Journal of Quantitative Criminology, Vol. 12, No. 1, 1996

From Poisson to the Present: Applying Operations Research to Problems of Crime and Justice

Michael D. Maltz[1]

In the 1830s Siméon-Denis Poisson developed the distribution that bears his name, basing it on the binomial distribution. He used it to show how the inherent variance in jury decisions affected the inferences that could be made about the probability of conviction in French courts. In recent years there have been a number of examples where researchers have either ignored or forgotten this inherent variance, and how operations research, in particular mathematical modeling, can be used to incorporate this variance in analyses. These are described in this paper, as well as other contributions made by operations research to the study of crime and criminal justice.

KEY WORDS: operations research; mathematical modeling; statistical modeling; Poisson distribution; crime; criminal justice.

1. INTRODUCTION

The application of operations research (OR), in particular mathematical and statistical modeling, to crime, justice, and law enforcement is a relationship of long standing, one that has benefited both OR and criminal justice practice over the years. In terms of statistical theory, Poisson's analysis of criminal justice data led to his derivation of

- the Poisson distribution (Stigler, 1986, p. 183)[2] and
- the law of large numbers (Hacking, 1990, p. 100).

Criminal justice practice has profited to an even greater extent from the mutual association. Mathematical and statistical modeling of criminal justice

[1]Department of Criminal Justice (and Department of Information and Decision Sciences), University of Illinois at Chicago, 1007 West Harrison Street, Chicago, Illinois 60607-7140.

[2]Poisson derived this distribution (as a limiting case of the binomial distribution) in the book, *Recherches sur la probabilité des jugements en matière criminelle et en matière civile.*

data has led to

- the treatment of homicide of young urban African–American males as an epidemic,
- enhanced fairness of the jury system,
- improved police scheduling,
- improvements in evaluating the effectiveness of correctional programs,
- increased knowledge of offender behavior,
- a better understanding of the dynamics of business-oriented crime, and
- improved methods of forecasting prisoner populations.

In short, the application of OR in this field has been, by and large, a success. That this has been the case should not be surprising. The criminal justice system is an information-intensive public bureaucracy (or, rather, group of bureaucracies) that runs on paper and generates mountains of data. Its input includes reports that are produced by police officers; they are added up to generate statistics; offenders' records are summarized on "rap sheets;" arrest reports lead to prosecutors' records, which lead to trial transcripts; probation and parole officers prepare reports on their clients' behavior; correctional institutions track the inmates in their care. The system's output is not a tangible product, but decisions: where to patrol, who are probable suspects, who is to be arrestsed, who is adjudged guilty, what type of sentence to impose, who should be released. Thus, although not all of the information is accessible to researchers, the system is overflowing with data, providing excellent opportunities for studies aimed at detecting patterns in the data: how offenders flow through the system; how (and how often) they commit crimes; how the agencies deal with finding and processing offenders.

There are also excellent opportunities for mathematical modeling, especially when it comes to inferring patterns of offender behavior. Unlike the officials in the criminal justice system, who react to offenders, the offenders themselves do not ordinarily provide researchers with records of their activity.[3] We have to assess their behavior from the rather limited information we do have about their activity, most often the crimes for which they have been arrested. It is in that context that mathematical models of offender behavior can be of great benefit.

1.1. Scope

This paper deals for the most part with mathematical models of offender and system behavior, and especially those based on the stochastic, or

[3] Ethnographers obtain such information (see, e.g., Johnson *et al.*, 1985), but usually not in ways that would permit a mathematical model to be constructed.

random, aspects of criminal justice data. Not all OR applications are covered; for example, police resource allocation methods and models (e.g., Larson, 1972) are not discussed—for a recent review of these applications, see Swersey (1994). And considerations of space precluded the review of OR studies of police, juries, and offense switching. Nor does this paper cover applications of the more standard statistical techniques (i.e., those found in the SPSS, SAS, BMDP, or SYSTAT software packages), such as time series, regression, logit, LISREL, and the prediction techniques based on them; those who are interested in such techniques may wish to consult Finkelstein (1978), Greenberg (1979), Saks and Baron (1980), Monahan and Walker (1984), Farrington and Tarling (1985), Copas and Tarling (1986), DeGroot *et al.* (1986), S. Gottfredson and D. Gottfredson (1986), or M. Gottfredson and D. Gottfredson (1987).

Not all topics are covered to the same depth. In part, this is due to different priorities accorded different topics; in part, to the incomplete knowledge of the author. It is not always possible to be sure of covering all sources of published material, so there are doubtless some gaps in this review.[4]

This is not the first time the application of OR to crime and justice has been reviewed. Tarling (1986) provides a relatively recent overview of the application of statistical models and methods in criminology, and different OR or mathematical modeling techniques are compiled by Chaiken *et al.* (1975), Broustein and Kamrass (1976), Greenberg (1979), and Fox (1981a, b). An excellent, although slightly dated, monograph on the application of information theory in criminology and criminal justice is found in Willmer (1970).

The symbols used in this paper are not always the same as those used in the cited works. In a rapidly developing field, different authors writing on the same topic often use different symbols to represent the same variable. In the interests of clarity, I have taken some liberty with the notation used in the papers and employ symbols that have in some sense become standard, so that the reader is not confused (in this paper, at least) by having one symbol defined in two different ways, or two symbols to denote the same phenomenon.

[4]For the most part, papers, reports, and dissertations not in general circulation are not included in this review. Although citations to work outside of North America are few in this review, the reason may be due in part to there not being as much quantitative research in criminology and criminal justice in other countries (Cohn and Farrington, 1990); see Farrington *et al.* (1996) for a more recent review of European quantitative research in criminal justice and criminology.

The beginnings of the intersection between OR and issues of crime and justice are rooted in the works of two nineteenth-century European statisticians, Adolphe Quetelet and Siméon-Denis Poisson, discussed briefly below (see also Pollock and Maltz, 1994). But the greatest attention is given to the work that has taken place more recently, primarily in the past 30 years, and for the most part, in the published literature. Much of this work was initiated by the President's Crime Commission in 1967, discussed in Section 2.3. Subsequent sections describe the research that that effort spawned, in studies of criminal justice systems models, criminal careers, incapacitation, recidivism, deterrence, population forecasts, and crime patterns.

1.2. OR Models and Statistical Methods

Before describing how OR models have been used, it is worth explaining the limitations and shortcomings of mathematical models. For the most part, the models start out as fairly simple, and the assumptions that are made, are made explicit; for example, that an offender (or all offenders, for that matter) has the same rate of commission of offenses. Cynical researchers might then discount the results of such models by pointing out the obvious: that all offenders do not have the same offense rates, and may change their rates over their lifetimes, and therefore, what good are models like these? Of course, they then go on to collect data and analyze them using methods based on models developed by Pearson or Cox or Jøreskog and Sørbom or Bryk and Raudenbush, models that may have been developed for problems very different than the ones they have at hand—and all too often they either ignore the models' assumptions or pay lip service to them (Maltz, 1994a). Who among these researchers could justify the "simplifying assumption" that offending rates could be estimated by adding 1 part age at first arrest to 2 parts parental income to 3 parts median census tract income?

The response that often comes back is that these models are "robust," a word which Safire (1995, p. 20) notes has been used to characterize "a result that is approximately correct despite the falsity of assumptions on which it is based." Thus, the pot is calling the kettle black, often without realizing that this is the case. In fact, simple models can provide a great deal of insight into complex problems, and when the assumptions are made explicit those assumptions can be tested (Hodges, 1991).

This paper shows how the standard statistical analyses often fall far short of correct, and how techniques based on mathematical models can often tell a far truer story, whether the issue at hand is the conviction rate in nineteenth-century France or the deterrent effect of executions on homicide or of incarceration on delinquents.

2. A HISTORICAL PERSPECTIVE

To set the stage for the 19th century work, and for the paper in general, I want to discuss two different views of a well-known statistical regularity, the male–female birth ratio. It was noticed by a number of early mathematicians and prompted John Arbuthnot to write a monograph in 1710 entitled, "An Argument for Divine Providence, taken from the Regularity Observed in the Birth of Both Sexes" (Daston, 1988). In other words, he equated this particular regularity as proof of the existence of God, for why else would the two sexes be almost equally likely? Of course, a statistician would point out that coin tossing provides the same regularity, but perhaps one would not attribute this regularity to divine providence. Yet similar ascriptions of natural laws to statistical regularities continue to this day.

2.1. The Nineteenth Century

According to Chevalier (1973), crime was one of the most pressing social problems in Paris in the early nineteenth century. The misery of the working classes was captured in Hugo's *Les Misérables* as well as in contemporary data: Chevalier reports how concerned the police were about fluctuations in the price of bread because "its curve followed the curve of all the ills that Paris was heir to" (p. 262). The misery was also reflected in French statistics that described the relationship between literacy and crime, a topic that is still a matter of concern. For example, in 1828 Taillander wrote that "67 out of 100 [prisoners] are able neither to read nor write. What stronger proof could there be that ignorance, like idleness, is the mother of all vices?" (quoted by Porter, 1986).

During this same period the Société Française de Statistique Universelle was deeply involved in the debate over the death penalty. In the first issue of its bulletin (1830–1831) the Société announced a prize for the best analysis calling for its abolition:

> By attaining statistical certainty that fewer crimes are committed where the penalty of death has been abolished, the influence that a gentler and more genuinely philosophical legislation exerts on the criminality of human actions can better be appreciated.

However, this call in France for a "kinder and gentler nation" was not heeded at that time (Porter, 1986).

The collection of criminal justice statistics in France began in 1825 (Daston, 1988; p. 286). The availability of judicial data permitted the development of tools for statistical analysis of social problems. Adolphe Quetelet (1835) initiated this movement; he noted the relative constancy in France

Table I. French Conviction Data used by Quetelet

Year	Number accused	Number convicted	Conviction rate
1825	7234	4594	0.6351
1826	6988	4348	0.6222
1827	6929	4236	0.6113
1828	7396	4551	0.6153
1829	7373	4475	0.6069
1830	6962	4130	0.5932

of annual conviction rates (the fraction of the population convicted annually) and analyzed these data using the newer mathematical tools then in current use in astronomy. In finding the same regularity in crime statistics that was found in astronomical observations, he argued that, just as there was a true location of a star (despite the variance in the location measurements), there was a true level of criminality: he posited the construct of *l'homme moyen* (the "average man") and, moreover, *l'homme moyen moral.* Quetelet asserted that the average man had a statistically constant "penchant for crime," one that would permit the "social physicist" to calculate a trajectory over time that "would reveal simple laws of motion and permit prediction of the future" (Gigerenzer *et al.*, 1989).

Both Quetelet and Siméon-Denis Poisson analyzed French conviction data. Table I (excerpted from Stigler, 1986, p. 189), shows some of these data.[5]

Quetelet noted that there was a general downward trend in the proportion of trials that resulted in conviction. He interpreted this as meaning that the penchant for crime was decreasing, which was based in part on his using incorrect numbers for 1825 (7234 and 4594, respectively; see Fig. 1), but also because he did not consider the stochastic aspects of the data. This can be seen more clearly in Poisson's analysis of the correct data.

After incorporating the correct data (6652 and 4037—see Fig. 2), Poisson took a different analytic view of these statistics (Stigler, 1986, p. 186). First, he developed a model of the trial process that assumed that jury behavior was constant over time, that the jury members made their decisions independently, and that each had the same probability (u) of reaching the correct verdict. Then the probability that no more than k of N jurors will decide correctly is

$$B(k, N, u) = \sum_{i=0}^{k} \binom{N}{i} u^i (1-u)^{N-i}$$

[5]This description is based largely on Stigler's account.

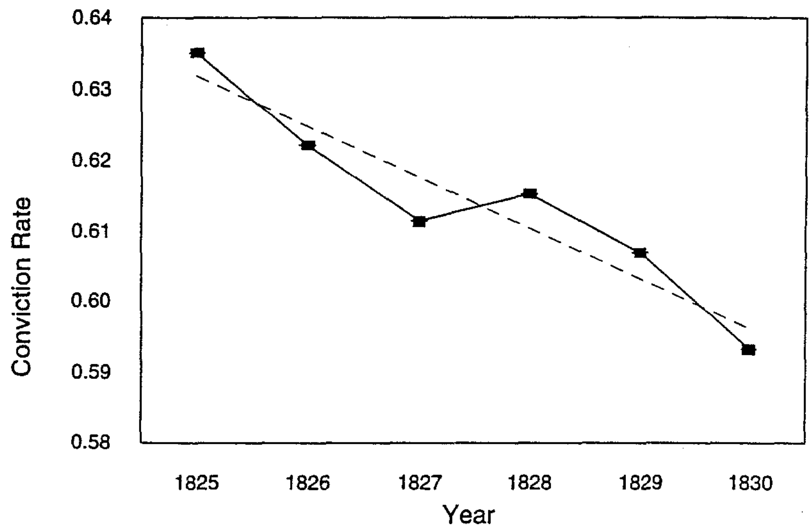


Fig. 1. Quetelet's analysis of French conviction data.

