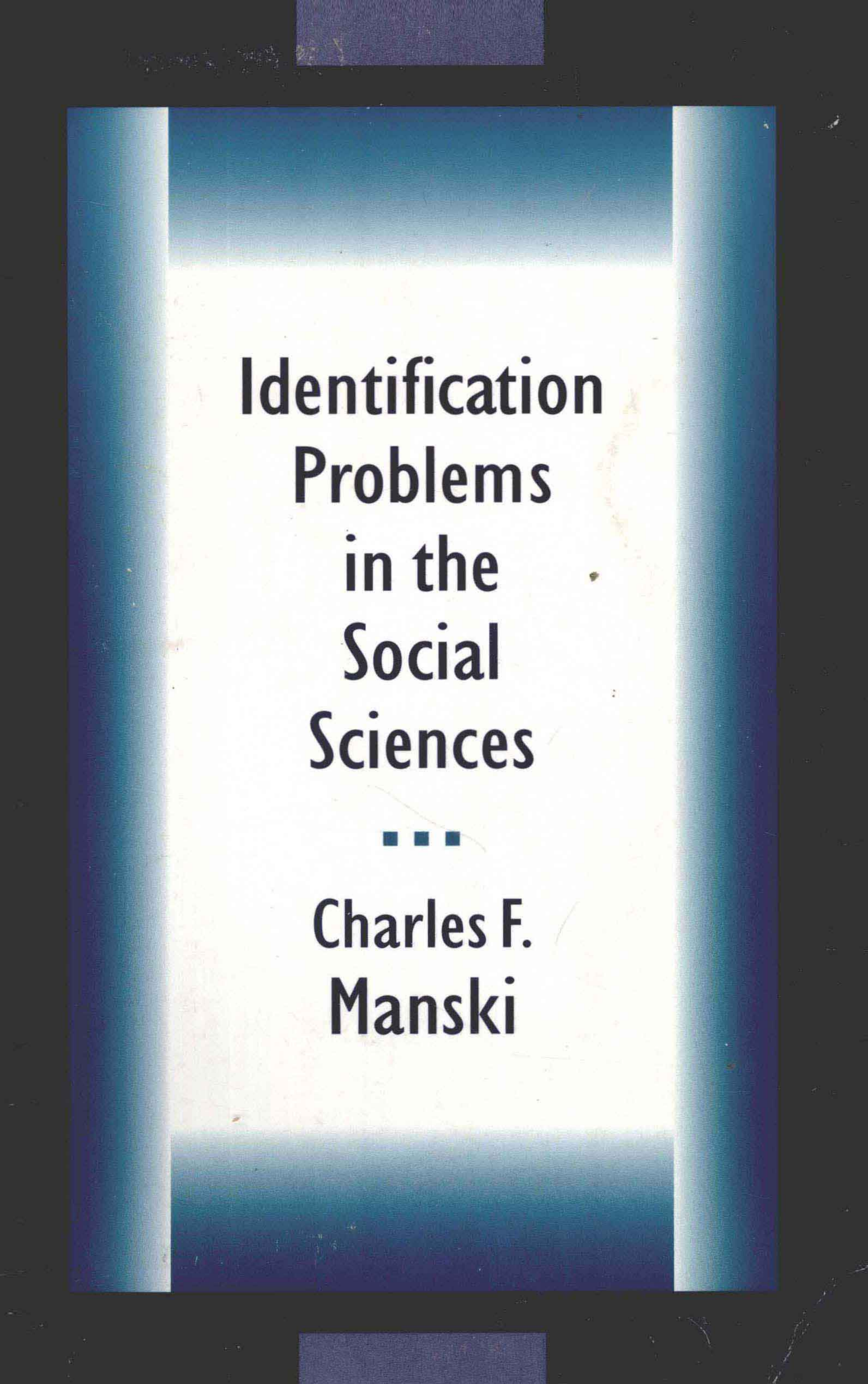

Identification Problems in the Social Sciences

* * *

Charles F. Manski

Identification Problems in the Social Sciences

Charles F. Manski

HARVARD UNIVERSITY PRESS
CAMBRIDGE, MASSACHUSETTS
LONDON, ENGLAND 1995

Copyright © 1995 by the President and Fellows of Harvard College
All rights reserved
Printed in the United States of America

This book is printed on acid-free paper.

