for the experimental design used, the expected travel time for cars in reactive beats would not increase very much. Since the experimental units were beat areas but what was manipulated were beat cars, it is unfortunate that any potentially useful information on reaction time is recorded by car rather than by beat area.

Let us return to the issue of visibility. Crimes where offenders are in private places are not accessible to the police for prevention efforts, and the offenders who commit these crimes are not affected by the visibility of police cars. Only street crime may be affected and, even in the case of purse snatchings and auto theft, visibility can aid the offender committing a crime if the visibility is fairly predictable. Because there was no direct measurement of visibility we must try to infer the amount of visibility associated with each of the three "levels of patrol." Using another "back-of-the-envelope" model, Larson (1976) estimates that the average frequency of preventive patrol passings in the control areas in 0.164 passes per hour or roughly one every 6 hours. Thus he notes that preventive patrol as usual in the experimental zone is not very much preventive patrol. Even if doubling the number of cars in the proactive areas doubles the number of times a randomly chosen spot is passed in a fixed period of time, this is still only once every 3 hours. Larson notes that the range of preventive patrol coverages in the KCPPE is considerably less than that for many other U.S. cities on a typical day, thus calling into question how useful the Kansas City results would be for other cities.

In summary, there are few (if any) direct measures of treatment conditions in the KCPPE, and those measurements that are available tend to point toward minimal differences in treatment effects across the three types of treatment.

D. *Manipulating manpower versus patrol strategy*

The Kelling et al. report on the KCPPE continually speaks of experimenting with police patrol strategies, but as we noted earlier the experiment itself deals only with the manpower levels for the proactive treatment and with one particular strategy, i.e., reducing preventive patrol for the reactive treatment. But manipulating what officers actually do while on patrol may well have as great an effect on the commission of crimes in an area as the manpower level of the patrol. For example, "stop and frisk" patrol strategies clearly involve more intrusive activities than do routine patrol strategies. By manipulating *both* manpower and patrol strategy we might produce alternatives to routine preventive patrol at current manpower levels, the effects of which are detectable. It is, of course, much easier for us to reach this conclusion after examining the KCPPE and its results than it would have been before it was carried out.

Whenever manpower is highly inefficient, or has low productivity relative to what it might have, increasing the number of workers will have relatively little impact on reducing the levels of crime. If a police officer makes only one car check a day when he could easily make 10, then doubling the manpower (assuming a constant ratio of manpower to activity) would produce only 2 per day but changing the patrol strategy could produce 5 times that number. In short, when we are faced, as we are in Kansas City, with very little preventive patrol strategy other than riding and sitting around, we can't expect changes in manpower to produce very much. Thus leaving the ordinary conditions of productivity constant, we must ask whether changes in manpower will make much difference. To the extent that one can increase levels of output per man by changes in patrol strategy, one can measure the effect of the strategy. An interesting question is whether changes in strategy can produce more effect than changes in manpower. Up to some point, there can be enormous gain by changes in strategy, and similarly only to some given point—usually spoken of as a saturation point—can changes in manpower have an effect. Under many conditions, tripling, etc., manpower does not have the expected multiplicative effect.

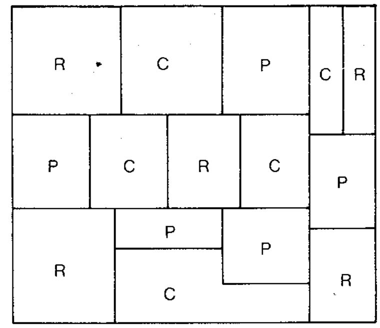
In an experiment like the KCPPE, whether we try to manipulate manpower, patrol strategy, or both should depend upon the level of effect expected per man for a given strategy relative to what might be achieved by increasing output per man. There is little reason, we believe, to increase police manpower when the current manpower is used so inefficiently relative to any strategy. The problem is complicated by the fact that events are not uniformly distributed in time so that manpower will have less effect at some times than others and ceiling effects are quickly reached. Between 2 and 4 a.m., there may be very few people to interrogate and very few cars to check in most beats of even a very large city. Thus strategy changes have different limits for different times. We frankly don't know how the experiment we propose in Section IV (or any other experiment) can take these difficulties into account, particularly because most strategies attempt to affect the offending populations, and we know so little about police interactions with such populations.

E. *Strategies in adjacent beats*

As we noted above, officers are dispatched across beat boundaries in answer to calls for service. With more manpower on routine preventive patrol in proactive areas, officers from proactive beats are much more likely to be able to answer calls in another beat than are officers from reactive or even control beats. Officers in one beat may in fact feel compelled to "help out" their colleagues in adjacent reactive beats and thus negate the intended effects of the experiment. Since controlling the dispatch of police cars is a near-impossible task, care must be taken in allocating treatments to beats to allow a measure of control relating to effective treatment differences in adjacent beats.

In this regard, we should reexamine the schematic representation of the 15-beat experimental area, reproduced here as Figure 1. The striking feature of the layout is the placement of the 5 reactive beats in the 4 corners and in the middle. Proactive and control beats intervene between every pair of reactive beats, and the *only* 4 beats with 2 sides adjacent to areas not in the experiment are all reactive. The probability that such an allocation of treatments to beats would happen by chance is quite small indeed.[1] In fact, as the authors of the KCPPE note in the Summary Report: "The geographical distribution of beats avoided clustering reactive beats together or at an unacceptable distance from proactive beats. Such clustering could have resulted in lowered [sic] response times in the reactive beats." The consequences of the geographical distribution are far more distressing, both because they indicate a lack of random assignment and because of the interactive effects described above. The arrangement of beats is of even greater concern once we realize the potential significance of a beat's being on the edge of the experimental area. The South Patrol Division is bordered at the west by Kansas City, Kansas (which is under different police jurisdiction), and on east and south by yet additional suburban Missouri police jurisdictions. Because police rarely if ever cross jurisdictional lines, let alone state lines, the geographical arrangement of beats in the KCPPE leads one to wonder how the results might have turned out had a different arrangement been used, especially one that controlled for what we might label "border" effects.

Figure 1.
Schematic representation of the 15-beat experimental area



P = Proactive
C = Control
R = Reactive

[1]In experimental situations such as the KCPPE there are several recognizable subsets of possible designs which for one reason or another may seem striking. While each of these designs might occur with small probability as a result of randomization, the probability of chosing a design coresponding to one of the recognizable subsets is substantially larger.

F. *Random assignment of treatments*

The layout of the experimental area and the discussions of it in Kelling et al. (1974) raise considerable doubt whether randomization was used at all in the assignment of treatments to beats. Is this a serious matter? We believe strongly that it is. The difficulties resulting from a nonrandomized field trial are many, but as Gilbert, Light, and Mosteller (1975) note, "The key problem is that the effect of the treatment is not distinctive in all cases, so that the treatment cannot be proved to have caused the effect of interest. Selection effects and variability of previous experience have often led to biases and misinterpretations."

Now it is true that by making certain assumptions regarding models and the nature of the experimental variability we can hope to surmount the obstacles posed by the lack of proper randomization in an experiment or controlled field trial. Yet this hope all too often remains just that, and the only sure way to be able to make causal statements after we complete an experiment is to randomize at some stage of the allocation of treatments to units. The nonrandomized controlled field trial can be an effective evaluation device, but primarily in the presence of what Gilbert, Light, and Mosteller call "slam-bang effects," hardly a phrase that can be used in describing the results of the KCPPE.

It may well be argued that a simple random assignment of treatments to beats in the KCPPE would have been inappropriate. But experimental designs that balance the allocation of treatments, and yet also involve some form of random assignment, are available.

IV. Designing a new patrol strategy experiment

It is possible to design a new experiment to examine alternative police patrol strategies in the spirit of the KCPPE that would overcome the pitfalls elaborated upon in the preceding section, although the task is not at all an easy one. The difficulty comes when we attempt to control for a variety of different sources of variability. Although such control in principle increases the efficiency of an experiment, we must retain sufficient degrees of freedom for the error term in order to keep the power against interesting alternative hypotheses high. Because the number of police beats in a given city is never very large (e.g., 69 beats in Kansas City, Missouri, and only 24 beats in Minneapolis, Minnesota), we need to strike a balance between the control of variability and the estimation of residual error.

In this section we describe an experimental design for a new patrol strategy field trial and several variants of added or diminished complexity. The basic design described involves 108 beats and thus cannot be implemented in any but the largest cities in the United States. By assuming the absence of all interaction effects we can reduce the required number of beats to 54. By utilizing natural boundaries such as the Missouri River and Missouri-Kansas state boundary line, we can cut these requirements to 70 beats for the full experiment and 28 for the reduced one (assuming completely regular beat sizes and shapes).

For those not familiar with the sophisticated details of the design of statistical experiments, the experiment described here may seem so complex that we have no hope of using it to detect interesting results that we might have discovered had we only tried to manipulate one factor and answer one question. The distinguished statistician, Sir R. A. Fisher (1926), once pointed out that he believed exactly the opposite to be true:

> In most experiments... the comparisons involving single factors... are of far higher interest and practical importance than the much more numerous possible comparisons involving several factors. This circumstance, through a process of reasoning... leads to the remarkable consequence that large and complex experiments have a much higher efficiency than simple ones. No aphorism is more frequently repeated in connection with field trials, than that we must ask Nature few questions, or ideally, one question, at a time. The writer is convinced that this view is wholly mistaken. Nature, he suggests, will best respond to a logical and carefully thought out questionnaire; indeed, if we ask her a single question, she will often refuse to answer until some other topic has been discussed.