In France before 1831, a majority of 7 of the 12 jurors was required for a determination of guilt, so the probability of conviction of a guilty defendant is $B(5, 12, 1-u)$ and that of an innocent defendant is $B(5, 12, u)$. If G is the probability that the defendant is in fact guilty, then the probability of conviction P_C is

$$P_C = GB(5, 12, 1-u) + (1-G)B(5, 12, u)$$

Poisson assumed that G and u are constant over time, so P_C is also constant over time. He could therefore model the juries' decisions as a binomial process and determine the stochastic variation in the data. Based on his estimates (Fig. 2 shows the estimate ± 2 standard deviations), he found that the variation in rates over time was "not so large as to support a belief that there has been a notable change in the causes" (quoted by Stigler, 1986, p. 190). In other words, the year-to-year differences were small compared to the expected variation, or were not what we would now call *statistically significant*. This was doubtless the first application of the concept of statistical significance, one that in this case did not falsify the null hypothesis of no change. [In a subsequent analysis comparing the French statistics with those of the Department of the Seine (Paris), however, Poisson did discern a significant difference in the conviction rate. As Hacking (1990, p. 101) put it, Poisson thought he had "detected a real change in court behavior. 1830 was a year of revolution. Either the court was bringing too many criminals

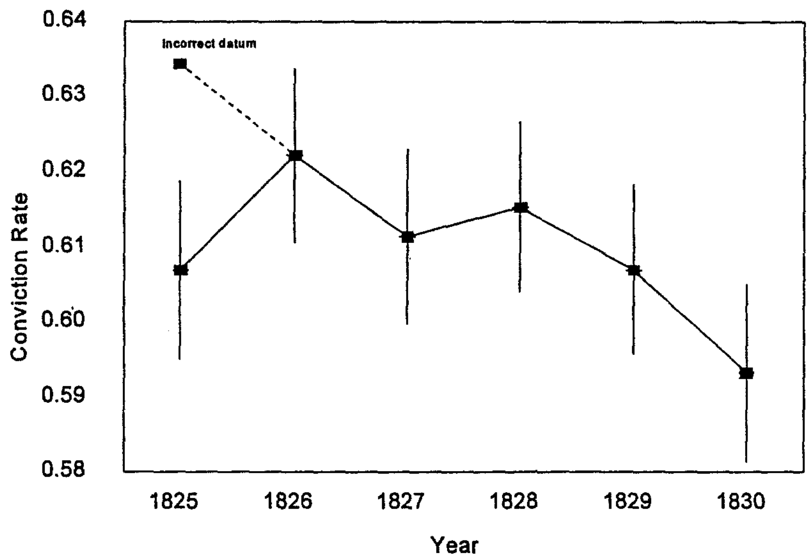


Fig. 2. Poisson's analysis of (corrected) French conviction data.

to trial in a draconian attempt to maintain order, or jurors were, in revolutionary spirit, declining to bring in convictions."]

Thus, Poisson did not just analyze the data as raw numbers but as outputs of a mathematical model of the jury process. Although his assumptions regarding the independence and constancy of jurors' decisions may have been in error,[6] it put the analysis of jury behavior on a firm analytic footing, with his assumptions made explicit. Thus, his models may have been one of the first applications of OR (at least as defined herein) to problems of crime and justice.[7] Accordingly, we may have the French Ministry of Justice and its *Compte général de l'administration de la justice criminelle en France* to thank for the first application of operations research in this field.[8]

[6]For example, his model assumes that each juror's reliability is the same regardless of the truth, that is, that P (guilty vote|actual guilt) = P(not guilty vote|actual innocence) = u. This would not ordinarily be the case, especially if the benefit of the doubt is given to one side, as in the presumption of innocence.

[7]Stigler (p. 186) refers to earlier models of jury behavior, by Condorcet and Laplace, but they are not described.

[8]Criminal justice statistics were also instrumental in Poisson's development of the law of large numbers. According to Hacking (1990, p. 101), "Some jurors are more reliable than others. *Poisson's law of large numbers was devised in order to solve just this problem.* He studied as a model the situation in which reliability varies from juror to juror, but in which there is some law (Poisson's word) or probability distribution (our phase) for reliability among French jurors" (emphasis in the original).

There are important lessons to be learned here, one that "number crunchers" often forget. First, one cannot just assume that the variation in the data as depicted or tabulated (Fig. 1 and Table I) represents *all* the variation, because for any stochastic process, each datum may have an inherent variation; this is returned to later in the paper. And second, the fact that the model is simple (and even incorrect) does not negate its value.

2.2. The Early Twentieth Century

There was little OR related to problems of crime and justice between that time and the 1960s. In the United States, it was only when crime began to increase rapidly that criminologists and government officials began to take notice of the relationship between crime and the criminal justice system. This happened twice during this period of time, soon after Prohibition and again in the late 1960s. As a consequence of these increases, the United States Government did what it seems to do best in the face of pressing problems: it formed a commission to study them. The first commission, the National Commission on Law Observance and Enforcement (also known as the Wickersham Commission), published its findings in the early 1930s. One outcome of its work was the beginning of the systematic collection of police statistics (i.e., crimes reported to the police, arrests) by the International Association of Chiefs of Police and, subsequently, by the FBI (Maltz, 1977). At this time prosecutorial, court, and correctional statistics were still not collected with any degree of uniformity.

Although France and other European countries had been collecting criminal justice data for years, similar statistics for the United States were not so easy to obtain (Pollock and Maltz, 1994): whereas France had (and has) a centralized criminal justice system, in the United States virtually every state, county, and municipality has its own criminal justice agency, making standardization of crime statistics a major obstacle to their collection, one that has not been overcome to this day.[9]

2.3. The 1960s and the President's Crime Commission

These problems of data collection and coordination were apparent from the start of probably the most significant investigations into problems of crime and justice in this century, the President's Commission on Law

[9]This problem is still with us: a 1990 audit of the Illinois Computerized Criminal History system "found that nearly 85 percent of the arrests sampled were missing one or both of the state's attorney or court dispositions, while nearly 60 percent were missing both" (Illinois, Criminal Justice Information Authority, 1991, p. 3), yet the Illinois system is one of the better state systems.

Enforcement and Administration of Justice (1967). The President's Crime Commission, as it was known, spawned a number of task forces, among them a Task Force on Science and Technology, headed by Alfred Blumstein. The publication of the Science and Technology Task Force Report (1967) marks the beginning of the modern era of applying operations research to problems of crime and justice, an era that continues to the present day. The topics covered by that report set the priorities and direction for most subsequent operations research studies of the criminal justice system. They include

- system-wide models and simulations,
- population projections, and
- probability of arrest and response time.

In addition, a number of other topics have been the subject of operations research investigations, based on research subsequent to that report. They include models of criminal careers, incapacitation, recidivism, deterrence, crime measurement, business-oriented crime analysis, and jury scheduling. The rest of this paper examines the application of operations research to many of these topics.

3. SYSTEMWIDE MODELS

One of the first facts of life to confront the researchers of the Commission's Science and Technology Task Force (STTF) in their study of the criminal justice system was that there was no system. The separation of powers in governments at all levels in the United States, designed to ensure the independence of the executive, legislative, and judicial branches of government, also meant that the police (usually a municipal agency) paid little attention to the courts (a county agency), which paid little attention to the correctional system (a state agency), which paid little attention to the police. A primary task of the STTF, then, was to describe the system, or as it was often called, the "nonsystem," in sufficient detail to permit decision-makers to see how problems in one part of the system affected the rest of the system. No one knew, for example, how a change in police resources (e.g., an increase in the number of police officers, shifting more patrol officers to the detective division) would affect

- the workload of prosecutors (perhaps leading them to use plea bargaining more often),
- the number of judges and courtrooms needed,
- the number of new correctional institutions needed at different security levels, and
- the number of probation or parole officers needed, etc.

Nor could anyone provide useful forecasts of the crime rate based on population statistics. The STTF therefore called for a *systems analysis* of the criminal justice system, to determine how changes in criminality affected agency activity, or how changes in one criminal justice agency affected other agencies.

One of the first such systems analyses, designed to show how the parts were interrelated, was conducted by the New York State Information and Intelligence System (NYSIIS, 1967) in its report entitled NYSIIS: System Development Plan. The first figure in that report was a 6-ft-long foldout figure entitled, "The criminal justice process for adult felonies." It detailed how a criminal case proceeds through the criminal justice system from the initial crime to the incarceration of an offender and this eventual discharge or release.[10] Its primary goal was to develop the specifications for a statewide computerized information system that would collect, store, and disseminate criminal justice information throughout the state. Its developers believed that, based on this information system, decisions would be made on a more rational basis; people and problems would no longer "fall through the cracks;" and the total system would function more effectively and efficiently. (Fig. 3).

These beliefs in the ultimate triumph of technology in curing urban ills should be seen from the historical context of the era: we were about to put a man on the moon, and the cry from everyone was, "If we can put a man on the moon, why can't we solve *X*?" where *X* was whatever urban problem was pressing at the moment. In addition, the war in Vietnam was still being fought, and there were a number of systems analysts who had developed command-and-control systems for the military who felt that the same techniques could be brought to bear on solving the crime problem. Furthermore, "crime in the streets" had been a major focus on the Presidential campaign of 1964, and computers and other high-tech solutions were just beginning to be seen (often inappropriately, as it turned out) as the answer to social as well as technical problems.

Other systems analyses of criminal justice began to turn up in the academic and professional literature (e.g., Blumstein and Larson, 1967, 1969). These analyses underscored the fact that offenders who wend their way through the criminal justice system often wend their way again and again through the same system. This observation that recidivism provides the positive feedback to the system (e.g., Blumstein and Larson, 1971) gave rise to studies to see how recidivism affected the criminal justice system (Belkin *et al.*, 1973). As seen in Fig. 4, a very simplified system flow diagram,

[10]Figure 3, taken from the STTF (1967), is a similar depiction of the flow of criminal cases through the criminal justice system.

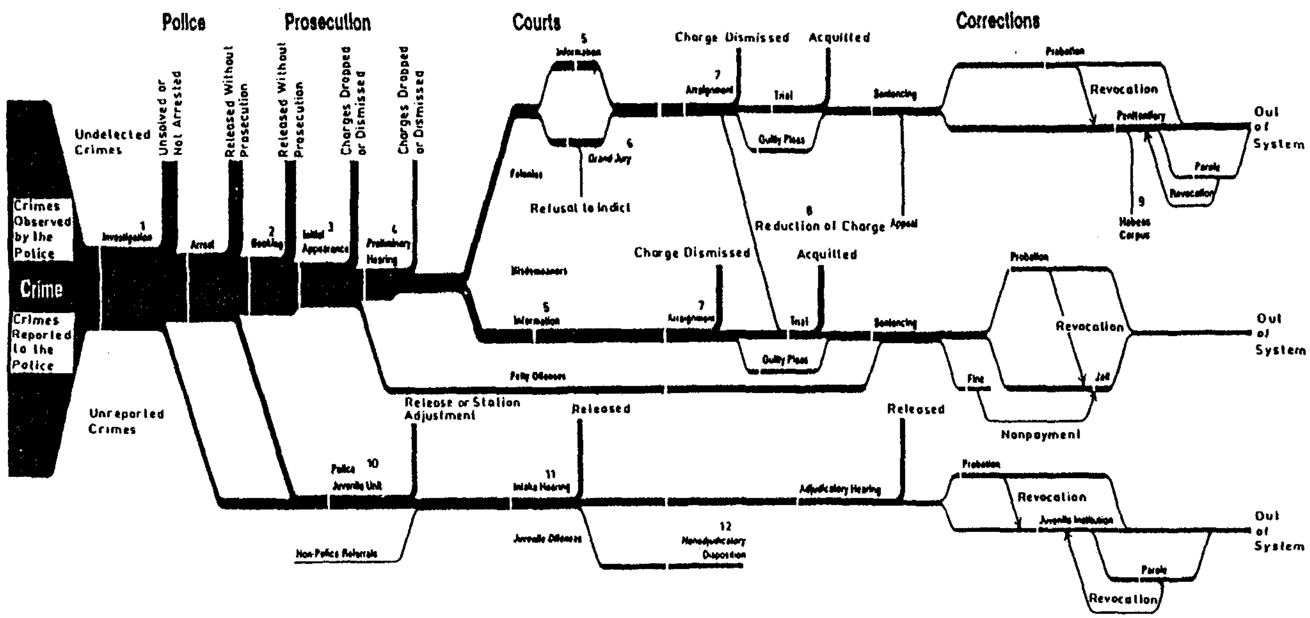


Fig. 3. A general view of the criminal justice system. This chart seeks to present a simple yet comprehensive view of the movement of cases through the criminal justice system. Procedures in individual jurisdictions may vary from the pattern shown here. The *differing* thicknesses of the lines indicate the relative volumes of cases disposed of at various points in the system, but this is only suggestive, since no nationwide data of this sort exist. [Source: Science and Technology Task Force (1967, pp. 58–59).]