Library of Congress Cataloging-in-Publication Data

Manski, Charles F.
Identification problems in the social sciences / Charles F. Manski.
p. cm.
Includes bibliographical references and index.
ISBN 0-674-44283-0
1. Social sciences—Statistical methods. 2. Estimation theory. I. Title.
HA29.M2465 1995
300'.5'51—dc20
94-29313
CIP

Preface

As a new Ph.D. student almost a quarter of a century ago, I was introduced to the study of identification by Franklin Fisher, who devoted a large part of the required course in econometrics to the problem of identification of linear simultaneous equations models. Some of my research from the early 1970s through the middle 1980s concerned the identification of discrete response models. But the perspective on identification presented in this book took shape more recently.

In the spring of 1987, Irving Piliavin stimulated me to examine the basic probabilistic structure of the selection problem associated with survey nonresponse. Later, after assuming the directorship of the Institute for Research on Poverty, I found myself grappling with the selection and mixing problems that arise in the evaluation of social programs and with the reflection problem faced in the study of neighborhood effects. My research on these subjects yielded a stream of new findings, and I gradually came to see a set of themes that warranted development into a book. Throughout this period of discovery I had the immense good fortune of being able to discuss ideas with, and receive encouragement from, my colleague Arthur Goldberger.

In 1992 I was invited by the American Sociological Association to present some of my work on identification at its annual meeting and to write a synthesis article for *Sociological Methodology* (Manski, 1993a). I am grateful to Clifford Clogg, Robert Mare, Peter Marsden, and William Mason for their roles in creating this opportunity for me to organize the disparate strands of my research into a more coherent whole.

With the article completed, writing a book seemed less daunting. The perfect environment for composing a first draft was provided by the Center for Advanced Studies in the Behavioral Sciences, where I was able to work with close to total concentration during the academic year 1992–93. I appreciate the financial support that made this possible, obtained from National Science Foundation grant SES90-22192 and from the University of Wisconsin Graduate School.

Arthur Goldberger, Jeffrey Dominitz, and Elizabeth Uhr have offered many constructive substantive comments and editorial suggestions on a draft of the manuscript. It has been a pleasure to work with Michael Aronson and Elizabeth Gretz of Harvard University Press.

June 1994

此为试读，需要完整PDF请访问：www.ertongbook.com

Contents

Preface *ix*

Introduction *1*

Identification *3*

Tolerating Ambiguity 7

1 *Extrapolation* *10*

1.1. Predicting Criminality *10*

1.2. Probabilistic Prediction *11*

1.3. Inferring Conditional Distributions from Random-Sample Data *13*

1.4. Prior Distributional Information *16*

1.5. Predicting High School Graduation *18*

2 *The Selection Problem* *21*

2.1. The Nature of the Problem *21*

2.2. Identification from Censored Samples Alone *23*

2.3. Bounding the Probability of Exiting Homelessness *27*

2.4. Prior Distributional Information *31*

2.5. Identification of Treatment Effects *37*

2.6. Information Linking Outcomes across Treatments *43*

2.7. Predicting High School Graduation If All Families Were Intact *47*

3 *The Mixing Problem in Program Evaluation* *51*

3.1. The Experimental Evaluation of Social Programs *51*

3.2. Variation in Treatment *54*

3.3. The Perry Preschool Project 58
3.4. Identification of Mixtures Using Only Knowledge of the Marginals 62
3.5. Restrictions on the Outcome Distribution 65
3.6. Restrictions on the Treatment Policy 67
3.7. Identifying Combinations of Assumptions 72

4 *Response-Based Sampling* 73
4.1. The Odds Ratio and Public Health 74
4.2. Bounds on Relative and Attributable Risk 78
4.3. Information on Marginal Distributions 81
4.4. Sampling from One Response Stratum 82
4.5. General Binary Stratifications 85

5 *Predicting Individual Behavior* 88
5.1. Revealed Preference Analysis 88
5.2. How Do Youth Infer the Returns to Schooling? 95
5.3. Analysis of Intentions Data 98