A. *Treatments*

In place of simply varying the manpower level of routine preventive police patrol, we propose to vary the actual patrol activities. We do so in part because of what we have learned about varying manpower from the KCPPE.

Before suggesting specific treatments for our proposed experiment, we offer a list of treatments built upon the idea of testing alternative police patrol strategies and tactics:

(1) *Information processing patrol strategies.* These strategies rest on the assumption that sharing information gathered in police patrol among the patrol officers policing a common area will reduce crime by increasing the risk of apprehension of offenders.

Tactic I: *Information officer coordination of beat patrol.* This system, pioneered in England, rests on the principal that each beat officer shares information on crime and offenders in his beat with an information officer who also gathers information from the wider criminal justice system. This information is processed and serves to direct the patrol activities of each beat officer with a view to increasing apprehension of offenders or to increasing their risk of being caught and thus reducing their level of offending.

Tactic II: *Neighborhood team policing.* This tactic emphasizes the creation of a team of officers who regularly patrol a given territory in which they develop a close relationship with potential victims and a knowledge of potential offenders. (Bloch & Bell 1976; Schwartz and Clarren 1977; Sherman et al. 1973)

(2) *Environmental intervention strategies.* These strategies involve active patrol intervention in the physical and social environment of patrol areas or in the activities of citizens as they move about the area.

Tactic I: *Stop & Frisk.* This tactic was examined in patrol areas of Chicago, Boston, and Washington, D.C., in studies by the National Crime Commission in 1966. The assumption is that the identification of potential offenders, particularly those who possess weapons, reduces the risk of their use in crimes. (Reiss 1967)

Tactic II: *Stop and interrogate (question).* This tactic assumes that questioning suspicious individuals deters potential offenders. (Boydstun 1975)

Tactic III: *Directed deterrent patrol.* This includes a variety of tactics designed to specify the kind of activity patrol will undertake when assigned to preventive patrol, e.g., identify vehicles parked in an area, or check the security status of residences or buildings. (Larson 1972; Tien et al. 1976)

Tactic IV: *Tactical squads.* There is a large number of possible types of tactical squads whose tactics include proactive policing. The location-oriented patrol (LOP) and perpetrator-oriented patrol (POP) tactics attempted in Kansas City following the completion of the KCPPE are examples (Pate et al. 1976).

(3) *Visibility strategies.* These strategies are based on the presumption that degree of visibility of police patrol affects the rate of offending in an area. The tactics appear superficially to be contradictory, with unmarked cars as well as marked cars presumed to have effects increasing the apprehension of offenders.

Tactic I: *Unmarked patrol cars.* This tactic is based on the assumption that large manpower allocated to unmarked patrol cars increases the risk of apprehension of offenders committing crimes.

Tactic II: *Reactive patrol.* Here visibility results solely from reactive dispatches into an area in response to complaints. This is a low visibility condition.

Tactic III: *Foot patrol.* This tactic is based on the assumption that foot patrol officers have knowledge of their beat and how to secure it so as to reduce the risk of victimization by crime and the apprehension of offenders.

Tactic IV: *High visibility.* This tactic assumes that offenders are deterred by the open presence of police officers, either on patrol or functioning in some other fashion.

(4) *Crime-specific strategies.* These strategies are based on the assumption that no patrol strategy is effective against all crimes and that the greatest deterrent and apprehension effects arise from tactics that are undertaken for a specific type of crime. Among those that have been tried are:

Tactic I: *Decoy squads for street robbery.* A police team with one or more members serves as decoys by representing potential victims of street robbery. This is intended to increase the apprehension of career street robbers.

Tactic II: *Operation barrier for apprehending robbers.* This tactic dispatches patrol to intercept robbers in flight and is designed especially for commercial and street robbers.

Tactic III: *Identification tactics*: Various tactics of surveillance including technological surveillance (e.g., burglar alarm systems linked to patrol operations) are used to increase the apprehension of criminals by police on patrol.

(5) *Patrol investigation strategy.* This strategy assumes that police patrol can increase the apprehension rate as well as the effectiveness of detective investigation if the police on patrol immediately undertake some investigation of the crime rather than leaving investigation primarily to the detective division.

Tactic I: *Team patrol to manage investigations.* This tactic includes the assignment of detectives to patrol teams (Bloch and Bell 1976).

Tactic II: *Combining detective investigation with patrol function.* Each patrol officer does substantial detective investigation as part of routine patrol.

The experimenter will find that these strategies and the tactics related to them are not readily separable from one another. For example, it is no simple matter to separate foot patrol from the special activities undertaken by police officers on foot patrol that occur in other strategies and tactics.

Were sample size not a severe and overriding problem in the present situation, we would most likely propose a factorial or fractional factorial structure for treatments aimed at measuring the effects of three or more strategies and their interactions. Since the available number of police beat areas is so small, and for purposes of comparability with the KCPPE, we have decided to consider only the following three "treatments." We recognize that the proposed interventions require changes in police behavior that may be very difficult to achieve.

(A) *A search patrol strategy,* involving high visibility (see Strategy 3, Tactic IV) of the participating police officers and intrusive forms of patrol involving field interrogation of suspicious individuals and other forms of environmental intervention (see Strategy 2, Tactics I and II).

(B) *A basic visibility patrol strategy,* involving a standard level of routine preventive patrol.

(C) *A reactive no-patrol strategy* (see Strategy 3, Tactic II), utilizing police officers wherever possible to handle complaints related to particular types of crime via scheduled appointments and in unmarked patrol cars (see Strategy 3, Tactic I).

If the effects of these treatments do differ, and if we assume that there are no interactive effects of visibility and interrogation (i.e., they are additive), then the differences between beats with treatments B and C will be a measure of the effects of the *visibility* component of patrol, and the differences between beats with treatments A and B will be primarily a measure of the effects of interrogation. If the experiment we propose here is successful and detects sizeable differences in the effects of treatments, then subsequent experiments might explore the interrogation versus visibility structure in further detail.

B. *Do we monitor treatments?*

We propose to monitor various aspects of police activities at the beat level and to collect basic response information similar to that in the KCPPE. The purpose of monitoring police activities is to ensure that there is a true difference in treatments, and thus we are interested in having a direct measure of police car visibility. One way to measure visibility is to use, at a variety of locations, counter devices that can distinguish among patrol cars. Such devices would also provide measures of the presence of cars from other beats, e.g., in answer to calls for service.

An (expensive) alternative to counter systems is an automatic vehicle monitoring system such as the FLAIR (Fleet Location and Information Reporting) system currently being implemented in St. Louis on a trial basis. The St. Louis FLAIR system couples a simplified form of "internal guidance" technique with real-time computer tracking to monitor the location of each of 25 patrol cars to within what is claimed to be an average accuracy of 50 feet. Such a system would not only allow for the direct measurement of visibility, but it would also allow for better control of police operations such as dispatching.

We are also interested in monitoring the amount of interrogation, especially in beats using the search patrol strategy. We suggest the filling out of punched cards by officers for each field interrogation or car check.

In addition to the basic response data used in the KCPPE, such as UCR and victimization data and citizen response data, we suggest the collection of additional information such as weapons production. What we are seeking here is the offender protection potential of each strategy. For example, in car checks, we would like to know how often we produce a hot car, a wanted person on warrants, etc. If we increase that strategy, do we increase the numbers of such events? For searches of the person, we would like to know how often we produce a weapon, burglary tools, narcotics, etc. In short, what we want is direct evidence from the activity itself. Do searches produce evidence of a direct criminal sort, e.g., burglary tools, or of a more indirect sort, such as weapons since they could be for self-protection, stolen goods, persons wanted on warrants, etc.? We would also want to know whether weapons production has any effect on particular kinds of crime rates. Indeed, we should expect that increased productivity of weapons taken might be related to reductions in carrying weapons over time.

We also would record separate measurements by "beat-watch." Since there are three standard watches for each beat this would give us a trivariate response variable. We could then examine differences among watches within beats, as well as differences among beats within watch.

C. *How long an experiment?*

The KCPPE ran for a 1-year time period. Because of the important seasonal variation in the levels of various types of crimes and in exposure, the KCPPE could only use one level of preventive patrol in each beat. We propose an experiment that will run for 3 years. This expanded time frame will (1) allow for the use of all three patrol strategies in each beat for a full year, (2) permit the assessment of effects beyond those considered in the KCPPE, and (3) provide an acceptable number of degrees of freedom for an estimate of experimental error. This approach in effect allows each beat to serve as its own control, and thus enables us to control more accurately for beat-to-beat variations.

The time period of 1 year per treatment is dictated by the need to accumulate sufficient data by beat-watch of rare or low productivity events. Such a period also removes the need to adjust directly for seasonal effects, and allows police officers sufficient time to adapt to different patrol strategies. If such surveys were not needed to provide various response measures, more interesting and efficient experimental designs could be proposed. For example, if one were to use a time period of 4 months per treatment, one could apply each treatment twice to each beat (so the length of the experiment becomes 2 years), thus increasing substantially the degrees of freedom associated with the estimate of experimental error. Such a plan would assume that the effects of treatment changes could be assessed after only 4 months. Moreover, in this 2-year experiment each treatment would not be applied to each beat for each of the 4 seasons. Hawthorne effects, resulting from the seemingly continuous changes in treatment, might also result.