"virgin" (i.e., first-time) arrestees, $V(t)$, and recidivating arrestees, $R(t)$, enter the criminal justice system. A fraction 1-Ω of them never recidivate, and the remaining Ω of them reenter the system after remaining free for a (mean) time delay τ.

Aside from explicitly recognizing the role of recidivism as a feedback loop to the system, this model also led to development of JUSSIM, a computerized simulation of the criminal justice system (Belkin *et al.*, 1972). This system and its variants are still in use (e.g., Cassidy, 1985; McEwen and Guynes, 1990; Institute for Law and Justice, 1991). Similar work has been carried out in the United Kingdom (Rice, 1984; Morgan, 1985; Pullinger, 1985).

Aggregate flow models of this sort (which can become quite complicated, given the complicated structure of the criminal justice system) are very useful in developing workload statistics that answer "what if" questions relating to changes in either the input (e.g., the "baby boom") or the system (e.g., elimination of plea bargaining). They are of limited value, however, in determining what the workload actually consists of. For example, study after study (e.g., Wolfgang *et al.*, 1972; Chaiken and Chaiken, 1982; Reiss and Farrington, 1991) show that a surprisingly small fraction of all juveniles accounts for a very high proportion of all crimes and arrests. Models that relate the flow to offenders' behavior are required, to determine the effect of each behaviorally distinct subpopulation on the criminal justice system.

Attempting to identify the characteristics of this small subpopulation became the focus of much research. It was felt that their identification (prospectively, before they committed the bulk of their criminal acts) would mean that the more dangerous individuals could be identified for early intervention programs or, should intervention prove ineffective, could be incarcerated for longer periods of time than those offenders who are less dangerous.[11]

A first step in this direction was embodied in a very simple model proposed by Shinnar and Shinnar (1975; see also Avi-Itzhak and Shinnar, 1973). Assume that a person commits crimes at a Poisson rate λ of per year. If not arrested he/she would commit λT crimes during his/her career length T years. If the probability of arrest for each crime is q, the probability of imprisonment (given arrest) is Q, and the average sentence length (given imprisonment) is S, the offender would on average be free for $1/[\lambda q Q]$

[11]Section 5, Incapacitation, discusses this line of research. There are some major ethical issues imbedded in it. First, it implies that it is proper to sentence an individual not only for the crime(s) already committed, but also for the crimes to come, which is antithetical to the philosophy of our criminal justice system; see von Hirsch (1985). Second, because no method of prediction is perfect, a sizable fraction of the predicted population would be punished for crimes they would not have committed, and a sizable fraction of the dangerous offenders would not be so labeled and would avoid the longer incarceration.

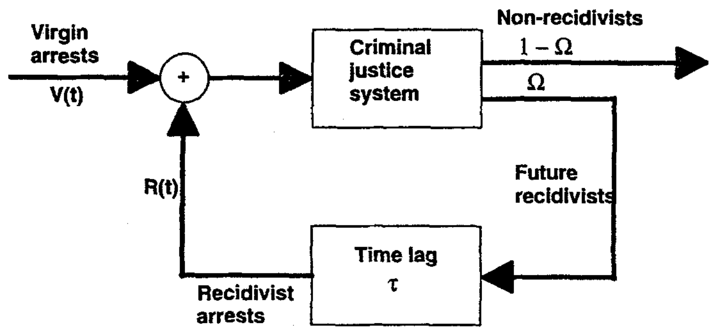


Fig. 4. A simplified criminal justice feedback model.

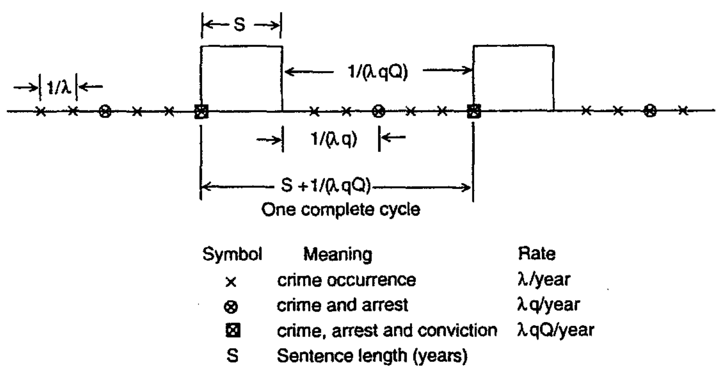


Fig. 5. A (deterministic) criminal career.

years between the prison sentences of S years (see Fig. 5 for a deterministic representation of such a career). Thus, the offender's *effective* crime rate is $\lambda/[1+\lambda qQS]$, which provides a quantitative estimate of the incapacitative effect of prison. Although there are some gross simplifications in this model, it encapsulates in one equation the effects of the primary elements of the criminal justice system: the offender is represented by λ, his/her offense rate; policy by q, the probability of arrest; the courts by Q, the probability of commitment to prison; and the correctional system by S, the sentence length.[12] Cohen (1978) reviewed this and other models of incapacitation,

[12]In this simple model the variable T can be argued away by assuming that the birth and death processes balance out. In other words, as one offender's career terminates, another offender arises to take his/her place. Other, more realistic, models are discussed below.

and provided some guidance for future research on modeling the incapacitative effect of imprisonment.[13]

The simple model of offender behavior depicted in Fig. 5 produced a number of important and controversial lines of research. The first relates to the description of criminal careers; the second relates to the possibility of incapacitating high-rate offenders to reduce the crime rate. The next section describes some of the research in criminal careers.

4. CRIMINAL CAREERS

The criminal career paradigm was detailed in a series of reports and articles (Blumstein *et al.*, 1986, 1988a, b; Blumstein and Cohen, 1987). They showed the utility of modeling an offender's career using four dimensions:

- participation, the probability that an individual will commit a crime;
- frequency, the rate at which he/she commits crimes while criminally active;
- seriousness, the nature of the crimes he/she commits; and
- length, the time interval between his/her first and last offenses.

Although it was certainly an oversimplification to assume that all offenders would fit the same model of behavior, one could apply this model to different subpopulations and aggregate them to come up with the overall characteristics of the population under study.

Although the two are often confused, there is a difference between the concepts of the "criminal career" and the "career criminal." The *career criminal* is an individual for whom crime is a profession. A *criminal career* is merely a pattern of offending, delineating the progression of incidents from the one that initiates the individual's career, through the last, at career termination. Having a criminal career, a property of everyone who commits at least one crime, does not imply that the individual is a career criminal.[14]

One of the key variables relating to an offender is his/her rate of commission of offenses λ. This must ordinarily be inferred from his/her arrest rate $\mu = \lambda q$, where q is the probability of being arrested. The arrest rate μ is based on the "rap sheet," a listing of an individual's arrest, but not crime commission, history; and inferring the offense rate from this information is not a straightforward process—see Blumstein and Cohen (1979). Some argue

[13]Greenberg (1975) proposed a model of offender behavior that is conceptually similar but different in its details. Using aggregate statistics, his model permits the estimation of various criminal career parameters. See Cohen (1978) for a review and comparison of these and other models.

[14]All of my students have had scholarly careers; only a few of them (including the editor of this journal) are career scholars.

that arrests represent the crimes which the offenders failed to commit successfully and, therefore, are a biased sample of the offender's crime history. Self-report studies, however (e.g., Chaiken and Chaiken, 1982), seem to indicate that the crimes for which an offender is arrested are generally indicative of the types of crimes he/she generally commits.

In the first studies it was assumed that λ was constant for all offenders during their period of active offending.[15] Later models included variations in offense rates (Maltz and Pollock, 1980b; Rolph *et al.*, 1981; Lehoczky, 1986; Copas and Tarling, 1988).

The distribution of λ, offenders' crime commission rates, is an important aspect of a criminal career. Although offenders are not all expected to have the mean crime commission rate, the particular distribution of rates does not fit a well-behaved parametric distribution either. Using data collected from prisoners in three states (Chaiken and Chaiken, 1982), Rolph *et al.* (1981) developed a method to estimate the distribution of offense rates among offenders. They analyzed data provided by imprisoned offenders concerning the number of offenses they reported having committed while free, and then used Turnbull's (1976) method to estimate the empirical distribution of offense commission rates. In attempting to fit the data with various distributions they found that the thickness of the tail of the distribution made this process difficult (see Fig. 6).

Since the data were obtained from imprisoned individuals, certain biases were present. A sample of prisoners is not representative of the general offender population; for example, all other factors being equal, it would contain 10 times as many individuals imprisoned for 10 years as it would individuals imprisoned for 1 year. For that reason, they developed a model-based method for extrapolating to other populations of interest. Using a number of simplifying assumptions, they estimated that the probability that an offender is in prison is $\lambda qQS/(1+\lambda qQS)$ (see Fig. 5), then used the inverse of this probability as sampling weights. Although the assumptions were admittedly somewhat simplistic, the poor quality of the data militated against developing a more complicated model.

These methods were used to analyze data from a major study (Chaiken and Chaiken, 1982) that described the past criminal experience of convicted offenders in three states, in which subpopulations of offenders were identified based on their offending histories. In addition, prior longitudinal studies of

[15] This was assumed for mathematical tractability and convenience, not because analysts showed that this was actually the case. This concept, of a constant (Poisson) offense rate, has been misconstrued by some: they take this to mean that the model implies that offenders are very regular in their criminal activity, unaware, perhaps, that a Poisson process can be realized in an infinite variety of ways, most of which would not appear to be very regular.

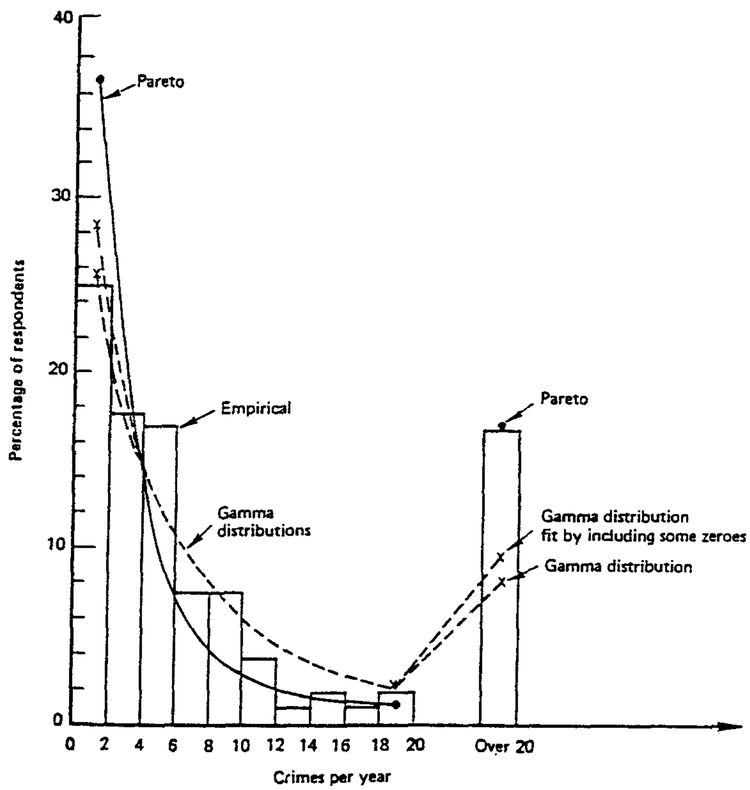


Fig. 6. Distribution of crime rates for business robberies. [Source: Rolph *et al.* (1981, p. 28).]

prospective criminal activity were investigated to determine the characteristics of subpopulations and their career parameters (Blumstein and Cohen, 1987; Blumstein *et al.*, 1985, 1988; Barnett *et al.*, 1987, 1989), including the use of hierarchical models (Ahn *et al.*, 1990).