6 *Simultaneity* 110
6.1. "The" Identification Problem in Econometrics 110
6.2. The Linear Market Model 112
6.3. Equilibrium in Games 116
6.4. Simultaneity with Downward-Sloping Demand 119

7 *The Reflection Problem* 127
7.1. Endogenous, Contextual, and Correlated Effects 127
7.2. A Linear Model 129
7.3. A Pure Endogenous Effects Model 133
7.4. Inferring the Composition of Reference Groups 134
7.5. Dynamic Analysis 135

Notes 139

References 155

Name Index 165

Subject Index 169

Introduction

Suppose that you observe the almost simultaneous movements of a man and his image in a mirror. Does the mirror image cause the man's movements or reflect them? If you do not understand something of optics and human behavior, you will not be able to tell.

A like inferential problem arises if you try to interpret the common observation that individuals belonging to the same group tend to behave similarly. Two hypotheses often advanced to explain this phenomenon are

> *endogenous effects,* wherein the propensity of an individual to behave in some way varies with the prevalence of that behavior in the group

and

> *correlated effects,* wherein individuals in the same group tend to behave similarly because they face similar institutional environments or have similar individual characteristics.

Similar behavior within groups could stem from endogenous effects; for example, group members could experience pressure to conform to group norms. Or group similarities might reflect correlated effects; for example, persons with similar characteristics might choose to associate with one another. If you do not know something about the way groups form and the way their members interact, then you cannot distinguish between these hypotheses.

Why might you care whether observed patterns of behavior are generated by endogenous effects, by correlated effects, or in some other way? A good practical reason is that different processes have differing implications for public policy. For example, understanding how students interact in classrooms is critical to the evaluation of many aspects of educational policy, from ability tracking to class-size standards to racial integration programs.

Suppose that, unable to interpret observed patterns of behavior, you seek the expert advice of two social scientists. One, perhaps a sociologist, asserts that pressure to conform to group norms makes the individuals in a group tend to behave similarly. The other, perhaps an economist, asserts that persons with similar characteristics choose to associate with one another. Both assertions are consistent with the empirical evidence, so you have no objective way to assess their validity. All you can do is judge the persuasiveness of the arguments offered. If you are persuaded by one social scientist more than by the other, it is only because one is a more skilled advocate for his or her position.

The situation just depicted is frustratingly familiar. There are many good reasons to want to know why the members of groups tend to behave similarly. Nevertheless, researchers have been unable to resolve the question.

Social scientists rarely seem able to settle questions of public concern. Consider, for example, the never-ending American debate about Aid to Families with Dependent Children (AFDC), the social insurance program commonly referred to as welfare. A central issue is the effect of AFDC on marriage, fertility, and labor supply behavior. Almost everyone has an opinion on the matter, but the opinions vary widely. Researchers have worked hard to understand how individuals respond to the incentives embedded in AFDC (see Moffitt, 1992a, and Manski and Garfinkel, 1992). But disagreements about the behavioral effects of welfare persist.

Social scientists have similarly worked hard to understand how the threat of punishment deters crime (see Blumstein, Cohen, and Nagin, 1978), how class size and composition affect student learning (see Hanushek, 1986, and Gamoran, 1992), how neighborhoods affect their inhabitants (see Jencks and Mayer, 1989), and how family structure affects children's outcomes (see Hayes and Hofferth, 1987, and

McLanahan and Sandefur, 1994). In these and so many other areas, progress is painfully slow. Research accumulates but does not converge toward a consensus.

Why do social scientists so often provide conflicting perspectives on questions of public interest? The core problem is the inherent difficulty of studying human behavior. The conclusions that can be drawn from any analysis are determined by the assumptions made and by the data brought to bear. The range of plausible assumptions about human behavior is wide. The available data are limited to observations that can be made without undue intrusion.[1] Researchers combining limited data with different maintained assumptions can, and often do, reach different logically valid conclusions.

A contributing problem is the frequent failure of social scientists to face up to the difficulty of their enterprise. Researchers sometimes do not recognize that the interpretation of data requires assumptions. Researchers sometimes understand the logic of scientific inference but ignore it when reporting their own work. The scientific community rewards those who produce strong novel findings. The public, impatient for solutions to its pressing concerns, rewards those who offer simple analyses leading to unequivocal policy recommendations. These incentives make it tempting for researchers to maintain assumptions far stronger than they can persuasively defend, in order to draw strong conclusions.