Although 3 years is a very long period of time for the implementation of a social experiment, we feel the advantages here outweigh the potential hazards.

D. *The basic crossover design format*

The basic experimental design that we propose has two key features:

(1) Different treatments are applied to the same experimental unit during different periods, and this results in a *crossover* or *changeover* design (see Cochran and Cox 1957).

(2) The use of *primary* and *secondary* experimental units, with the secondary units providing a *double insulation* between "adjacent" primary units.

The initial experiment analysis is aimed at the consideration of the primary units, although the effects of the treatments applied to the secondary units are taken into account. Depending on the results of this initial analysis, subsequent analyses may make use of the secondary units in addition to the primary ones. The units of analysis always remain the individual police beats.

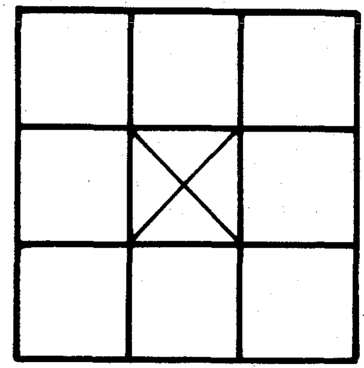
We assume that *all beats* are *square* and of the *same size*. While this is far from true in practice, we expect that modest variations in size and shape will not seriously affect the implementation of the design.

We begin by dividing the beats into 3 × 3 blocks of 9. The basic configuration of beats within blocks has the primary unit surrounded by 8 secondary units as shown in Figure 2.

For any given experimental period we intend to use the same treatment (i.e., patrol strategy) in all 8 secondary beats within a block. Because our initial analyses will use data from only the primary beats, it may seem extremely wasteful to use up a total of 9 beats with potential information to yield only the primary beat. We do so in order to be able to measure the direct effects of strategies in the secondary beats on the primary beat which they surround. The inability to make such measurement appears to have been a serious problem in the KCPPE. If these effects involving adjacent beats are negligible, then we can still carry out subsequent analyses using measurements on all beats.

In the unlikely event that an effect is more likely to be observed in secondary rather than in primary beat units, the analysis could be redone taking into account the units surrounding each secondary unit.

Figure 2.



The reason that this possibility exists is that each beat contains a somewhat different mix of commercial and residential properties, and we expect to find different crimes at different times of the day associated with the two types of properties. For example, the southwest part of the Southern Patrol Division in Kansas City consists primarily of upper middle class residential properties, but one beat in this area contains a modern shopping area, and we would expect the mix of crimes associated with this beat to differ from the other beats nearby. At some point in the analysis the combination of commercial and residential properties within beats should be examined as a possible concomitant variable.

The 3-year experimental time frame is divided into three consecutive 1-year periods, and for each primary beat we use a different patrol strategy during each period. This is the crossover feature referred to above, and each primary beat serves as its own control. We also vary the use of the three patrol strategies in the secondary beats within each block in a similar fashion, so that within each block all 8 secondary beats first use one strategy, then a second, and finally a third. Thus there are 36 possible patrol strategy combinations for a given block of 9 beats, since there are 6 possible orderings of the strategies for the primary beats, and for each of these 6 orderings there are 6 possible orderings of the strategies for the secondary beats. This suggests an experiment involving 36 blocks or a total of 36 × 9 = 324 beats, clearly an impossibility!

By carefully balancing treatment combinations in sets of blocks we can design experiments involving 6, 9, or 12 blocks of interest. This leads us to variants on sets of Graeco-Latin squares treated in the form of crossover designs. Alternatively we can choose a set of treatment combinations at random from the 36 possible ones. We explore both of these possibilities below.

To keep our descriptions of designs compact, henceforth we use the following shorthand notation for patrol strategies. Capital Latin letters refer to strategies used in the primary beats within a block, and Greek letters refer to strategies used in the secondary beats:
A(α) = Search strategy.
B(β) = Visibility strategy.
C(γ) = Reactive no-patrol strategy.

We denote the three successive 1-year time periods as T_1, T_2, and T_3, respectively.

For all the designs we discuss below, we are interested in effects due to:
(1) blocks of beats.
(2) time period.
(3) patrol strategies in primary units.
(4) patrol strategies in adjacent secondary units.

We would also like to be able to test for the presence of various interaction effects, such as:
(5) primary by secondary patrol strategies.
(6) time by patrol strategy (either in the form of residual or carryover effects), for both primary and secondary beats.

A residual effect during a particular time period depends only on the treatment (strategy) used in the preceding period. A carryover effect, as used here, depends on both the treatment in the preceding period and the treatment in the present one. Thus the "carryover" model allows for different effects resulting from the preceding treatment, when the transition is from search to control or from search to reactive, whereas the model based on residual effects forces the effect of the preceding treatment to be the same in both cases.

Of all the effects listed here the one we consider to be least substantial is that of the interaction between the patrol strategy in the primary beat and the patrol strategy in the secondary beat. If we were forced to choose between including this interaction in our model and design or achieving superior experimental balance with additional degrees of freedom for estimating the error term, we would opt for the latter. This is because the magnitude of detectable effects suggested by all previous controlled studies is not that great, and the physical (as opposed to parametric) interactive effect of secondary units on adjacent primary units is already taken into account by the model. The next least important among the effects would seem to be residual effects due to secondary-beat strategies.

The model that goes with this experimental structure assumes that each response variable (or some function of it, like the logarithm or log-odds in the case of proportions) is the sum of effects attributable to each source. Thus if we are looking at a beat with primary strategy A and secondary strategy β during period T_2, when strategies C and α were used in T_1, the model with residual effects postulates that the *response variate*
= (beat effect)
+ (main effect due to period T_2)
+ (main effect due to A)
+ (main effect due to β)
+ (residual effect due to C)
+ (residual effect due to α)
+ (interaction effect due to A × β)
+ (random error).

If we use carryover effects, then in this model we replace the residual effect due to C and the residual effect due to α by a carryover effect of C preceding A and a carryover effect of α preceding β.

Note that the model includes a term labelled "beat effect." All beats do not start from the same level of the response variable, e.g., beats differ in their rate of automobile theft. The crossover design described here uses each beat as its own control, and attempts to remove the initial differences among beats by incorporating an additive effect due to beats in the model. The fact that beats with low crime productivity would not be expected to respond to particular treatments in the same "additive way" as beats with high crime productivity suggests that when crime rates are used as response variables, the additive model is more appropriate for the logarithm of the crime rate. The use of such transformations is standard practice in modern statistical data analysis.

Most crossover experimental designs that assume the presence of residual effects do so only for the actual time frame of the experiment, i.e., they assume that there are no residual effects for period T_1. Although the discussion below is based on such an assumption, it would seem to be inappropriate in the proposed experiment. In particular, we might expect to find a residual effect during T_1 in the beat with primary strategy C and secondary strategy γ, since prior to the experiment all beats supposedly are using the current preventive patrol strategy (i.e., B and β). The analyses described below can take such effects into account but the modifications result in some additional nonorthogonalities in the associated analysis of variance breakdowns.

E. *Random crossover designs*

In order to be able to test for all of the effects listed above we require a minimum of 9 or 10 blocks (depending on whether we use residual or carryover effects). To ensure a sufficient number of degrees of freedom for error we probably need a minimum of 12 blocks unless we are prepared to assume the absence of interaction effects.

In this subsection we describe two examples of "random crossover" designs, one with 12 blocks that can be used to get tests for all of the different effects and a second one with 6 blocks that can be used to get tests for the main effects.

One possible random 12-block design can be described as follows:

Block	T_1	T_2	T_3
1	Aγ	Bβ	Cα
2	Cγ	Aβ	Bα
3	Bβ	Aγ	Cα
4	Cβ	Bα	Aγ
5	Bα	Cγ	Aβ
6	Bβ	Cα	Aγ
7	Cβ	Aα	Bγ
8	Bα	Aβ	Cγ
9	Bβ	Cγ	Aα
10	Cα	Aγ	Bβ
11	Aα	Bβ	Cγ
12	Cγ	Bα	Aβ

Table 1. ANOVA layout for 12-block cross-over design with carry-over effects

Source	df
Blocks (beats)	11
Time	2
Primary strategy	2
Secondary strategy	2
Primary x secondary	4
Primary carry-over	3
Secondary carry-over	3
Error	8
Total	35

Table 2. ANOVA layout for 12-block cross-over design with residual effects

Source	df
Blocks (beats)	11
Time	2
Primary strategy	2
Secondary strategy	2
Primary x secondary	4
Primary residual	2
Secondary residual	2
Error	10
Total	35

Table 3. ANOVA layout for 6-block random cross-over design

Source	df
Blocks (beats)	5
Time	2
Primary strategy	2
Secondary strategy	2
Error	6
Total	17

Table 4.