A similar (hierarchical model) approach was taken by Blumstein *et al.* (1993) to estimate the expected variation in λ of the population studied by the Chaikens. Because high-rate offenders are more likely than low-rate offenders to be incarcerated due to their greater exposure to risk—a process they call "stochastic selectivity"—modeling is needed to estimate the characteristics of the offending population from which the inmates were drawn. They posited that the incarcerated offenders committed offenses at an offense rate λ_i that was drawn from a probability distribution that is a mixture of r exponential distributions,

$$f(\lambda) = \sum_{k=1}^{r} p_k \beta_k e^{-\beta_k \lambda}$$

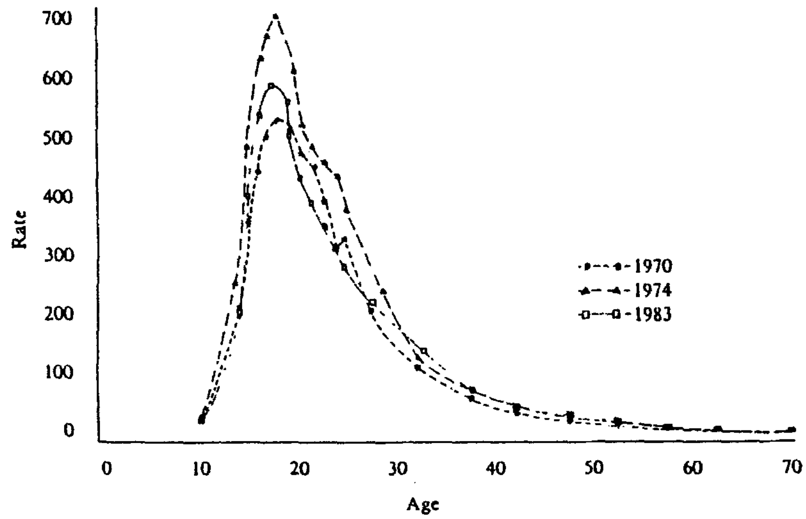


Fig. 7. Number of robbery arrests per 100,000 males. [Source: Gottfredson and Hirschi (1986, p. 223).]

where the mixing parameters p_k sum to 1. Using maximum-likelihood estimation, they show that the robbery offenses for the three states' offenders can be fit by a mixture of three exponentials: about 90% of the offenders have an expected robbery commission rate of 1 per year, about 10% of the offenders have an expected rate of about 9 robberies per year, and less than 1% have an expected rate of over 100 robberies per year. This distribution of robbery offenders "in the wild" contrasts with the estimate it provides of incarcerated robbers: about half come from the low-rate subpopulation, one-third from the middle-rate subpopulation, and one-sixth from the high-rate subpopulation.

4.1. A Competing Theory: The Age–Crime Relationship

The criminal career paradigm was challenged by some researchers (e.g., Gottfredson and Hirschi, 1986, 1987, 1988, 1990; Tittle, 1988; for counter-arguments see Blumstein *et al.*, 1988a, b), who felt that it was unnecessarily complicated. They noted the long-term stability of the relationship between arrest rate and age (Fig. 7). The stability of this relationship, they argued, is based on offenders' lack of self-control, which leads directly to criminality, and this lack of self-control increases through adolescence and then diminishes with age. They maintained that "differences in the propensity to engage in criminal acts are established before the high-crime-rate years, persist during those years, and indeed maintain themselves throughout life" (Gottfredson and Hirschi, 1989). They considered this formulation of a stable criminality equivalent to a natural law, making it superfluous to consider models of offender behavior that incorporate four variables (i.e., participation, frequency, seriousness, and length) into a model. They argued that their cross-sectional approach to analysis of criminality held more promise than the longitudinal approach, especially considering the cost of longitudinal research.

4.2. Arguments Against the Age–Crime Explanation

This reliance on a stable pattern of *aggregate behavior* to make inferences about *individual behavior* is reminiscent of the approach used by Quetelet and Durkheim (Maltz, 1994a), who saw in the constancy of numbers (conviction rates and suicides, respectively) from year to year an innate "*penchant au crime*" or "suicidogenetic currents." The age–crime stability is, however, based on better empirical evidence than that used by Quetelet or Durkheim. In particular, not only are the patterns stable over time, but also the relationship of these patterns to covariates (e.g., race, education, income) is stable over time. Thus, there is much to be learned from the cross-sectional approach, yet there are also some significant problems with this approach; inferring from a stable *aggregate* pattern a characteristic of *every member of a population* sweeps away the variation within that population and relegates it to "error."

This theory of criminality, based on an age–crime constancy, has been challenged on other grounds as well. Harada (1991) points out that the concept of a stable criminality implies that the hazard rate[16] of every individual in a group follows the same general course, the only difference between individuals being a constant of proportionality. He used arrest records of a sample of Japanese youth of junior high-school age to determine whether hazard rates actually followed this pattern. His findings do not support the stable criminality hypothesis. Furthermore, there is additional evidence (from Visher and Linster, 1990) (see Fig. 8) that individuals with different characteristics have very different hazard rates; see also Moffitt (1993). In other words, the fact that a *group* exhibits a certain stability in its statistics does not guarantee that each *individual* in the group has the same general characteristic, which is implied by Gottfredson and Hirschi's theory. For example, after studying longitudinal data on criminal careers, Blumstein *et al.* (1982) found that the drop in arrests with age reflected not less activity per year among active offenders but, rather, a growing fraction of offenders who terminated their criminal careers.

[16]The hazard rate at time t is the probability that an individual will fail (i.e., be arrested) in the time interval $(t, t+dt)$, conditioned on his/her not having failed up to time t.

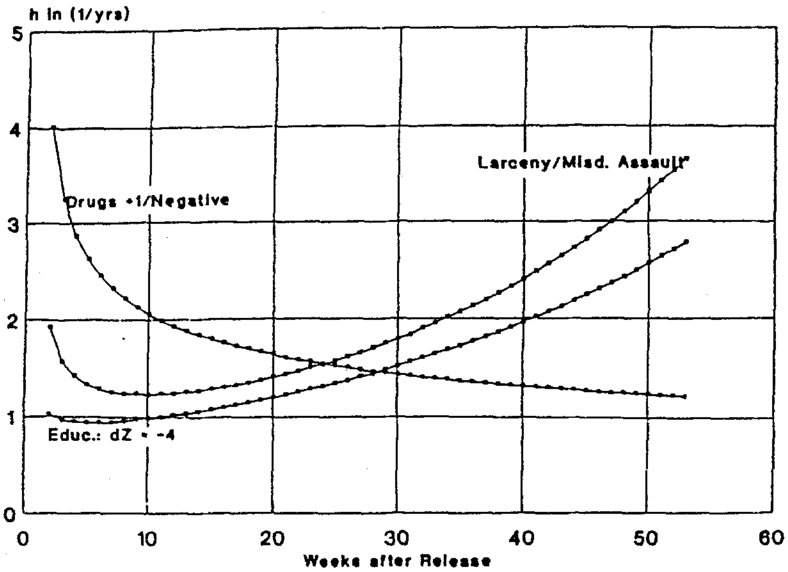


Fig. 8. Hazard rate for pretrial arrest, for individuals with different characteristics. [Source: Visher and Linster (1990, p. 168).]

The debate over the adequacy of different models of criminality does not seem to be diminishing. In fact, when a journal editor asked another researcher to compare the two approaches, that researcher posited a new model of offender behavior (Greenberg, 1991). And when yet another journal editor found the controversy continuing between the developer of this new model and proponents of the criminal career paradigm (Barnett *et al.*, 1992; Greenberg, 1992), another reviewer was asked to resolve this newer controversy—and this reviewer proceeded to posit yet another model (Land, 1992)! It brings to mind Persig's Postulate, one of Murphy's laws (Bloch, 1977, p. 87)—"The number of rational hypotheses that can explain any given phenomenon is infinite."

In point of fact, the new models are not as well grounded as the ones developed using the criminal career paradigm, which are based on a much more thorough analysis of the data. It should also be noted that some of the insights possible using the longitudinal approach would not have been gained in a cross-sectional study of criminal careers. One of the more significant results is that, although *participation* in criminal activity is strongly correlated with race or ethnicity, the subsequent criminal careers of offenders of different racial and ethnic backgrounds are very similar: for example, although one study (Nagin and Smith, 1990) showed that participation and frequency appear to be closely associated, two other studies (Farrington and

Hawkins, 1991; Shover and Thompson, 1992) indicate that participation and desistance are predicted by different characteristics. And a cross-sectional approach makes it difficult to disentangle the effects of age, period, and cohort on a group's criminal behavior (Mason and Fienberg, 1985).

Furthermore, some characteristics of an offender's career cannot be observed in a cross-sectional study. For example, Reiss and Farrington (1991) use longitudinal data to show that "the incidence of co-offending[17] decreases with age primarily because individual offenders change and become less likely to offend with others rather than because of selective attrition of co-offenders or persistence of those who offend primarily alone" (p. 393). A cross-sectional study would not have been able to discern this pattern of behavior, since individual offenders are not tracked over time.

4.3. Limits of These Models

Awash as we are in crime statistics, we sometimes forget that the numbers we use have severe limitations. Models of criminality, whether they are based on the criminal career paradigm or the age–crime paradigm, are often based on administrative statistics, e.g., number of arrests by age, race, sex, type of offense. Administrative data, however useful they are in estimating the flow of offenders through the system, cannot be used to infer behavioral relationships; they are of benefit in explaining *what* happens, but not *why* it happens.

5. INCAPACITATION

The correctional system has many goals, among them deterrence, education, retribution, rehabilitation, and incapacitation. It is sometimes difficult to determine the extent to which these goals are met. OR and mathematical modeling have played a part in estimating the extent to which they may be achieved. In this section models relating to incapacitation are explored.

Models of incapacitation are all based to some extent on those developed by Avi-Itzhak and Shinnar (1973) and Shinnar and Shinnar (1975). Greenberg (1975) used aggregate data to develop estimates of the extent to which imprisonment affects the general population's crime rate through incapacitation. He also distinguished between *collective* incapacitation—the effect on the overall crime rate[18] of prison sentences in general—and *selective*

[17]Co-offenders are persons who act together in a crime, as juvenile gang members are wont to do.

[18]In the following discussion I distinguish between the overall crime rate, such as the number of crimes per year in a jurisdiction, and the individual crime rate λ, the estimated number of crimes per year committed by a given offender.

incapacitation—the effect on the overall crime rate of giving longer sentences to selected individuals whose characteristics make them more likely to have a higher *individual* crime rate (i.e., λ) upon release.

5.1. Collective Incapacitation

If incapacitation can reduce the overall crime rate, there should be a means of determining the extent to which different sentencing policies affect it: for example, is it better to sentence half as many offenders to twice the sentence, or twice as many to half the sentence? Such determinations depend on the characteristics of the offending population, and on the effect that prison is likely to have on their subsequent careers. Blumstein and Nagin (1978) developed a model that related the assumed characteristics of offenders to the sentencing process, to determine how sentencing policies might affect overall crime rates. It estimated the crime-control potential of different policies and the trade-off that can exist between mean sentence length and overall crime rate. This model appears to be in agreement with Langan's (1991) analysis that showed that the prison population, which had soared from the mid-1970s due to changes in sentencing practices, was accompanied by a reduction in crime, at least through the 1980s, that might have been due to the increased incapacitation of offenders.

5.2. Selective Incapacitation

A proposal for serious consideration of selective incapacitation as a penal policy was put forward by Greenwood (1982). It was based on combining the findings from two studies. The first study, an analysis of interviews of incarcerated offenders in California, Texas, and Michigan (Chaiken and Chaiken, 1982), calculated the individual offense rates (while free) for these individuals; they demonstrated that rates for some offenders (called "violent predators") were extremely high (see Fig. 6). The second study (Hoffman and Beck, 1974; Hoffman and Adelberg, 1980) developed a "Salient Factor Score," a relatively simple method of scoring risk levels posed by Federal prisoners if given parole. It uses data concerning an inmates' conviction and incarceration record, including number of prior convictions and incarcerations; age at first commitment; whether the current conviction was for auto theft; parole, drug, educational, and employment history; and living plans at release.