Identification

Methodological research is concerned with the logic of scientific inference. The objective is to learn what conclusions can and cannot be drawn given specified combinations of assumptions and data.

Empirical researchers usually enjoy learning of positive methodological findings. Particularly pleasing are results showing that conventional assumptions, when combined with available data, imply stronger conclusions than previously recognized. Negative findings are less welcome. Researchers are especially reluctant to learn that, given the available data, some conclusion of interest cannot be drawn unless strong assumptions are invoked. Be this as it may, both positive and negative findings are important to the advancement of science.

For over a century, methodological research in the social sciences has made productive use of statistical theory.[2] One supposes that the empirical problem is to infer some feature of a population described by a probability distribution and that the available data are observations extracted from the population by some sampling process. One combines the data with assumptions about the population and the sampling process to draw statistical conclusions about the population feature of interest.

Working within this framework, it is useful to separate the inferential problem into statistical and identification components. Studies of identification seek to characterize the conclusions that could be drawn if one could use the sampling process to obtain an unlimited number of observations. Studies of statistical inference seek to characterize the generally weaker conclusions that can be drawn from a finite number of observations.

Statistical and identification problems limit in distinct ways the conclusions that may be drawn in empirical research. Statistical problems may be severe in small samples but diminish in importance as the sampling process generates more observations. Identification problems cannot be solved by gathering more of the same kind of data. These inferential difficulties can be alleviated only by invoking stronger assumptions or by initiating new sampling processes that yield different kinds of data.

To illustrate the distinction, consider Figures I.1 and I.2. Both figures concern a researcher who wants to predict a random variable y conditional on a specified value for some other variable x. The available data are a random sample of observations of (y, x) drawn from a population in which x only takes values in the intervals $[0, 4]$ and $[6, 8]$. In Figure I.1, the researcher has 100 observations of (y, x) and uses these data to draw a confidence interval for the expected value of y conditional on x. In Figure I.2, the researcher has 1000 observations and similarly draws a confidence interval.

Inspect the intervals $[0, 4]$ and $[6, 8]$. The wide confidence interval of Figure I.1 is a statistical problem. Gathering more data permits one to estimate the conditional expectation of y more precisely and so narrows the confidence interval, as shown in Figure I.2. Now inspect the interval $(4, 6)$. The confidence interval is infinitely wide in Figure I.1 and remains so in Figure I.2. This is an identification problem. The sampling process generates no observations in the interval $(4, 6)$,

此为试读，需要完整PDF请访问：www.ertongbook.com

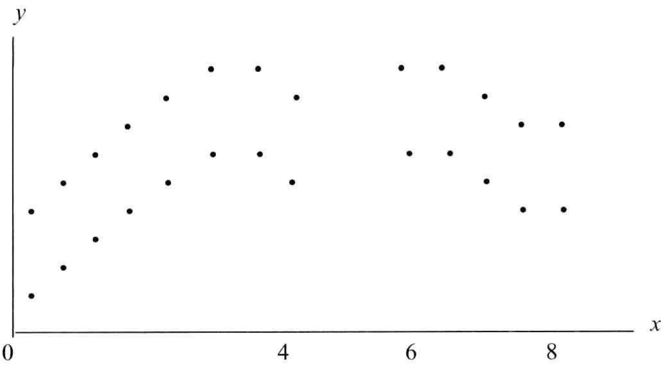

Figure I.1 Confidence interval based on 100 observations.

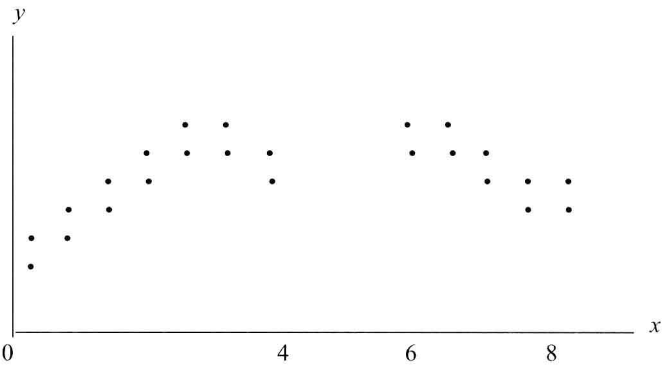

Figure I.2 Confidence interval based on 1000 observations.

so the researcher cannot possibly infer the expected value of y there.