Block	T_1	T_2	T_3
1	Aα	Bβ	Cγ
2	Bγ	Cα	Aβ
3	Cβ	Aγ	Bα
4	Bα	Aγ	Cβ
5	Cγ	Bβ	Aα
6	Aβ	Cα	Bγ
7	Aγ	Bβ	Cα
8	Bα	Cγ	Aβ
9	Cβ	Aα	Bγ
10	Aα	Cβ	Bγ
11	Bβ	Aγ	Cα
12	Cγ	Bα	Aβ

As indicated above, each patrol strategy is used exactly once for the primary unit in each block and exactly once for the secondary units. Because we randomly selected 12 of 36 strategy combinations A appears only twice in T_1, while B and C appear 5 times each. This lack of balance leads to a nonorthogonal analysis of variance where the ordering of the effects to be tested may make a difference. Of course other random crossover designs may be better than this one, while many others will definitely be worse.

All 6 effects in the original list in Section IV-D are estimable with this particular random design, and there is not too great a disparity in the precision of the estimates if we use a standard Analysis of Variance (ANOVA) model meeting the usual restrictions on parameters. Table 1 lists the effects and the associated degrees of freedom for the model with carryover effects and Table 2 lists these for the model with residual effects.

As we remarked in the preceding subsection, modifications in the standard ANOVA analyses are required to take into account carryover effects from the standard treatment (B and β) used prior to the experimental period.

The first 6 blocks in the preceding design can be used to illustrate a 6-block random design that can be used to test the four sets of main effects. The resulting ANOVA is given in Table 3.

F. *Balanced crossover designs*

Rather than trust the estimability of the various effects of interest to a random selection of 12 out of the 36 treatment sequences, we can try to prepare a design that balances the assignment of treatments to experimental units in a crossover experiment. Patterson (1952) has given the following list of 7 conditions for balance in a general crossover design:

(1) No treatment occurs in a given sequence more than once.

(2) Each treatment occurs in a given time period an equal number of times.

(3) Every two treatments occur together in the same number of sequences.

(4) Each ordered succession of two treatments should occur equally often in sequences.

(5) Every two treatments occur together in the same number of curtailed sequences formed by omitting the final period.

(6) In those sequences in which a given treatment occurs in the final period, the other treatments occur equally often.

(7) In those sequences in which a given treatment occurs in any but the final period, each other treatment occurs equally often in the final period.

These conditions for balance are for crossover designs involving only one type of treatment, i.e., only one for the Greek or Latin labelled effects used above. Conditions (1), (2), and (3) ensure that main effects are estimable, while if both conditions (1)–(3) and (4)–(7) are satisfied, then both main effects and residual effects are guaranteed to be estimable and obtainable in a simple form. To apply these conditions to the design problem considered here, we consider them once for treatments assigned to primary beats and a second time for treatments assigned to secondary beats. As far as we know, this will still not necessarily ensure that the primary by secondary interaction parameters are estimable, so for any particular balanced design satisfying (1)–(7) we must make a special check regarding the estimability of these parameters.

The random design described in the previous subsection satisfies conditions (1), (3), and (6), but not the others. Clearly any set of sequences built using Graeco-Latin squares of the form

Aα	Bβ	Cγ
Bγ	Cα	Aβ
Cβ	Aγ	Bα

(in which each Greek and Latin letter appears only once in each row and column, and each Greek letter appears only once with each Latin one), must satisfy conditions (1), (2), and (3). The design in Table 4 was built with 4 such Graeco-Latin squares, and the reader may verify that the design satisfies Patterson's 7 conditions for both Greek and Latin letters. In this particular design the primary by secondary treatment interaction parameters are also estimable.

The ANOVA layouts for the balanced designs are the same as in Tables 1 and 2 for the random design, but the Graeco-Latin square structure will make several of the sums of squares orthogonal to one another, a situation that did not exist for the random design.

The first six sequences in the above design (or the last six) may be used by themselves to give a balanced design for estimating only the main effects as in Table 3. These 6-block designs have the added feature that they allow for the estimation of both primary and secondary residual effects. If we choose to break both of these residual sums of squares out of the error sum of squares, the resulting error sum of squares has only 2 degrees of freedom, and thus any ANOVA tests will have very low power.

In order to ensure that the analyses we are about to describe are proper, the allocation of symbols (both Greek and Latin) to treatments, and sequences to blocks of beat *must be made at random* (Patterson 1952). A systematic assignment of treatments simply will not do.

G. *Comparing the balance and random designs*

Although there is an aesthetic appeal to the balanced design described above, we have no guarantee that it is in fact superior to the random, unbalanced design. In Tables 5 and 6 we give the variances associated with each of the estimated effects in the two complete (12-block) models, the one with residual effects and the one with crossover effects. Each variance is some multiple of the variance, σ^2, of the random error term. In Table 7 we give similar expected variances for the two 6-block designs (random and balances) assuming the presence of neither residual, carryover, nor interaction effects.

For the 12-block experiment and the model with residual effects the balanced design seems superior to the random one, except for a few primary by secondary interaction parameters. For the model with carryover effects, however, the results are less clear. The balanced design is superior for the main effects due to treatments in the primary and secondary beats, but otherwise the random design is by and large superior. All things considered, we prefer the balanced design. For the 6-block experiment the balanced design is apparently superior again.

Further comparison between the designs has been carried out in terms of the statistical power of the various tests for effects resulting from the ANOVA. The calculations are summarized in a series of tables presented in the Appendix. The power calculations seem to substantiate the superiority of the balanced design and suggest that we have a very high chance of detecting most effects that are about the size of the standard deviation of the random error term in the model.

H. *Response variables to be measured*

The KCPPE used three types of response variables to assess the effects of different patrol strategies:

Table 5. Variances of parameter estimates in the model with residual effects

		Balanced	Random
Treatments	A	$.074\sigma^2$	$.190\sigma^2$
	B	.074	.190
	C	.076	.114
Insulation	α	.076	.131
	β	.074	.098
	γ	.074	.171
Residual treatment	A	.172	.333
	B	.172	.205
	C	.149	.247
Residual insulation	α	.149	.231
	β	.172	.212
	γ	.172	.242
Interaction effects	Aα	.241	.483
	Aβ	.602	.562
	Aγ	.602	.437
	Bα	.241	.734
	Bβ	.602	.442
	Bγ	.602	.806
	Cα	.296	.507
	Cβ	.241	.559
	Cγ	.234	.806

		Balanced	Random
Block effects	B_1	.492	.700
	B_2	.656	.722
	B_3	.492	.697
	B_4	.492	.493
	B_5	.492	.763
	B_6	.656	.698
	B_7	.647	1.16
	B_8	.517	.842
	B_9	.517	.462
	B_{10}	.517	.581
	B_{11}	.647	.436
	B_{12}	.517	.763
Time	T_1	.074	.089
	T_2	.185	.070
	T_3	.185	.088

Table 6. Variances of parameter estimates in the model with cross-over effects

		Balanced	Random
Treatments	A	$.063\sigma^2$	$.140\sigma^2$
	B	.063	.185
	C	.056	.132
Insulation	α	.056	.145
	β	.063	.109
	γ	.063	.166
A vs B before C		.618	.579
A vs C before B		.619	.653
B vs C before A		.619	.353
α vs β before γ		.619	.482
α vs γ before β		.619	.408
β vs γ before α		.618	.467
Interaction effects	Aα	.741	.594
	Aβ	.852	.701
	Aγ	.852	.489
	Bα	.741	.915
	Bβ	.852	.596
	Bγ	.852	.930
	Cα	.296	.606
	Cβ	.741	.751
	Cγ	.741	.946

		Balanced	Random
Block effects	B_1	1.18	.789
	B_2	1.08	1.03
	B_3	1.18	.843
	B_4	1.18	1.22
	B_5	1.18	.921
	B_6	1.08	.777
	B_7	1.08	1.32
	B_8	1.18	.960
	B_9	1.18	.881
	B_{10}	1.18	.672
	B_{11}	1.08	.784
	B_{12}	1.18	.903
Time	T_1	.074	.117
	T_2	.185	.080
	T_3	.185	.090

(1) Survey and questionnaire data, including results from both community and commercial surveys regarding victimization and attitudes.

(2) Interviews with officers and participant observations.

(3) Departmental data such as those on reported crime, traffic, and arrests.

Given the nature of the crossover design being proposed, carrying out multiple victimization surveys seems highly impractical because of the large sample sizes required to get useful information (the sample sizes in the KCPPE were clearly much too small to produce useful information). Therefore we do not propose to carry out any surveys as part of an effort to directly monitor the ongoing effects of the experiment.

It would be extremely useful to have data on response time to calls for service. The KCPPE collected such data through a pair of surveys (one of observers and the other of citizens). If an automatic vehicle monitoring system such as FLAIR is used to track police vehicles during the experiment as we propose, then departmental data on response times would be available for analysis. Analysis need not be based on complete data for each time period, but measurements at three points for each condition (beginning, middle, end) seem desirable. The fact that the beginning of one condition is the end of another simplifies matters here somewhat.

Not only do we propose to analyze direct measures of reported crime and police production (e.g., arrests) which are part of

Table 7. Variances of parameter estimates in 6-beat design

		Balanced	Random
Treatments	A	$.111\sigma^2$	$.233\sigma^2$
	B	.111	.192
	C	.111	.125
Insulation	α	.111	.215
	β	.111	.131
	γ	.111	.192
Block effects	B_1	.278	.278
	B_2	.278	.278
	B_3	.278	.278
	B_4	.278	.278
	B_5	.278	.278
	B_6	.278	.278
Time	T_1	.111	.159
	T_2	.111	.111
	T_3	.111	.159

departmental records, but we would also like to have several offender-related response measures since we anticipate that the treatment tactics will have "deterrent" effects on potential offenders. Some possibilities are:

(a) *Data on modus operandi for offenses.* These are hard to get and will be the least sensitive to changes.