Greenwood (1982) analyzed data from the California high-rate offenders and found that, by using only a few elements from the Salient Factor Score, it was possible to identify (retrospectively) a significant number of the high-rate offenders. He concluded that it might be possible to have an effect on the crime rate by identifying the high-rate offenders (prospectively) using the Salient Factor Score and then selectively incapacitating them by having them serve longer sentences than others not so identified. He estimated that a 20% reduction in California robberies would ensue if sentences for the scale-identified high-rate offenders were to approximately double while sentences for other offenders were held to 1 year.

5.2.1. Utility of Selective Incapacitation

Regardless of the nature of the estimator, data on past offenses cannot be assumed to be a good predictor of future offenses. Significant errors result; even with retrospective data, 55% of those identified as high-rate offenders were not in fact high-rate offenders, and 46% of true high-rate offenders were not so classified by the scale. This proposed policy had other difficulties as well (Blumstein, 1983; Cohen, 1983). And Visher (1986) showed that applying the sentencing policy, based on Greenwood's analysis of California inmates, to Michigan's inmates would *increase* robberies by 30%.

In an effort to explore whether the prediction technique used by Greenwood would select high-rate offenders, Barnett and Lofaso (1985) tested it with data from a cohort of Philadelphia youths (Wolfgang *et al.*, 1972). They showed that the arrest rates among active delinquents do not vary that much—the most active have about 2.4 times the rate of the least active. Their study could not estimate the extent of Type 1 and Type 2 errors, because their data set was truncated when the youths reached 18 and there was no way of knowing the extent to which offense activity continued (or desisted) beyond that age.

Subsequently, Greenwood and Turner (1987) studied the offenders who had been labeled high-rate offenders in the first (Greenwood) study. They found that it was not possible to use the Salient Factor Score to provide a sufficiently accurate prediction as to who would continue to be a high-rate offender: offense rates may have highly dispersed distributions (such as in Fig. 9, based on offender self-reports); arrest data do not exhibit the same dispersion.

However, not all of the evidence is negative. Block and van der Werff (1991), in their study of a cohort of individuals arrested in The Netherlands in 1977, found that they were able to identify, prospectively, the active and dangerous offenders who are responsible for a disproportionate number of serious offenses. They used only official data, but their data included juvenile records, which may limit the utility of such a method in the United States: currently in many jurisdictions, juvenile records may not be used in adult proceedings. But it does suggest that the goal of identifying those who may become dangerous offenders may be achievable to some extent.

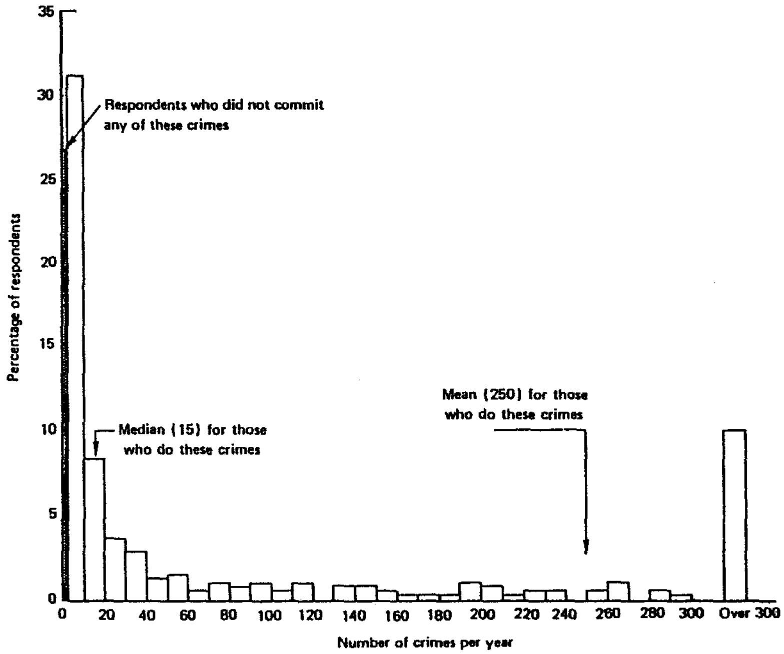


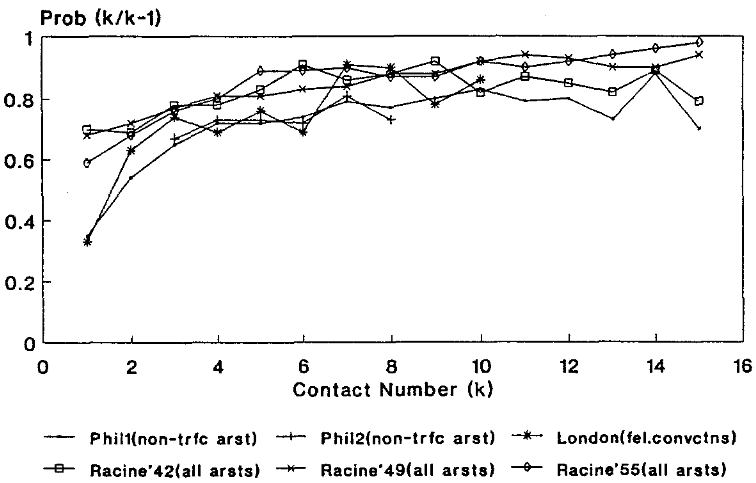
Fig. 9. Observed crime commission rate for burglary, assault, theft, auto theft, forgery, and fraud. [Source: Rolph, Chaiken, and Houchens (1981, p. 13).]

5.2.2. *Offender Desistance and Selective Incapacitation*

Identifying those *likely to become* high-rate offenders on the basis of juvenile and adult records is not equivalent to actually identifying these offenders. As discussed earlier, the variables associated with termination of a criminal career are not the same as those associated with its initiation, frequency, or seriousness. One means of investigating career termination is to obtain sufficiently detailed longitudinal data about a group of offenders so that their residual career length can be calculated. One can then analyze the data to determine the variables that appear to be associated with residual career length, to see the extent to which policy changes might affect it.

Toward this end, a reanalysis of a number of data sets was undertaken (Blumstein and Moitra, 1980; Blumstein *et al.*, 1985, 1986). They contained information relating to the characteristics of the individuals and records of their contact with the criminal justice system.[19] Based on this information

[19]The definition of what constituted an event varied from study to study. In one, it was all nontraffic police contacts; in another, it was all police contacts; in a third, it was conviction.



Data from Blumstein et al (1986: 90)

Fig. 10. Conditional probability of persistence in offending, for a variety of data sets. Probability of a kth contact given $k-1$.

they calculated the probability of having $k+1$ contacts, given k contacts (Fig. 10).

This study did not disaggregate the data by offender characteristics, but it shows that after a small number of police contacts, regardless of the definition of contact, the probability of additional contacts $[P(k+1 \mid k)]$ is relatively constant and quite high—about 0.8.[20]

It is noteworthy that, using so many different definitions of contact, populations, and sources of data, the probability trajectories are so similar. However, it is difficult to infer very much from these data. First, the observation times are different for each cohort, with the cutoff age ranging from 22 to 32 years of age. Additional offenses that occur after the cutoff age are not included in the data. This means that, if a person is arrested for a fifth crime at age 23, in one study he/she would be considered a person who desisted after his/her fourth contact, while in another he/she would be considered as having five contacts.

Second, the seriousness of events varies considerably from study to study: the Racine cohorts include all offenses, traffic as well as nontraffic, under the rubric "police contact," whereas the London cohort includes only convictions for indictable offenses. Moreover, seriousness can confound

[20]Broadhurst and Maller (1991), however, did not find the same constancy in their analysis of reconviction data from a large database of Western Australian prisoners.

interpretation *within* a study. For example, a Racine youth incurring 15 police contacts through the cutoff age has probably not committed offenses as serious as a Racine youth with two felonies who is still incarcerated at the cutoff age.

Research on the correlates of desistance is increasing (Rhodes, 1989; Farrington and Hawkins, 1991; Shover and Thompson, 1992). Results to this point, however, do not indicate whether an exponential model of desistance (implied by the relative constancy of the probabilities in Fig. 10) is warranted.[21]

5.3. Selective Incapacitation and Allocating Prison Space

Another use to which (a variant of) the Avi-Itzhak and Shinnar model has been put is in attempting to determine the extent to which early termination of criminal careers while they are in prison will "waste" the space that might be occupied by a person who will still be criminally active, and therefore incapacitated, while in prison. If we assume that (1) offenders commit crimes during their criminal careers at rate λ, (2) their probability of arrest and subsequent conviction for an offense is qQ, (3) their career length is exponentially distributed with mean T, and (4) when incarcerated they receive a fixed sentence S, it can be shown (Avi-Itzhak and Shinnar, 1973; Barnett and Lofaso, 1986) that the number of crimes committed by an offender, N, is

$$N = \frac{\lambda}{(1/T) + \lambda q Q(1 - e^{-S/T})}$$

Even with this idealized model of offender behavior, Barnett and Lofaso show that it is not possible to develop clear-cut rules of thumb regarding allocation of prison space. Using a number of examples, they estimate the extent to which crimes are averted and conclude that using selective incapacitation for "identifying the high-rate offenders and locking them up for enormous periods" is very wasteful of resources.

6. RECIDIVISM

The recidivism of an offender is the *sine qua non* of criminal career studies, since it is by virtue of his/her reappearance, at fairly regular intervals, in the arrest, conviction, and correctional processes that criminal

[21]Note that the abscissa in Fig. 10 is number of police contacts, not time. Therefore, it is certainly not possible to infer an exponential model in time for these data. Note also that the offense histories represented in this figure are time-truncated at different times, depending on whether an individual was incarcerated when data collection stopped. Thus a person given a long sentence after two contacts (presumably because he/she is more dangerous) might be represented in this figure as someone who "desisted" after $k=2$ contacts.

careers can be studied. Recidivism has been analyzed in two ways in the literature. It is most often looked on as a failure that occurs at a specific point in time, interrupting the offender's time on the street. (This is especially true when reincarceration is the definition of recidivism.) It has also been investigated as a continuing process of being arrested, since arrests do not always interrupt a person's street time.

6.1. Recidivism as a Failure Event

Recidivism was originally measured for the purpose of evaluating correctional treatments. Two groups were given different treatments, and their postrelease activity was monitored. The benchmark of comparison was the fraction of offenders who failed (e.g., were rearrested, had their parole revoked) during the first 12 months after release. A simple difference-of-proportions test was used to determine whether the difference was statistically significant.

However, individuals fail at different times, and comparing failure proportions at only one point in time is not warranted. As can be seen from Fig. 11, although the "satisfactory" group failed at lower rates than the "unsatisfactory" group at all three measured points of time, a reasonable extrapolation beyond the given data would suggest that the unsatisfactory group would ultimately have a lower fraction of failures. The risk of failure varies over time, and a benchmark at a single point in time can be misleading (Berecochea *et al.*, 1972). Stollmack and Harris (1974; see also Harris and Stollmack, 1976) studied recidivism as a failure process over time and showed how to apply failure-rate techniques to study the effectiveness of correctional programs. In particular, they compared two groups of releasees from DC prisons and showed that the (exponential) failure rates of the two groups were significantly different from each other. Models based on the reliability of electromechanical components (e.g., Barlow and Proschan, 1975; Kalbfleisch and Prentice, 1980) were thus invoked to study recidivism.

Reliability models, however, have their limitations. Whereas we can expect the eventual failure of every light bulb or hard disk drive,[22] not every person can be expected to recidivate: some people do turn over new leaves and desist, and do not reappear in the criminal justice system. A new class of recidivism models was proposed (Maltz and McCleary, 1977; see also Partanen, 1969; Carr-Hill and Carr-Hill, 1972; Greenberg, 1978) that specifically includes the possibility of nonfailures.

In particular, the failure process can be modeled by an incomplete distribution (one whose cumulative probability of failure does not reach 1

[22]Even during the initial preparation of this review, unfortunately.