It is tempting to connect the segments found in intervals [0, 4] and [6, 8] to cover (4, 6) as well. The researcher could do this if he or she were willing to assume that the expected value of y varies linearly with x over the entire interval [0, 8]. But the researcher might assume instead that the expected value of y remains constant as x varies between 4 and 6; perhaps it stays midway between its values at x = 4 and x = 6. With the available data, there is no objective way to extrapolate the segments.

Extrapolation is a particularly common and familiar identification problem. Distinguishing endogenous effects from correlated effects is another identification problem. A classic identification problem in economics is confronted when one tries to use data on market-equilibrium prices and quantities transacted to infer the supply behavior of firms and the demand behavior of consumers.

The American debate about the incentive effects of welfare also stems from an identification problem. It has been observed that the marriage, fertility, and labor supply behavior of welfare recipients tends to differ from that of nonrecipients. Unless one knows a good bit about human behavior, one cannot tell whether welfare programs influence recipients to behave in certain ways or whether individuals who behave in those ways choose to receive welfare.

These and other identification problems in the social sciences are the subject of this book. Empirical research must, of course, contend with statistical issues as well as with identification problems. Nevertheless, the two types of inferential difficulties are sufficiently distinct for it to be fruitful to study them separately. The study of identification logically comes first. Negative identification findings imply that statistical inference is fruitless: it makes no sense to try to use a sample of finite size to infer something that could not be learned even if a sample of infinite size were available. Positive identification findings imply that one should go on to study the feasibility of statistical inference.

The usefulness of separating the identification and statistical components of inference has long been recognized. Koopmans (1949, p. 132) put it this way in the article that introduced the term *identification* into the literature:

> In our discussion we have used the phrase "a parameter that can be determined from a sufficient number of observations." We shall now define this concept more sharply, and give it the name *identifiability* of a parameter. Instead of reasoning, as before, from "a sufficiently large number of observations" we shall base our discussion on a hypothetical knowledge of the probability distribution of the observations, as defined more fully below. It is clear that exact knowledge of this probability distribution cannot be derived from any finite number of observations. Such knowledge is the limit approachable but not attainable by extended observation. By hypothesizing nevertheless the full availabil-

> ity of such knowledge, we obtain a clear separation between problems of statistical inference arising from the variability of finite samples, and problems of identification in which we explore the limits to which inference even from an infinite number of observations is suspect.

I focus on identification problems that arise when we attempt to make conditional predictions; that is, when we attempt to answer questions of the form "What if?" This book examines the conditional predictions that can and cannot be made given specified assumptions and empirical evidence.

Not all research is concerned with prediction, so focusing the book on prediction does influence its content. Scientists sometimes conduct research as an effort to improve our "understanding" of a subject, and they argue that this is a worthwhile objective even if there are no interesting implications for prediction. For example, in a text on statistical methods in epidemiology, Fleiss (1981, p. 92) states that the retrospective studies of disease that are a staple of medical research do not yield policy-relevant predictions and so are "necessarily useless from the point of view of public health." Nevertheless, the author goes on to say that "retrospective studies are eminently valid from the more general point of view of the advancement of knowledge." Justifications of this sort will not be found in the present book.

The book contains seven chapters. Chapters 1 through 4 examine observational problems that arise in all scientific work, whether in the social or the natural sciences. The central concern is to explain how the sampling process affects the predictions that can be made. Chapters 5 through 7 examine identification problems particular to the prediction of individual behavior and social interactions. A recurring interest is to compare the distinct approaches taken by different social science disciplines.

Chapters 1 and 2 cover basic material that is often referred to subsequently. Chapters 3 through 7 can be read independent of one another. Although the book examines a wide range of identification problems, it makes no pretense of being encyclopedic. Much of the book draws on my own recent research.

Tolerating Ambiguity

In addition to analyzing specific identification problems, this book develops a general theme. Social scientists and policymakers alike seem driven to draw sharp conclusions, even when these can be generated only by imposing much stronger assumptions than can be defended. We need to develop a greater tolerance for ambiguity. We must face up to the fact that we cannot answer all of the questions that we ask.