(b) *Data on rearrest rates and place of occurrence of offense for rearrested offenders.*

(c) *Data on characteristics of offenders for crimes against the person*, e.g., changes in race, age, and sex of offenders in beats. Thus, for example, blacks might commit fewer offenses in white areas with an interrogation procedure or juveniles might be a smaller proportion of all offenders.

(d) *The ratio of crimes against persons to property.* This is again a highly indirect measure but offenders might shift more to crimes of stealth under increased proactivity. On the whole this would be a difficult measure to interpret.

(e) *Changes in the population of offenders with a first arrest and in proportion of offenders with a first arrest.* Both have defects as measures, yet they could be affected by proactive tactics.

(f) *Changes in distance between place of occurrence of offense and residence of offender.* If an entire city is used, this should not change if resources were constant, but they are not. Offenders might be more inclined to go greater distances where the police have a larger territory to cover with increased proactivity.

The foregoing is a tentative list of possible response variables to be measured and analyzed. Considerably more work is required to refine this part of the proposal.

I. *Boundary effects*

In Section III-E we discussed problems in the KCPPE which resulted from some beats in the experimental area being on the boundaries of the city of Kansas City, Missouri. By making a few plausible assumptions, we can turn what was formerly a disadvantage into a benefit and, in the process, cut down on the actual number of police beats required to carry out the experiment.

Let us suppose that those secondary beats in a block sharing a side with the primary beat each contribute equally to the appropriate effect terms, while the four secondary beats that share only a corner with the primary beat contribute equal but lesser amounts (say ½ of the former). Then we can alter the number of beats in those blocks on a boundary, by eliminating as many secondary beats as necessary. For example, suppose we have a design with nine blocks laid out so that the experimental area takes the form of a 9 × 9 layout of square beats. If we have an actual area of roughly the same size but possessing 49 beats in a 7 × 7 area, then we eliminate the secondary beats all around the original design resulting in the layout shown in Figure 3.

Note that the 4 blocks in the corner have only 3 secondary beats, while 4 others have exactly 5, and only one block has 8.

Analyzing designs laid out in a fashion similar to the one above requires modifications to the more or less standard analyses for the "complete" designs described earlier, but the modifications are straightforward, corresponding directly to the assumptions regarding the effects from the secondary beats described above.

In Kansas City, not only can we reduce the number of beats required for our designs by taking advantage of the city boundaries, we can make similar reductions with blocks of beats bordering the Missouri River. The river slices across the city and, according to reports we have had, police patrolling beats on one side of the river rarely answer calls for service on the other side. Thus we may well be able to reduce the 108 beats required for implementing our 12-block balanced design to the 69 beats that actually exist in Kansas City.

Some care must be taken in using the approach outlined here in order to ensure that all of the primary beats do not fall adjacent to city and natural boundaries. For example, if we have 6 blocks resulting in a 6 x 9 layout of beats, eliminating all of the exterior secondary beats leads to a layout with all 6 primary beats on the boundary. Some type of compromise would probably be wise in such circumstances.

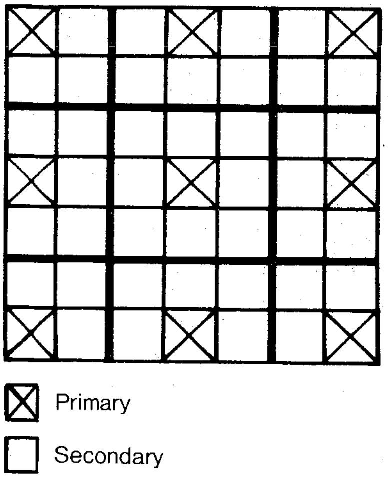
V. Further considerations and summary

For policy decisions regarding police deployment and manpower allocation we require as precise estimates as possible of the effects of one policy versus another. Thus once we carry out all of the tests corresponding to the various sources in the ANOVA tables of the previous section, we should bolster our estimates of various effects by directly utilizing data from the secondary beats. Since reported crimes are already recorded by beat there would be no added cost for collecting additional data on the secondary beats.

The experimental designs described in Section IV may easily be adapted for other experiments involving police deployment and manpower allocation, such as ones dealing with comparisons between the use of one-man and two-man patrol cars.

We have noted earlier that the proposed experiment involves neither changes in police manpower nor the spatial redistribution of patrol resources within a city, as was the case in the KCPPE. Thus no policy implications regarding levels of manpower can possibly result. Moreover, since the different patrol strategies will be applied to small areas, even if the results are striking the policy implications for overall police patrol strategy in a city may be difficult to formulate due to the use of the different strategies in adjacent beats in the experiment. For example, if there is no detectable difference among treatment effects, we must ask if the results suggest a policy of "business as usual" or a new overall strategy of low visibility patrol. The answer seems unclear.

Figure 3.



Finally we must ask how stable the estimates of parameters in the experimental model will be *over time, and across cities.* Variables that influence the behavior of offenders may interact with the experimental variables and with the peculiarities of cities and times in such a way as to make estimates highly unstable. If this is the case, then changes and refinements in the experimental design are required. Further consideration must be given to the potential policy implications of the experiment as part of the design.

Many so-called social experiments are not really experiments in the formal statistical sense. All too often, when trying to evaluate whether some innovation or social program has the intended effect, investigators fail to provide a direct means for separating out or controlling for the effects of changes in concomitant variables. To overcome the difficulties inherent in such studies, social scientists have turned to the use of randomized controlled field trials. Both the randomization and the control are crucial to such studies.

Many social experiments, as a result of their failure to control for various sources of variability, end up by proving the null hypothesis. The analysis of experimental results is structured toward rejecting the null hypothesis (typically that of no difference) in favor of some alternative, by giving heavy weight to the null hypothesis unless it is demonstrably false. The level of significance (the probability of falsely rejecting the null hypothesis) is set at some low value and then the probability of failing to detect the alternative hypothesis when it is true is made problematic. Unless that probability also is small, the experiment will be weighted in favor or the null hypothesis. Such, we have argued, was likely the case in the KCPPE because of its weak experimental design, i.e., its failure to control for important sources of variability. At least as crucial, in our view, was the failure of the KCPPE investigators to use proper randomization of treatments to experimental units.

The problem of what design to adopt for social experiments is not easily resolved. Here we call attention to a particular class of designs that are useful not only for experiments relating to patrol strategies, but also to other types of social experiments. Our basic design has two key features:

(1) Different treatments are applied to the same experimental units during different time periods, resulting in a crossover or changeover structure where each unit is its own control.

(2) Additional control is exercised through a special device whereby primary experimental units are insulated from one another by the addition of secondary experimental units, with effects being examined in both the primary and secondary units. This paper explores different orders of complexity in this design and discusses means for evaluating tradeoffs among them.

It is apparent that more complex experimental designs controlling for various sources of variation will in general be more costly, take longer periods of time, and involve greater use of repeated measures than the simple experimental design adopted for the KCPPE. If public policy is to rest on social experiments where causal inference is essential, the gain in efficiency associated with complex designs should ordinarily outweigh any increased cost.

Appendix:
Power calculations for comparing balanced and random designs for ANOVA models in Section IV

The tables in this appendix give illustrations of the magnitude of effects capable of detection through the use of standard F-tests with an 0.05 level of significance, with power = Pr (reject | H_A true) = 0.50, 0.70, 0.90, and 0.95. Table A-1 is for the 12-block experiment and the model with carryover effects. Table A-2 is for the 12-block experiment and residual effects. Table A-3 is for the 6-block experiment. At least two illustrations are given for each effect.

For treatment and main effects, the tests were based on adjustments first for other main effects (this is the standard form of analysis), and then for all other effects including interactions, carryover effects, and residual effects.

How to read the tables. Suppose we are interested in compared experiments using an 0.05 level of significance. Then the first set of rows in the table tell us that, if A = -0.55σ, B = 0, and C = $+0.55\sigma$ (where σ^2 is the variance of the random error term), the usual F-tests for treatments would have detected these effects for the balanced design with power = 0.50.

Acknowledgments

This paper grew out of material discussed in the Workshop on Criminal Justice Statistics, held in Washington, D.C., July 1975, and sponsored by the Social Science Research Council Center for Coordination of Research on Social Indicators and the Law Enforcement Assistance Administration. We are indebted to David W. Britt, William Fairly, Richard Larson, Katherine C. Lyall, and Richard Sparks for several important suggestions and comments. An abbreviated version of an earlier draft of this paper appeared under the same title in *Evaluation 3* (1976), 124–131.

During the preparation of this paper, Stephen E. Fienberg's research was partially supported by grants from The Robert Wood Johnson Foundation and the Commonwealth Fund to the Center for Analysis of Health Practices, Harvard School of Public Health, and from the National Science Foundation Grant SOC72-05257 to the Department of Statistics, Harvard University. Albert J. Reiss' research was partially supported by a grant from the Russell Sage Foundation.