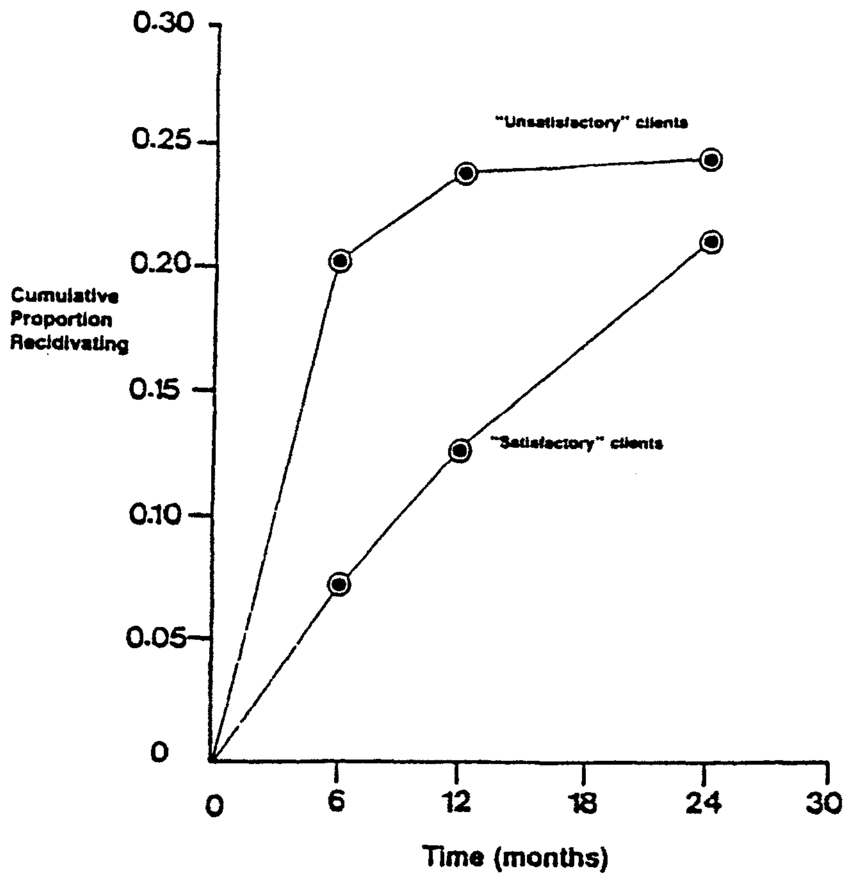


Fig. 11. Recidivism rates at three times, for two groups of Minnesota halfway-house releasees. [Source: Maltz (1984, p. 73).]

as time approaches infinity). The cumulative distribution function for the incomplete exponential distribution (one that is suggested by Figs. 11 and 12) is

$$F(t) = \Omega(1 - e^{-\phi t}), \qquad \text{where} \qquad 0 < \Omega \leq 1$$

This model of offender behavior is closely related to the "innocents, desisters, and persisters" model of Blumstein *et al.* (1985). The desisters consist of the $1 - \Omega$ nonrecidivists, and the remaining Ω are persisters; the innocents, never having committed a crime in the first place, never have the opportunity to recidivate.

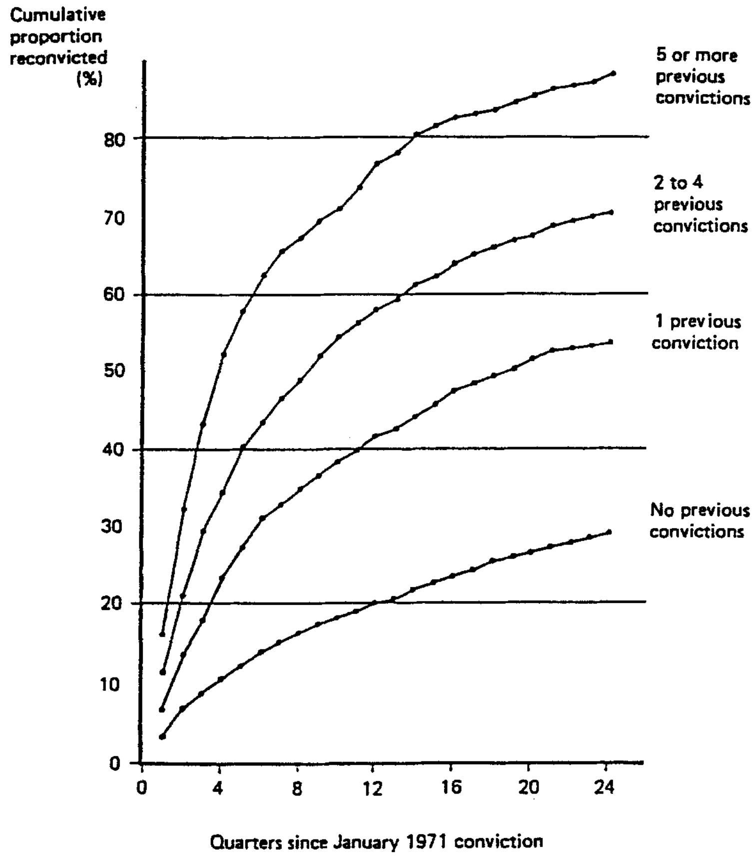


Fig. 12. Proportion reconvicted as a function of prior record. [Source: Philpotts and Lancucki (1979, p. 23).]

Maximum-likelihood methods for estimating Ω and ϕ were developed[23] (Maltz and McCleary, 1977, 1978; Maltz *et al.*, 1979), and their empirical validity has been confirmed using data from a number of sources (Maltz, 1984; Blumstein *et al.*, 1985; Barnett *et al.*, 1987, 1989; Schmidt and Witte, 1988; Broadhurst *et al.*, 1988; Rhodes, 1989; Visher and Linster, 1990). They imply that the incomplete exponential distribution is an appropriate model when failure time is based on arrest but that the incomplete lognormal distribution may be the model of choice when the failure time is based on return to prison. A more complete discussion of the different distributions used to model recidivism is given by Maltz (1994b).

[23]Note that ϕ in the above equation is equivalent to $1/\tau$ in Fig. 1.

6.1.1. Population Heterogeneity

The incomplete distribution represents but one class of models for studying recidivism. It implies that a group can be split into two subgroups, one which does not recidivate and the other which recidivates according to a particular distribution (e.g., exponential, Weibull, lognormal). There are, however, other modes of failure. For example, both groups may fail, but at different rates:

$$F(t) = \Omega f_1(t) + (1 - \Omega) f_2(t)$$

where f_1 and f_2 are two probability density functions. Harris *et al.* (1981) investigated a mixture of exponentials (with three parameters, Ω, ϕ_1, and ϕ_2) and Broadhurst and Maller (1991) considered a mixture of two Weibull functions (with five parameters).

There are also other ways to model population heterogeneity. Lehoczky (1986) suggested the use of hierarchical models. Gerchak and Kubat (1986) considered the use of β and γ mixtures of exponentials. Schmidt and Witte (1988) performed a covariate analysis of recidivism data. [They found, however, that models using individual covariates are hardly better than models that do not use explanatory variables (Schmidt and Witte, 1988, p. 117).] And Ellerman *et al.* (1992) modeled population heterogeneity using a Burr model, a continuous mixture of Weibull distributions.

It is always difficult to decide how many parameters to use to model a process. Koopman (1986, p. 379) invoked the aphorism, "With six parameters I could model an elephant," to state his aversion for overly complicated models. Models with few parameters can provide a great deal of insight, at the expense of grouping many different phenomena together, while those with a greater number of parameters fit the data better, at the expense of providing less insight; and "the more complex the model, the further do its data requirements outstrip the numbers that are readily available" (Barnett, 1987, p. 31). The inclination to keep the number of parameters as small as possible, using the principle of Ockham's razor, seems to have some validity from a Bayesian standpoint (Jefferys and Berger, 1992).

Regardless of the model used, it seems clear that the models which best fit the data appear to be those with decreasing hazard rates. There is a theoretical explanation for this as well; since some offenders may have high hazard rates and others low rates, the resulting pooled data will always exhibit a decreasing hazard rate—those with high hazard rates fail early, so as time increases, the remaining members have a lower and lower mean hazard rate (Proschan, 1963).

6.1.2. The Proportional Hazards Method

Other methods have also been used for studying recidivism. Barton and Turnbull (1979, 1981) used Cox's proportional hazards method (Cox and

Oakes, 1984). This requires the hazard function $h_i(t; x)$ for individual i to be of the form

$$h_i(t; x) = h_0(t) f_i(x)$$

where

$$\ln[f_i(x)] = \alpha_0 + \alpha_1 x_{i1} + \alpha_2 x_{i2} + \cdots + \alpha_k x_{ik}$$

and x_{ij} is the value of covariate j for individual i. If h_i is of this form, then the specific form h_0 of the hazard function does not enter into the likelihood-maximization procedure, and the values of the α_j's that maximize the likelihood function can be readily determined.

Although this method is nonparametric, it incorporates two assumptions, one reasonable, one questionable. The first assumption is embodied in the functional form of $f_i(x)$, that the covariates are loglinearly related; this is done for analytical tractability, so that the likelihood function can readily be maximized over the α's.

The second assumption, however, is much more restrictive: specifically, all individuals are assumed to have the same basic hazard rate, a multiplicative constant being the only difference among them. For example, it assumes that hazard rates do not cross.[24] This assumption may in fact be suitable for some biostatistical applications, in which there is only one failure mechanism (a disease) that affects all individuals in the same way. However, there are indications that it is not true for correctional applications. Visher and Linster (1990) depict hazard distributions for some individuals in the cohort they studied (Fig. 8); some of them have a high hazard of failing early in their careers, whereas others have low initial hazard rates that grow over time. Gruenwald and West (1989) also found the proportional hazards model unsuitable in studying juvenile delinquency.

6.2. Recidivism as a Sequence of Failures

The revolving door of recidivism seems to turn faster for juveniles than for adults, since penalties for juveniles are of considerably shorter duration than for adults. As a consequence, chronic juvenile delinquents may have a long record of arrests with little incarceration time after court intervention to interrupt them. Murray and Cox (1979) analyzed the police contact data for a group of chronic delinquents and found that the postintervention police contact rates were substantially lower than their preintervention rates. They called this phenomenon the "suppression effect," because offenses appeared

[24]Recall that this is equivalent to the assumption of "stable criminality" proposed by Gottfredson and Hirschi and investigated by Harada (1991); see Section 4.1.

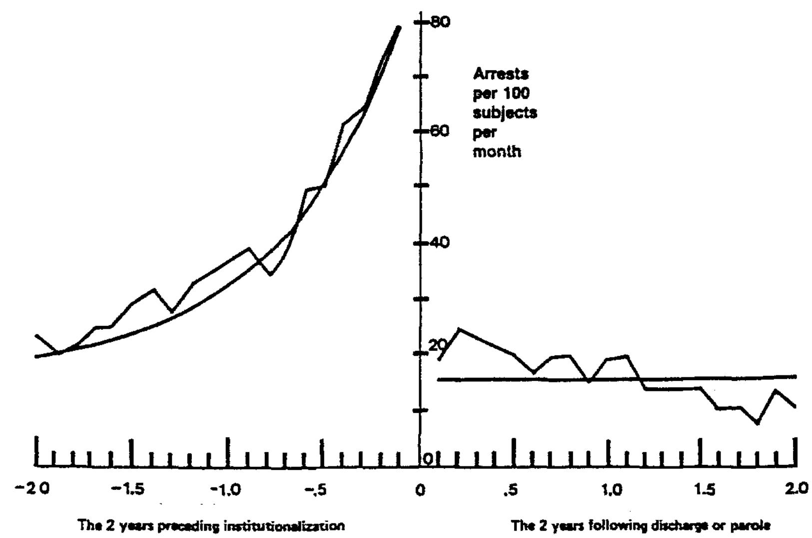


Fig. 13. Pre- and postintervention police contact rates, superimposed with a Markov model of the intervention process. [Source: Maltz (1984, p. 34).]

to be suppressed by the court-mandated intervention (Fig. 13). They argued that since arrests are a proxy for all crimes, a decrease in the arrest rate is tantamount to a decrease in the offense rate.

Maltz and Pollock (1980b) showed that this "suppression effect" is probably not based on true suppression (that is, a person actually increasing his/her offense rate before court intervention and curtailing it afterward) but could be adequately explained by the "regression to the mean" phenomenon.[25] They posited that, since an intervention *requires* a higher frequency of police contacts prior to intervention (Why else would an intervention be warranted?), and since police contacts are stochastic (i.e., not every crime results in a contact), the combination of these factors could generate the apperance of a steep rise immediately prior to the intervention.