The pressure to produce answers, without qualifications, seems particularly intense in the environs of Washington, D.C. A perhaps apocryphal, but quite believable, story circulates about an economist's attempt to describe his uncertainty about a forecast to President Lyndon B. Johnson. The economist presented his forecast as a likely range of values for the quantity under discussion. Johnson is said to have replied, "Ranges are for cattle. Give me a number."

A thoughtful news magazine article written at the height of the 1992 presidential campaign compared the predictions made by participants in the campaign to those of the fictional prophet Carnac played by the popular television comedian Johnny Carson. Whitman (1992, p. 36) wrote: "An unpublished commandment of presidential campaigns can be stated simply: Thou shalt never say, 'I don't have an answer to this crisis.' Instead, both candidates and pundits feel obliged to act out their Carnac complex in election years, professing, like Johnny Carson's famous seer, to be all seeing and all knowing."

Social scientists should recognize how hard it is to provide firm answers to complex social questions. Some scientific conventions, notably the reporting of sampling confidence intervals in statistical analysis, do promote the expression of uncertainty. But other scientific practices encourage misplaced certainty.

One problem is the fixation of social scientists on point identification of parameters. Empirical studies typically seek to learn the value of some parameter characterizing the population of interest. The conventional practice is to invoke assumptions strong enough to identify the exact value of this parameter. Even if these assumptions are implausible, they are defended as necessary for inference to proceed. Yet identification is not an all-or-nothing proposition. Weaker and more plausible assumptions often suffice to bound parameters in informative ways.

A larger problem is the common view that a scientist should choose one hypothesis to maintain, even if that means discarding others that are a priori plausible and consistent with the available empirical evidence. This view was expressed in an influential methodological essay written by Milton Friedman over forty years ago. Friedman (1953) placed prediction as the central objective of science, writing, "The ultimate goal of a positive science is the development of a 'theory' or 'hypothesis' that yields valid and meaningful (i.e. not truistic) predictions about phenomena not yet observed" (p. 5). He went on to say, "The choice among alternative hypotheses equally consistent with the available evidence must to some extent be arbitrary, though there is general agreement that relevant considerations are suggested by the criteria 'simplicity' and 'fruitfulness,' themselves notions that defy completely objective specification" (p. 10).

I do not see why a scientist must choose one hypothesis to hold, especially when this requires the use of "to some extent . . . arbitrary" criteria. Indeed, using arbitrary criteria to choose a single hypothesis has an obvious drawback in predicting phenomena not yet observed: one may have made a wrong choice. Social scientists are notorious for making sharp predictions that turn out to be incorrect. The credibility of social science would be higher if we would strive to offer predictions under the range of plausible hypotheses that are consistent with the available evidence.

1

Extrapolation

1.1. Predicting Criminality

In 1982 the RAND Corporation released a study of criminal behavior as reported in 1978 by a sample of 2,200 prison and jail inmates in California, Michigan, and Texas (Chaiken and Chaiken, 1982, and Greenwood, 1982). Most respondents reported that they had committed five or fewer crimes per year in the period before their current arrest and conviction. A small group reported much higher rates of crime commission, in some cases more than one hundred per year.

The researchers found that, using limited information on a person's past convictions, drug use, and employment, they could predict well whether that person had been a high-rate offender. This finding suggested to at least part of the research team that *selective incapacitation* should be encouraged as a crime-fighting tool (Greenwood, 1982). Selective incapacitation calls for the sentencing of convicted criminals to be tied to predictions of their future criminality. Those with backgrounds that predict high rates of offenses would receive longer prison terms than those with other backgrounds.

The RAND study generated much controversy, especially when a prediction approach devised by Greenwood found its way into legislative proposals for selective incapacitation (see Blackmore and Welsh, 1983, and Blumstein et al., 1986). Some of the controversy concerned the normative acceptability of selective incapacitation, but much of it concerned the validity of extrapolations from the RAND findings.

The findings characterize the empirical association between background and reported crime commission within one cohort of inmates imprisoned in three states under the sentencing policies then in effect.

此为试读，需要完整PDF请访问：www.ertongbook.com