Table A-1. Carry-over analysis

(a)		50% Power Balanced	50% Power Random	70% Power Balanced	70% Power Random	90% Power Balanced	90% Power Random	95% Power Balanced	95% Power Random
Treatment effects (adjusting only for main effects)	A	−.55	−.57	−.69	−.72	−.89	−.93	−.99	−1.03
	B	0	0	0	0	0	0	0	0
	C	+.55	+.57	+.69	+.72	+.89	+.93	+.99	+1.03
Treatment effects (adjusting for all effects)	A	−.56	−.80	−.70	−1.02	−.91	−1.31	−1.01	−1.45
	B	0	0	0	0	0	0	0	0
	C	+.56	+.80	+.70	+1.02	+.91	+1.31	+1.01	+1.45
Treatment effects (adjusting only for main effects)	A	+.63	+.66	+.80	+.83	+1.03	+1.07	+1.14	+1.19
	B	−.32	−.33	−.40	−.41	−.51	−.54	−.57	−.60
	C	−.32	−.33	−.40	−.41	−.51	−.54	−.57	−.60
Treatment effects (adjusting for all effects)	A	+.67	+.98	+.85	+1.24	+1.09	+1.60	+1.21	+1.78
	B	−.34	−.49	−.42	−.62	−.55	−.80	−.61	−.89
	C	−.34	−.49	−.42	−.62	−.55	−.80	−.61	−.89
Insulation effects (adjusting only for main effects)	α	−.55	−.57	−.69	−.72	−.89	−.93	−.99	−1.03
	β	0	0	0	0	0	0	0	0
	γ	+.55	+.57	+.69	+.72	+.89	+.93	+.99	+1.03
Insulation effects (adjusting for all effects)	α	−.56	−.96	−.70	−1.20	−.91	−1.55	−1.01	−1.73
	β	0	0	0	0	0	0	0	0
	γ	+.56	+.96	+.70	−1.20	+.91	+1.55	+1.01	+1.73
Insulation effects (adjusting only for main effects)	α	+.63	+.64	+.80	+.81	+1.03	+1.04	+1.14	+1.16
	β	−.32	−.32	−.40	−.40	−.51	−.52	−.57	−.58
	γ	−.32	−.32	−.40	−.40	−.51	−.52	−.57	−.58
Insulation effects (adjusting for all effects)	α	+.63	+.99	+.80	+1.25	+1.03	+1.61	+1.14	+1.79
	β	−.32	−.50	−.40	−.62	−.51	−.81	−.57	−.90
	γ	−.32	−.50	−.40	−.62	−.51	−.81	−.57	−.90
(b)									
Treatment carry-over effects	A before C	+1.18	+1.84	+1.47	+2.30	+1.89	+2.95	+2.08	+3.24
	B before C	−1.18	−1.84	−1.47	−2.30	−1.89	−2.95	−2.08	−3.24
	A before B	+.59	+.92	+.74	+1.15	+.95	+1.48	+1.04	+1.62
	C before B	−.59	−.92	−.74	−1.15	−.95	−1.48	−1.04	−1.62
	B before A	0	0	0	0	0	0	0	0
	C before A	0	0	0	0	0	0	0	0
Treatment carry-over effects	A before C	0	0	0	0	0	0	0	0
	B before C	0	0	0	0	0	0	0	0
	A before B	0	0	0	0	0	0	0	0
	C before B	0	0	0	0	0	0	0	0
	B before A	+1.53	+1.53	+1.91	+1.91	+2.45	+2.46	+2.69	+2.70
	C before A	−1.53	−1.53	−1.91	−1.91	−2.45	−2.46	−2.69	−2.70
Insulation carry-over effects	α before γ	+1.18	+1.38	+1.48	+1.73	+1.90	+2.20	+2.07	+2.44
	β before γ	−1.18	−1.38	−1.48	−1.73	−1.90	−2.20	−2.07	−2.44
	α before β	+.59	+.69	+.74	+.86	+.95	+1.10	+1.04	+1.22
	γ before β	−.59	−.69	−.74	−.86	−.95	−1.10	−1.04	−1.22
	β before α	0	0	0	0	0	0	0	0
	γ before α	0	0	0	0	0	0	0	0
Insulation carry-over effects	α before γ	0	0	0	0	0	0	0	0
	β before γ	0	0	0	0	0	0	0	0
	α before β	0	0	0	0	0	0	0	0
	γ before β	0	0	0	0	0	0	0	0
	β before α	+1.52	+1.84	+1.90	+2.30	+2.44	+2.96	+2.68	+3.25
	γ before α	−1.52	−1.84	−1.90	−2.30	−2.44	−2.96	−2.68	−3.25
(c)									
Interaction effects	Aα	+5.16	+3.72	+6.37	+4.59	+8.19	+5.91	+8.95	+6.45
	Aβ	−.65	−.47	−.80	−.57	−1.02	−.74	−1.12	−.81
	Aγ	−.65	−.47	−.80	−.57	−1.02	−.74	−1.12	−.81
	Bα	−.65	−.47	−.80	−.57	−1.02	−.74	−1.12	−.81
	Bβ	−.65	−.47	−.80	−.57	−1.02	−.74	−1.12	−.81
	Bγ	−.65	−.47	−.80	−.57	−1.02	−.74	−1.12	−.81
	Cα	−.65	−.47	−.80	−.57	−1.02	−.74	−1.12	−.81
	Cβ	−.65	−.47	−.80	−.57	−1.02	−.74	−1.12	−.81
	Cγ	−.65	−.47	−.80	−.57	−1.02	−.74	−1.12	−.81
Interaction effects	Aα	+3.39	+1.90	+4.18	+2.34	+5.38	+3.01	+5.87	+3.29
	Aβ	−.97	−.54	−1.19	−.67	−1.54	−.86	−1.68	−.94
	Aγ	−.97	−.54	−1.19	−.67	−1.54	−.86	−1.68	−.94
	Bα	−.97	−.54	−1.19	−.67	−1.54	−.86	−1.68	−.94
	Bβ	+1.94	+1.08	+2.39	+1.34	+3.07	+1.72	+3.36	+1.88
	Bγ	−.97	−.54	−1.19	−.67	−1.54	−.86	−1.68	−.94
	Cα	−.97	−.54	−1.19	−.67	−1.54	−.86	−1.68	−.94
	Cβ	−.97	−.54	−1.19	−.67	−1.54	−.86	−1.68	−.94
	Cγ	+.48	+.27	+.60	+.33	+.77	+.43	+.84	+.47
Interaction effects	Aα	+2.00	+1.12	+2.47	+1.38	+3.17	+1.78	+3.47	+1.94
	Aβ	−1.00	−.56	−1.23	−.69	−1.59	−.89	−1.73	−.97
	Aγ	−1.00	−.56	−1.23	−.69	−1.59	−.89	−1.73	−.97
	Bα	−1.00	−.56	−1.23	−.69	−1.59	−.89	−1.73	−.97
	Bβ	+2.00	+1.12	+2.47	+1.38	+3.17	+1.78	+3.47	+1.94
	Bγ	−1.00	−.56	−1.23	−.69	−1.59	−.89	−1.73	−.97
	Cα	−1.00	−.56	−1.23	−.69	−1.59	−.89	−1.73	−.97
	Cβ	−1.00	−.56	−1.23	−.69	−1.59	−.89	−1.73	−.97
	Cγ	+2.00	+1.12	+2.47	+1.38	+3.17	+1.78	+3.47	+1.94