They showed how this could be accomplished *even with constant* (but stochastic) *police contact rates both before and after intervention*, by either of two decision rules. In the first, offenders are assumed to commit crimes and incur arrests at a constant mean rate by switching between two rates

[25] *Regression to the mean* is caused by the *selection of extreme cases*. It can best be explained by a simple example. Suppose that 1000 fair coins are tossed three times each and 125 of them (or 1/8th, as expected) come up with all tails. We can call these coins "delinquent" coins, subject them to a treatment, and reexamine them by tossing them three more times. Note that most of them now are "cured;" only about 15 or so coins (1/8th again) are still "delinquent."

("active" and "quiescent") stochastically in a Markov process; in such a case, since the intervention occurs upon arrest, the arrest rate increases exponentially just prior to intervention, as depicted by the smooth curve in Fig. 13.

The second process that could explain the "suppression effect" is based on judicial selection criteria. If the process whereby offenders are selected for intervention is based generally on "k arrests in the past M months," then a constant prior arrest rate will again generate a steep rise just prior to intervention. This is easily explained: if preintervention activity must exceed a threshold for intervention to occur, then the average level prior to intervention will exceed the postintervention level. This "selection of the extreme" cases, followed by a "regression to the mean," makes it appear that the rate has changed drastically between before and after (Maltz, 1980; Maltz *et al.*, 1980; Tierney, 1983; Pollock and Farrell, 1984; Hwang, 1990).

7. DETERRENCE

Two types of deterrence are distinguished by Zimring and Hawkins (1973): *special* (or individual) deterrence and *general* deterrence. Special deterrence is the reduction in criminal activity by known offenders, due to the criminal sanctions previously imposed on them or to the threat of similar penalties should they continue to break the law. Operationally, it cannot be distinguished from rehabilitation, since it is difficult to determine whether the carrot of rehabilitation or the stick of punishment caused the change. Measuring the extent of special deterrence is accomplished by investigating the recidivism behavior of the offenders being studied; the techniques used are described in Section 6.

7.1. General Deterrence

General deterrence is the reduction in criminal activity by the general public (that is, actual and potential offenders) attributable to a given intervention. It is not associated with specific individuals, so it cannot be measured by tracing offenders' careers. Rather, it must be inferred from the available aggregate data. Some examples of interventions that may affect crime rates are the effect of the death penalty on homicide rates and the effect that increased penalties for gunrelated crimes might have on armed robbery—for a general overview, see Blumstein *et al.* (1978).

Mathematical modeling plays an important role in the study of general deterrence, although all too often its role is ignored. Inferring that a change in offense rates is attributable to a particular intervention requires that the effects of other possible explanatory variables be accounted for, which means that their effect on the outcome must be either known or modeled and that the measured change is sufficiently large that it can be distinguished from random variation.

One problem with inferring deterrence from data is that it can be confounded with an incapacitation effect: if a sizable fraction of the offending population is incarcerated, then offense rates will decrease due to incapacitation of the offenders, which cannot be distinguished from the reduction in offenses by those deterred. Cohen (1978, 1983) and Nagin (1978) used the Shinnar and Shinnar model to estimate the magnitude of the incapacitative effect on crime rates.

7.2. Linear Models of Deterrence

Nagin also reviewed a number of studies of general deterrence (see also Nagin, 1981). Almost all of the studies used official data and were cross-sectional in nature; almost all used linear models of one form or another (correlation, regression, ordinary least squares, two-stage least squares) to "control" for the effect of explanatory variables (such as income and age distributions of a jurisdiction, percentage nonwhite, and education levels). The justification for using linear models to relate these variables apparently was not discussed in any of the studies, perhaps because linear models are tractable and are assumed to be sufficiently "robust" (Maltz, 1994a). They are very often (implicitly) justified by virtue of the fact that, should a different relationship actually relate the variables, if that relationship is well behaved, then the Taylor series expansion of the function in a small neighborhood around any point is linear to a first-order approximation.[26]

For the most part, these studies appealed only to audiences of academics who were trying to determine whether changes in arrest rates, sentencing patterns, or police resources affected crime rates. This changed, and the study of deterrence and deterrents took on a greater importance, after a paper by Isaac Ehrlich (1975) on the deterrent effect of the death penalty hit the headlines.

7.3. Deterrence and the Death Penalty

Ehrlich analyzed data on homicide rates and the death penalty and concluded that the death penalty deterred homicide: his estimate was that every execution reduced the number of homicides by eight. This finding was

[26]This approximation, of course, holds only for the immediate neighborhood of the point. However, this limitation has not stopped people from extrapolating the solution to regions of the solution space for which the approximation does not hold.

not only published in an academic journal, it was cited by the U.S. Supreme Court in *Fowler v. North Carolina* and featured in newspaper stories.

His analysis was based on an econometric model and used national data from the Uniform Crime Reports from the 1930s to the 1970s. The functional form he used was the Cobb–Douglas econometric model, wherein his "murder supply function" was

$$M = kP_{\mathrm{a}}^{\alpha_1} P_{\mathrm{c|a}}^{\alpha_2} P_{\mathrm{e|c}}^{\alpha_3} U^{\beta_1} L^{\beta_2} Y_{\mathrm{p}}^{\beta_3} A^{\beta_4} e^{v}$$

where M is the murder rate, k is a constant, P_{a} is the probability of apprehension for a homicide, $P_{\mathrm{c|a}}$ is the probability of being convicted of homicide given apprehension, $P_{\mathrm{e|c}}$ is the probability of execution for homicide given conviction, U is the (civilian) unemployment rate, L is the fraction of adult civilians in the labor force, Y_{p} is an estimate of real per capita income, A is the fraction of the residential population between 14 and 24 years of age, and v represents the noise or random fluctuations. Taking the logarithm of this equation results in an equation that is linear in the model's parameters (the α's and β's). The probabilities were estimated by dividing, respectively, arrests for homicides by number of homicides, homicide convictions by homicide arrests, and executions for homicide by homicide convictions.

His finding was hotly contested, as others found weaknesses in the model, the data, and the method of analysis he used (Bowers and Pierce, 1975; Passell, 1975; Klein *et al.*, 1978; Barnett, 1978; Beyleveld, 1982; Nagin, 1978; Fisher and Nagin, 1978). For the most part, however, his critics believed that the deterrent effect of executions on homicides could be modeled, should the problems with data, model, and analytic tools be cleared up. Some of the model's problems included the following (Barnett, 1978).

- The Cobb–Douglas model may have been chosen more for its analytic tractability than its appropriateness.[27]
- How the variables were entered into the equation was not explained—Ehrlich used, for example, P_{a} and not $1-P_{\mathrm{a}}$ (the probability of *avoiding* arrest) in the equation, without justifying his choice.[28]
- The model was assumed to be time-invariant over four decades, implicitly ignoring the potential effects of a depression, of a world

[27] "Buttressed by an elegant theory and vast computer packages, regression models have great appeal in criminology and elsewhere. But mathematical models of a process ideally should arise from attempts to quantify reasonable assumptions about it, rather than from an excessive concern about tractability and familiarity" (Barnett, 1978, p. 302).

[28] Riccio (1974) makes a similar point, that deterrence may depend on the probability of getting away with the offense, not on the probability of arrest. This is especially true when the probability of arrest is small: even a 10-fold increase in the arrest probability will have little effect on an offender if the increase is from 0.001 to 0.010, since the probability of avoiding arrest drops from 0.999 to 0.990, or less than 1%.

war, of major postwar internal migration patterns, of civil rights legislation, of the increasing availability of automobiles and guns, and of other factors that doubtless exert some influence on the homicide rate.

7.4. Stochastic Modeling of Homicide Data

There was an even more fundamental problem with Ehrlich's model, however. Like any other statistic, the homicide rate has a certain natural (stochastic) variation. Barnett (1978, 1981a, b; see also Barnett, 1983a, b) developed an estimate of the expected "noise" in the homicide data. His approach is reminiscent of Poisson's investigation of jury behavior (Section 2.1, above); he modeled the stochastic process that results in homicide by investigating the structure of the data and making realistic assumptions about it.

Barnett assumes that there is a random variable h_i ($i=1, \ldots, N$) associated with each of the N persons in a jurisdiction, equal to the number of homicides he/she will commit in a given year. If p_{ij} is the probability that person i will commit j homicides (i.e., that $h_i=j$), then person i is expected to commit

$$E(h_i)=\bar{h}_i=\sum_{j=1}^{\infty} jp_{ij}$$

offenses with a variance of

$$\sigma^2(h_i)=\sum_{j=1}^{\infty} j^2p_{ij}-\bar{h}_i^2$$

Barnett then makes two assumptions: that $\bar{h}_i^2$ is negligible compared to $\bar{h}_i$ and that $\bar{h}_i$ is independent of $\bar{h}_j$ for $i\neq j$, both of which he shows to be reasonable. Then the expected number of total homicides in the jurisdiction for that year is

$$E(H)=E\left(\sum_{i=1}^{N} h_i\right)=\sum_i\sum_j jp_{ij}=\sum_j\left[\sum_i p_{i1}+2\sum_i p_{i2}+\cdots\right]$$
$$=[q_1+2q_2+3q_3+\cdots]$$

where q_j is the sum of the probabilities that individuals will commit j murders in a year and, therefore, is the expected number of the N individuals who commit j murders in a year.

$$q_j = \sum_i p_{ij}$$

Using his assumptions, the variance can be approximated by

$$\sigma^2(H) = \sum_i \sigma^2(h_i) = \sum_i (q_1 + 4q_2 + 9q_3 + \cdots)$$

Based on data from a number of states and cities on double, triple, and quadruple murderers, he was able to develop the estimates $q_2/q_1 \approx 0.0046$, $q_3/q_1 \approx 0.00016$, $q_4/q_1 \approx 0.0005$, etc., so that

$$\sigma^2(H) \approx 1.04E(H)$$

(This shows that the process is almost Poisson, since the variance almost equals the mean.) Barnett then showed that, for a typical state with 200 homicides per year (and an associated variance of about $200 \times 1.04 = 208$), even with the most favorable assumptions about the validity of the deterrence model used, *none* of the purported deterrent effects rises above the noise level inherent in the data.

7.5. Deterrence and Other Offenses

Research on deterrence models is not restricted to investigations of the death penalty. Other crimes are also subject to deterrence (and to studies of deterrence), which would benefit from knowledge of the extent of variation in expected rates. Extending his work on modeling the magnitude of fluctuations expected in the homicide rate, Barnett (1981b) developed a model of robbery. This model can be used in developing program evaluation guidelines.

Consider the case of an individual, say a chronic offender, who would, absent police intervention, commit c crimes in a given year. His/her offense "career" during that year, however, might be interrupted by an arrest during any one of the crimes, with probability q. Under the assumption that the offender is incapacitated for the rest of the year after apprehension, the probability distribution of c_{eff}, the number of crimes actually committed, is

$$c_{\text{eff}} = \begin{cases} i & \text{with probability} \quad q(1-q)^{i-1}, \qquad i = 1, \ldots, c-1 \\ c & \text{with probability} \quad (1-q)^c \end{cases}$$

It can be shown (Maltz, 1994b) that, in general, for offenders with high offense rates and low arrest probabilities, the number of offenses would average $(1-q)/q$ with a variance of $(1-q)/q^2$.

Suppose we were able to identify a group of 100 offenders whose (unconstrained) offense rate was 200/year and whose arrest probability was 0.1. They would be expected to produce 900 offenses each year with a variance of 9000 (and standard error of 95). To be 95% confident (one-tailed) that a program actually caused a reduction in offenses, there would have to be 156 fewer offenses, or a reduction of 17%!

Note that, absent arrest, these offenders would be expected to produce about 20,000 offenses, but only produce about 900 because of arrest and incarceration, and that the high intrinsic variance makes assessing change difficult: *even for a very active group of offenders whose activity for the year was assumed to terminate after arrest, there would have to be a very large reduction in their activity if a program aimed at reducing their offense rate were to be distinguished from the natural variation in their offense activity.*

8. POPULATION PROJECTIONS

8.1. Prison Population Projections

Another area of importance to prison administrators is the projection of future populations of offenders. As discussed earlier, the fact that prison populations are determined by judges, who do not always take cognizance of prison resource constraints, means that rational planning for prisons could not be carried out effectively without the development of some sort of model. Although there are a number of time series and regression models for forecasting crime (e.g., Fox, 1978; Orsagh, 1981), Stollmack (1973) developed one of the first probabilistic models for projecting prison populations. If prison arrivals have a Poisson arrival rate A, and sentences are distributed exponentially with mean $1/\Theta$, then the number of prisoners $N(t)$ in year t can be estimated as

$$N(t) = A(1-e^{-\Theta})/\Theta + e^{-\Theta}N(t-1)$$

A similar approach was used by Blumstein *et al.* (1980), who based their projection on Stollmack's model. They allowed A to vary over time according to the expected number of annual admissions for different age–race–crime categories. The number of admissions for each age–race category was estimated by using census data for the total population in each category, then multiplying this by the fraction of each category admitted to prison in the past for each offense type. They then aggregated the number for each category to produce their overall forecast.