Table A-2. Residual analysis

		50% Power		70% Power		90% Power		95% Power	
(a)		Balanced	Random	Balanced	Random	Balanced	Random	Balanced	Random
Treatments (adjusting only for main effects)		−.53	−.55	−.66	−.68	−.85	−.89	−.94	−.98
		0	0	0	0	0	0	0	0
		+.53	+.55	+.66	+.68	+.85	+.89	+.94	+.98
Treatments (adjusting for all effects)		−.62	−.81	−.77	−1.00	−.99	−.85	−1.10	−1.44
		0	0	0	0	0	0	0	0
		+.62	+.81	+.77	+1.00	+.99	+.85	+1.10	+1.44
Treatments (adjusting only for main effects)		+.61	+.64	+.76	+.79	+.98	+1.03	+1.09	+1.14
		−.31	−.32	−.38	−.40	−.49	−.51	−.55	−.57
		−.31	−.32	−.38	−.40	−.49	−.51	−.55	−.57
Treatments (adjusting for all effects)		+.71	+1.09	+.88	+1.35	+1.14	+1.75	+1.26	+1.94
		−.35	−.54	−.44	−.68	−.57	−.87	−.63	−.97
		−.35	−.54	−.44	−.68	−.57	−.87	−.63	−.97
Insulation (adjusting only for main effects)		−.53	−.55	−.66	−.69	−.85	−.89	−.94	−.99
		0	0	0	0	0	0	0	0
		+.53	+.55	+.66	+.69	+.85	+.89	+.94	+.99
Insulation (adjusting for all effects)		−.62	−.91	−.77	−1.13	−.99	−1.46	−1.10	−1.62
		0	0	0	0	0	0	0	0
		+.62	+.91	+.77	+1.13	+.99	+1.46	+1.10	+1.62
Insulation (adjusting only for main effects)		+.61	+.62	+.76	+.77	+.98	+.99	+1.09	+1.10
		−.31	−.31	−.38	−.38	−.49	−.50	−.55	−.55
		−.31	−.31	−.38	−.38	−.49	−.50	−.55	−.55
Insulation (adjusting for all effects)		+.72	+.89	+.89	+1.11	+1.15	+1.44	+1.28	+1.59
		−.36	−.45	−.44	−.55	−.58	−.72	−.64	−.80
		−.36	−.45	−.44	−.55	−.58	−.72	−.64	−.80
(b)									
Treatment residual	After A	−.89	−1.24	−1.10	−1.54	−1.42	−2.00	−1.58	−2.21
	After B	0	0	0	0	0	0	0	0
	After C	+.89	+1.24	+1.10	+1.54	+1.42	+2.00	+1.58	+2.21
Treatment residual	After A	+1.07	+1.49	+1.33	+1.85	+1.72	+2.40	+1.91	+2.66
	After B	−.54	−.75	−.67	−.93	−.86	−1.20	−.96	−1.33
	After C	−.54	−.75	−.67	−.93	−.86	−1.20	−.96	−1.33
Insulation residual	After α	−.89	−1.11	−1.10	−1.38	−1.42	−1.79	−1.58	−1.98
	After β	0	0	0	0	0	0	0	0
	After γ	+.89	+1.11	+1.10	+1.38	+1.42	+1.79	+1.58	+1.98
Insulation residual	After α	+1.00	+1.25	+1.24	+1.54	+1.61	+2.00	+1.78	+2.22
	After β	−.50	−.62	−.62	−.77	−.80	−1.00	−.89	−1.11
	After γ	−.50	−.62	−.62	−.77	−.80	−1.00	−.89	−1.11
Interaction	$A\alpha$	+3.11	+3.45	+3.83	+4.24	+4.88	+5.40	+5.33	+5.91
	$A\beta$	−.39	−.43	−.48	−.53	−.61	−.68	−.67	−.74
	$A\gamma$	−.39	−.43	−.48	−.53	−.61	−.68	−.67	−.74
	$B\alpha$	−.39	−.43	−.48	−.53	−.61	−.68	−.67	−.74
	$B\beta$	−.39	−.43	−.48	−.53	−.61	−.68	−.67	−.74
	$B\gamma$	−.39	−.43	−.48	−.53	−.61	−.68	−.67	−.74
	$C\alpha$	−.39	−.43	−.48	−.53	−.61	−.68	−.67	−.74
	$C\beta$	−.39	−.43	−.48	−.53	−.61	−.68	−.67	−.74
	$C\gamma$	−.39	−.43	−.48	−.53	−.61	−.68	−.67	−.74
(c)									
Interaction effects	$A\alpha$	+1.90	+1.74	+2.35	+2.15	+2.98	+2.73	+3.26	+2.99
	$A\beta$	−.54	−.50	−.67	−.61	−.85	−.78	−.93	−.85
	$A\gamma$	−.54	−.50	−.67	−.61	−.85	−.78	−.93	−.85
	$B\alpha$	−.54	−.50	−.67	−.61	−.85	−.78	−.93	−.85
	$B\beta$	+1.09	+1.00	+1.34	+1.23	+1.71	+1.56	+1.87	+1.71
	$B\gamma$	−.54	−.50	−.67	−.61	−.85	−.78	−.93	−.85
	$C\alpha$	−.54	−.50	−.67	−.61	−.85	−.78	−.93	−.85
	$C\beta$	−.54	−.50	−.67	−.61	−.85	−.78	−.93	−.85
	$C\gamma$	+.27	+.25	+.34	+.31	+.43	+.39	+.47	+.43
Interaction effects	$A\alpha$	+1.14	+1.03	+1.40	+1.27	+1.78	+1.62	+1.95	+1.77
	$A\beta$	−.57	−.52	−.70	−.64	−.89	−.81	−.98	−.89
	$A\gamma$	−.57	−.52	−.70	−.64	−.89	−.81	−.98	−.89
	$B\alpha$	−.57	−.52	−.70	−.64	−.89	−.81	−.98	−.89
	$B\beta$	+1.14	+1.03	+1.40	+1.27	+1.78	+1.62	+1.95	+1.77
	$B\gamma$	−.57	−.52	−.70	−.64	−.89	−.81	−.98	−.89
	$C\alpha$	−.57	−.52	−.70	−.64	−.89	−.81	−.98	−.89
	$C\beta$	−.57	−.52	−.70	−.64	−.89	−.81	−.98	−.89
	$C\gamma$	+1.14	+1.03	+1.40	+1.27	+1.78	+1.62	+1.95	+1.77

Table A-3. 6-beat analysis

		50% Power		70% Power		90% Power		95% Power	
		Balanced	Random	Balanced	Random	Balanced	Random	Balanced	Random
Treatment effects	A	−.84	−.99	−1.04	−1.23	−1.37	−1.61	−1.51	−1.79
	B	0	0	0	0	0	0	0	0
	C	+.84	+.99	+1.04	+1.23	+1.37	+1.61	+1.51	+1.79
Treatment effects	A	+.96	+1.36	+1.20	+1.70	+1.58	+2.23	+1.75	+2.47
	B	−.48	−.68	−.60	−.85	−.79	−1.11	−.87	−1.24
	C	−.48	−.68	−.60	−.85	−.79	−1.11	−.87	−1.24
Insulation effects	α	−.84	−1.19	−1.04	−1.48	−1.37	−1.95	−1.51	−2.16
	β	0	0	0	0	0	0	0	0
	γ	+.84	+1.19	+1.04	+1.48	+1.37	+1.95	+1.51	+2.16
Insulation effects	α	+.96	+1.31	+1.20	+1.64	+1.58	+2.15	+1.75	+2.38
	β	−.48	−.66	−.60	−.82	−.79	−1.07	−.87	−1.19
	γ	−.48	−.66	−.60	−.82	−.79	−1.07	−.87	−1.19

References

Bloch, P. B., and J. Bell (1976)
Managing investigations: The Rochester system. Washington, D.C.: The Police Foundation.

Boydstun, J. E. (1975)
San Diego field interrogation: Final report. Washington, D.C.: The Police Foundation.

Brannon, B. C. (1956)
"A report on one-man police cars in Kansas City, Missouri." *Journal of Criminal Law, Criminology, and Police Science* 47:238–252.

Cochran, W. G., and G. M. Cox (1957)
Experimental designs (second edition), pp. 127–243. New York: Wiley.

Davis, E. M., and L. Knowles (1957)
"An evaluation of the Kansas City Preventive Patrol Experiment." *The Police Chief*, June 1957, pp. 22–27.

Fisher, R. A. (1926)
"The arrangement of field experiments." *Journal of the Ministry of Agriculture* 33: 503–513.

Gilbert, J. P., R. J. Light, and F. Mosteller (1975)
"Assessing social innovations: An empirical base for policy." In A. R. Lumsdaine and C. A. Bennet (eds.), *Evaluation and experiment: Some critical issues in assessing social programs*, New York: Academic Press.

Kelling, G. L., T. Pate, D. Dieckman, and C. E. Brown (1974)
The Kansas City Preventive Patrol Experiment: A technical report. Washington, D.C.: The Police Foundation.

Kempthorne, O. (1952)
The design and analysis of experiments, pp. 10–17. New York: Wiley.

Larson , R. C. (1972)
Urban police patrol analysis. Cambridge, Mass.: M.I.T. Press.

Larson, R. C. (1976)
"What happened to patrol operations in Kansas City: A review of the Kansas City Preventive Patrol Experiment." *Journal of Criminal Justice* 267–297.

Pate, T., R. A. Bowers, and R. Parks (1976)
Three approaches to criminal apprehension in Kansas City: An evaluation report. Washington, D.C.: The Police Foundation.

Paterson, H. D. (1951)
"Change-over trials." *Journal of the Royal Statistical Society* (B)12:256–271.

Patterson, H. D. (1952)
"The construction of balanced designs for experiments involving sequences of treatments." *Biometrika* 39:32–48.

Press, S. J. (1971)
"Some effects of an increase in police manpower in the 20th Precinct of New York City." R–704 NYC. New York: New York City RAND Institute.

Press, S. J. (1972)
"Police manpower versus crime." In J. M. Tanur et al. (eds.), *Statistics: A guide to the unknown*, San Francisco: Holden-Day.

Reiss, A. J., Jr. (1967)
"Studies in crime and law enforcement in major metropolitan areas." Report of a research study for the President's Commission on Law Enforcement and the Administration of Justice. University of Michigan.

Reicken, H. R., and R. F. Boruch (eds.) (1974)
Social experimentation. New York: Academic Press.

Schwartz, A. I., and S. N. Clarren (1977)
The Cincinnati Team Policing Experiment: A summary report. Washington, D.C.: The Police Foundation.

Sherman, L. W., C. H. Milton, and T. V. Kelly (1973)
Team policing; Seven case studies. Washington, D.C.: The Police Foundation.

Tien, J. M., R. C. Larson, and J. W. Simon (1976)
An evaluation report: Wilmington Split Force Patrol Program. Cambridge, Mass.: Public Systems Evaluation, Inc.

NCJ-62344

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

Single copies are available at no charge from the National Criminal Justice Reference Service, Box 6000, Rockville, Md. 20850. Multiple copies are for sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402.

National Crime Survey:

Criminal Victimization in the United States (annual):
- Summary Findings of 1977-78 Changes in Crime and of Trends Since 1973, NCJ-61368
- A Description of Trends from 1973 to 1977, NCJ-59898
- 1978 (final report), NCJ-66480
- 1977, NCJ-58725
- 1976, NCJ-49543
- 1975, NCJ-44593
- 1974, NCJ-39467
- 1973, NCJ-34732

The Cost of Negligence: Losses from Preventable Household Burglaries, NCJ-53527

Intimate Victims: A Study of Violence Among Friends and Relatives, NCJ-62319

Crime and Seasonality, NCJ-64818

Criminal Victimization of New York State Residents, 1974-77, NCJ-66481

Criminal Victimization Surveys in 13 American cities (summary report, 1 vol.), NCJ-18471
- **Boston,** NCJ-34818
- **Buffalo,** NCJ-34820
- **Cincinnati,** NCJ-34819
- **Houston,** NCJ-34821
- **Miami,** NCJ-34822
- **Milwaukee,** NCJ-34823
- **Minneapolis,** NCJ-34824
- **New Orleans,** NCJ-34825
- **Oakland,** NCJ-34826
- **Pittsburgh,** NCJ-34827
- **San Diego,** NCJ-34828
- **San Francisco,** NCJ-34829
- **Washington, D.C.,** NCJ-34830

Public Attitudes About Crime (13 vols.):
- **Boston,** NCJ-46235
- **Buffalo,** NCJ-46236
- **Cincinnati,** NCJ-46237
- **Houston,** NCJ-46238
- **Miami,** NCJ-46239
- **Milwaukee,** NCJ-46240
- **Minneapolis,** NCJ-46241
- **New Orleans,** NCJ-46242
- **Oakland,** NCJ-46243
- **Pittsburgh,** NCJ-46244
- **San Diego,** NCJ-46245
- **San Francisco,** NCJ-46246
- **Washington, D.C.,** NCJ-46247

Criminal Victimization Surveys in Chicago, Detroit, Los Angeles, New York, and Philadelphia: A Comparison of 1972 and 1974 Findings, NCJ-36360

Criminal Victimization Surveys in the Nation's Five Largest Cities: National Crime Panel Surveys in Chicago, Detroit, Los Angeles, New York, and Philadelphia, 1972, NCJ-16909

Criminal Victimization Surveys in Eight American Cities: A Comparison of 1971/72 and 1974/75 Findings—National Crime Surveys in Atlanta, Baltimore, Cleveland, Dallas, Denver, Newark, Portland, and St. Louis, NCJ-36361

Crimes and Victims: A Report on the Dayton/San Jose Pilot Survey of Victimization, NCJ-013314

Indicators of Crime and Criminal Justice: Quantitative Studies, NCJ-62349

Applications of the National Crime Survey Victimization and Attitude Data:

Public Opinion About Crime: The Attitudes of Victims and Nonvictims in Selected Cities, NCJ-41336

Local Victim Surveys: A Review of the Issues, NCJ-39973

The Police and Public Opinion: An Analysis of Victimization and Attitude Data from 13 American Cities, NCJ-42018

An Introduction to the National Crime Survey, NCJ-43732

Compensating Victims of Violent Crime: Potential Costs and Coverage of a National Program, NCJ-43387

Crime Against Persons in Urban, Suburban, and Rural Areas: A Comparative Analysis of Victimization Rates, NCJ-53551

Rape Victimization in 26 American Cities, NCJ-55878

Criminal Victimization in Urban Schools, NCJ-56396

National Prisoner Statistics:

Capital Punishment (annual):
- 1978, NCJ-59897
- 1979 advance report, NCJ-67705

Prisoners in State and Federal Institutions on December 31:
- 1978, NCJ-64671
- 1979 advance report, NCJ-66522

Census of State Correctional Facilities, 1974 advance report, NCJ-25642

Profile of State Prison Inmates: Sociodemographic Findings from the 1974 Survey of Inmates of State Correctional Facilities, NCJ-58257

Census of Prisoners in State Correctional Facilities, 1973, NCJ-34729

Census of Jails and Survey of Jail Inmates, 1978, preliminary report, NCJ-55172

Profile of Inmates of Local Jails: Sociodemographic Findings from the 1978 Survey of Inmates of Local Jails, NCJ-65412

The Nation's Jails: A report on the census of jails from the 1972 Survey of Inmates of Local Jails, NCJ-19067

Survey of Inmates of Local Jails, 1972, advance report, NCJ-13313

Uniform Parole Reports:

Parole in the United States (annual):
- 1978, NCJ-58722
- 1976 and 1977, NCJ-49702

Characteristics of the Parole Population, 1978, NCJ-66479

A National Survey of Parole-Related Legislation Enacted During the 1979 Legislative Session, NCJ-64218

Children in Custody: Juvenile Detention and Correctional Facility Census
- 1977 advance report:
  - Census of Public Juvenile Facilities, NCJ-60967
  - Census of Private Juvenile Facilities, NCJ-60968
- 1975 (final report), NCJ-58139
- 1974, NCJ-57946
- 1973, NCJ-44777
- 1971, NCJ-13403

Myths and Realities About Crime: A Nontechnical Presentation of Selected Information from the National Prisoner Statistics Program and the National Crime Survey, NCJ-46249

State and Local Probation and Parole Systems, NCJ-41335

State and Local Prosecution and Civil Attorney Systems, NCJ-41334

National Survey of Court Organization:
- 1977 Supplement to State Judicial Systems, NCJ-40022
- 1975 Supplement to State Judicial Systems, NCJ-29433
- 1971 (full report), NCJ-11427

State Court Model Statistical Dictionary, NCJ-62320

State Court Caseload Statistics:
- The State of the Art, NCJ-46934
- Annual Report, 1975, NCJ-51885
- Annual Report, 1976, NCJ-56599

A Cross-City Comparison of Felony Case Processing, NCJ-55171

Trends in Expenditure and Employment Data for the Criminal Justice System, 1971-77 (annual), NCJ-57463

Expenditure and Employment Data for the Criminal Justice System (annual)
- 1978 Summary Report, NCJ-66483
- 1978 final report, NCJ-66482
- 1977 final report, NCJ-53206

Dictionary of Criminal Justice Data Terminology: Terms and Definitions Proposed for Interstate and National Data Collection and Exchange, NCJ-36747

Justice Agencies in the U.S.: Summary Report of the National Justice Agency List, NCJ-65560

Criminal Justice Agencies in Region
- **1:** Conn., Maine, Mass., N.H., R.I., Vt., NCJ-17930
- **2:** N.J., N.Y., NCJ-17931
- **3:** Del., D.C., Md., Pa., Va., W.Va., NCJ-17932
- **4:** Ala., Ga., Fla., Ky., Miss., N.C., S.C., Tenn., NCJ-17933
- **5:** Ill., Ind., Mich., Minn., Ohio, Wis., NCJ-17934
- **6:** Ark., La., N.Mex., Okla., Tex., NCJ-17935
- **7:** Iowa, Kans., Mo., Nebr., NCJ-17936
- **8:** Colo., Mont., N.Dak., S.Dak., Utah, Wyo., NCJ-17937
- **9:** Ariz., Calif., Hawaii, Nev., NCJ-15151
- **10:** Alaska, Idaho, Oreg., Wash., NCJ-17938

Utilization of Criminal Justice Statistics Project:

Sourcebook of Criminal Justice Statistics 1979 (annual), NCJ-59679

Public Opinion Regarding Crime, Criminal Justice, and Related Topics, NCJ-17419

New Directions in Processing of Juvenile Offenders: The Denver Model, NCJ-17420

Who Gets Detained? An Empirical Analysis of the Pre-Adjudicatory Detention of Juveniles in Denver, NCJ-17417

Juvenile Dispositions: Social and Legal Factors Related to the Processing of Denver Delinquency Cases, NCJ-17418

Offender-Based Transaction Statistics: New Directions in Data Collection and Reporting, NCJ-29645

Sentencing of California Felony Offenders, NCJ-29646

The Judicial Processing of Assault and Burglary Offenders in Selected California Counties, NCJ-29644

Pre-Adjudicatory Detention in Three Juvenile Courts, NCJ-34730

Delinquency Dispositions: An Empirical Analysis of Processing Decisions in Three Juvenile Courts, NCJ-34734

The Patterns and Distribution of Assault Incident Characteristics Among Social Areas, NCJ-40025

Patterns of Robbery Characteristics and Their Occurrence Among Social Areas, NCJ-40026

Crime-Specific Analysis:
- The Characteristics of Burglary Incidents, NCJ-42093
- An Empirical Examination of Burglary Offender Characteristics, NCJ-43131
- An Empirical Examination of Burglary Offenders and Offense Characteristics, NCJ-42476

Sources of National Criminal Justice Statistics: An Annotated Bibliography, NCJ-45006

Federal Criminal Sentencing: Perspectives of Analysis and a Design for Research, NCJ-33683

Variations in Federal Criminal Sentences: A Statistical Assessment at the National Level, NCJ-33684

Federal Sentencing Patterns: A Study of Geographical Variations, NCJ-33685

Predicting Sentences in Federal Courts: The Feasibility of a National Sentencing Policy, NCJ-33686