In their basic form, such models suffer from the deficiency noted in Section 3, the inability to explore the characteristics of the workload. For example, a new policy may be to give all offenders 20-year sentences, thereby drastically reducing the number of recidivists; or a shortage of prison space might cause the sentence lengths to be reduced for many prisoners, increasing the number of potential offenders out on the street and, therefore, the number of individuals at risk for prison. In other words, these models rely on history to forecast the number of prison admissions, rather than allowing for the possibility that policy changes may affect admissions.

Barnett (1987; see also Rich and Barnett, 1985) developed a model that takes a number of these considerations into account. This model is based on the criminal career paradigm. It incorporates offense and arrest activity by specifying a mean arrest rate; offender recidivism behavior, by permitting specification of the age distribution of career termination; sentencing behavior, by specifying an age-specific probability distribution of imprisonment given arrest; and correctional behavior, by specifying an age-specific sentence length distribution.

Although this model substitutes one set of assumptions (past admission ratios are the best predictor of future admissions) with another (a constant arrest rate and stationarity in career termination and criminal justice behaviors), the new assumptions are more flexible and mirror reality more closely. This model, for example, shows how to incorporate the criminal career model into a population projection model. Second, this model accommodates offender characteristics and policy changes explicitly. Should changing employment patterns affect career termination, or should sentencing practices change,[29] this model can show how they affect the prison population.

Lattimore and Baker (1992) developed a model conceptually similar to that of Barnett, but based it on characteristics specific to the recidivism process. For example, since first-time offenders and recidivists have different recidivism characteristics (e.g., see Figs. 10 and 12), they disaggregated the prisoner population into "first-timers" and recidivists.[30] In addition, their model explicitly accounts for prison capacity constraints. Furthermore, they use the incomplete lognormal distribution to model return to prison; as noted in Section 6.1, this distribution is especially suitable for describing recidivism when return to prison is the indicator of failure.

When dealing with forecasts, the proof of the pudding can be set as far in the future as one wants, but eventually comes due. Although the Lattimore and Baker model provides a good forecast of commitments to prison, it

[29]Langan (1991) found that over half of the substantial growth in prison admissions over the past 15 years is attributable to the increased use of prison at sentencing. He suggests that this may also have contributed to the gradual reduction in the crime victimization rate during this period.

[30]The models posited by Barnett and by Blumstein *et al.* can also accommodate such dissaggregation.

stumbles over a common problem in forecasting prison populations, the lack of a stable policy and social environment. The changes that have occurred in recent years have drastically changed the criminal justice context: the increased prevalence of crack cocaine in urban areas and the consequent increase in drug crimes and in drug enforcement activity, the increased availability of money to focus on drug enforcement (via asset forfeiture), and the increased prison construction, to name but a few. Such changes make prediction in criminology a chancy occupation at best.

8.2. The "Stability of Punishment" Hypothesis

This same lack of stable conditions has also affected a conjecture about prison populations, that there is a "stability of punishment" in any given jurisdiction (Blumstein and Cohen, 1973). Similar to Quetelet's "*penchant au crime*" (see Section 2.1), it was hypothesized that a state or nation would tolerate the imprisonment of a certain percentage of its citizens: once a threshold or trigger level above this percentage is reached, the imprisonment rate decreases until it falls within the appropriate limits; and, similarly, if the imprisonment rate falls below a certain threshold, events conspire to increase the percentage to fall within the limits. In short, it was hypothesized that a society operated much like a thermostat, increasing or decreasing the punishment rate to keep it within the threshold limits of a set point. This phenomenon appeared to be present in the imprisonment statistics of a number of states and countries through the early 1970s (Blumstein *et al.*, 1976; Blumstein and Moitra, 1979).

Others disputed this interpretation of prison population statistics. They found different explanations that fit as well or found examples that did not fit the "stability of punishment" hypothesis (e.g., Greenberg, 1977; Rauma, 1981a, b; Blumstein *et al.*, 1981; Rauma, 1981b; Berk *et al.*, 1981, 1983). Since the time those articles were written, the growth in prison populations has been so high (Fig. 14) that they appeared to be "out of control" (Blumstein, 1988). Considering the recent changes in the prevalence of drug activity and enforcement of drug laws, one would not expect much stability in recent incarceration rates.

Although not a projection that would be of much benefit to prison administrators, in 1967 Christensen (1967) made a forecast of the offender population that made people aware of the major part the criminal justice system was beginning to play in the life of every citizen. He used a stochastic model of the offending population to project estimates of the fraction of the population that might experience an arrest or conviction for a criminal offense sometime during their lifetimes. Using a steady-state assumption about trends, he found that over half of the male population was eventually

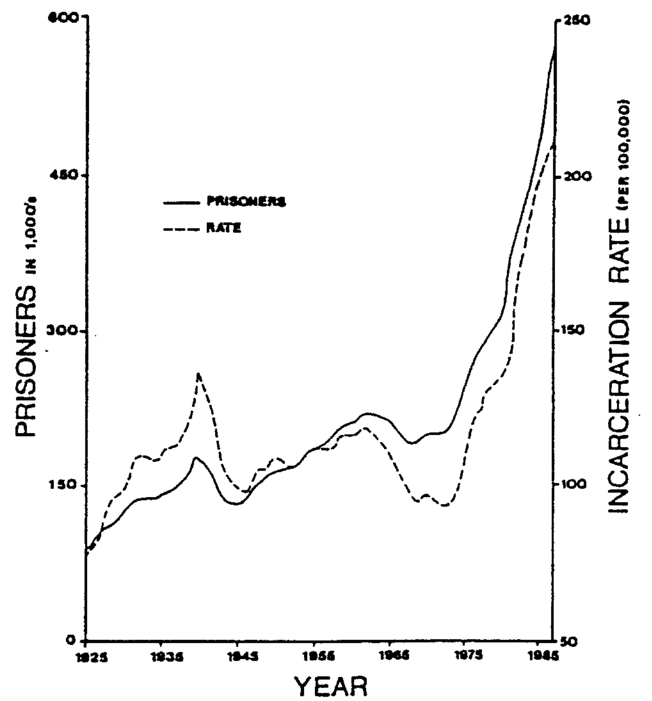


Fig. 14. Prison admissions and population in the United States. [Source: Blumstein (1988).]

likely to experience an arrest, a rate that rose to about 90% for nonwhite males. These statistics led others to explore the potential effect of high crime rates on prisons more closely.

8.3. Projecting the Lifetime Risk of Victimization

A line of research similar to that of Christensen's has focused on the probability that one would become a victim of crime. Using age-specific crime data from the New York City Police Department, Shinnar and Shinnar (1975) applied life-table techniques to show that the risk of becoming a victim of crime had increased substantially between 1960 and 1970 (with robbery increasing 10-fold during this period).

More recent work in this vein has focused on homicide. In a series of papers (Barnett *et al.*, 1975, 1980; Barnett, 1982; Barnett and Schwartz, 1989), Barnett and his colleagues showed that the lifetime risk of homicide for an individual born in an urban area is 1 in 68—and for a young urban African-American male the risk is far greater.

With statistics like these, economics can no longer claim to be the only "dismal science." Not only have they made headlines, but also these findings have led to policy changes. In part because of these studies, the Centers for Disease Control has recognized that homicides are the leading cause of death among young African–American adults and has begun to take an epidemiological approach to urban homicide.

8.4. Estimating the Size of the Criminal Population

Using the official statistics for Paris in 1825, Balzac estimated that "40,000 rogues, 15,000 petty thieves, 10,000 burglars and 40,000 ingenuous females living at someone else's expense add up to some 110 to 120,000 persons who are a great handicap to efficient administration" (Chevalier, 1973, p. 71). Although his estimate may have been based on faulty data and methods, the idea of expressing criminality in terms of such statistics brought a "statistical mentality" to the description of crime and the criminal population. Such estimates are important for other reasons as well.

For a jurisdiction experiencing 365 robberies a year, it makes a huge difference in terms of criminal justice policy if that number is the result of one offender plying his/her trade daily or 365 people pulling one job a year. In the first case, locking up one person is tantamount to solving the entire problem (by incapacitation); in the second case, locking up one person would make little difference. It is therefore important to develop methods for estimating the extent of participation of individuals in criminal careers and their frequency of committing crimes during their careers.

Greene and Stollmack (1981) developed such methods using models based on capture–recapture studies in ecology. They assumed that every offender experiences arrests according to the same Poisson distribution with parameter μ, such that the probability of a person experiencing k arrests in a year is

$$P(K=k)=\frac{\mu^k e^{-\mu}}{k!}$$

(Unlike the capture–recapture models, there is an implicit assumption that only the *offender* population is of interest, not all individuals. That is, the fact that most of the population would have an arrest rate μ of 0 is not explicitly handled, and no distinction is made between individuals for whom $\mu=0$ and those for whom $k=0$ but $\mu>0$.)[31]

Greene and Stollmack compared this with a model in which a fraction p is arrested at rate μ_1 and a fraction 1-p is arrested at rate μ_2, such that

[31] I thank Jacqueline Cohen for pointing this out.

the probability of a person experiencing k arrests is

$$P(K=k)=\frac{p\mu_1^k e^{-\mu_1}}{k!}+\frac{(1-p)\mu_2^k e^{-\mu_2}}{k!}$$

They used maximum-likelihood estimation techniques to fit the model to data from Washington DC, and found that the model that best fits the data is one that assumes that there are two subpopulations with different arrest rates: 4% of the offender population had a mean arrest rate of about 1.5 arrests per year (a mean time between arrests of 8 months), whereas the other 96% had a mean arrest rate of about 0.2 per year (for a mean time between arrests of 5 years).

Greene (1984) reanalyzed the same data under a different set of assumptions, including disaggregating the data into different age categories. He found that the model did not provide as good a fit under these conditions. Riccio and Finkelstein (1985) used arrest data to estimate the size of the burglar population, and Rossmo and Routledge (1990) used similar techniques, but different distributions, to estimate the size of the prostitute and fugitive population in Vancouver. Collins and Wilson (1990) used the Greene and Stollmack methods to estimate the number of automobile thieves in the Australian Capital Territory and found that the model based on two subpopulations provides the best fit to the data.

As with all models, this is predicated on assumptions about the population under study. This model assumes that there are no "births" or "deaths," i.e., the population is fixed, and the "capture" and "recapture" events occur independently of each other. Neither of these can be justified. The first assumption ignores the fact that in ecology the label "trout" is permanently attached to the animal, which is always subject to capture; but the label "criminal" applies only when the individual is active and subject to capture, at a rate dependent on his/her own λ, and ignores the possibility of his/her terminating his/her criminal career between capture and recapture.

The second assumption, independence, cannot be justified for two reasons. First, most police departments maintain files on prior arrestees, including their addresses and their *modi operandi*, just so they can increase the probability of arrest of these offenders should they commit additional crimes. It can also be argued that the more a person has been arrested the more he/she is liable to arrest, if only because his/her prior record reflects a lack of skill at his/her chosen career. (Of course, a third argument is that, in time, offenders may improve their skills and thus lower their probability of arrest.) The extent to which these explanations holds has not been ascertained; regardless, they serve to show that capture and recapture events are not independent.

8.5. Estimating the Number of Narcotics Users

Estimates of the prevalence of narcotics use in the United States vary widely, based on the political predilections of the estimators and on the methods they use as well. Hser *et al.* (1992) reviewed the different methods used to estimate the number of narcotics users. They distinguished between *static models* and *dynamic models*: static models are those that estimate the number of users at one point in time, while dynamic models generate estimates over time.

Among the static models they review are those based on straightforward *population projections* and those based on *closed-population capture models*. The former merely assumes that the usage patterns in an *unknown* population of known demographic characteristics are similar to those in a *known* population of known demographic characteristics, and estimates the number of users according to the relative demographic characteristics of the two populations.

The second static method, the closed-population capture model, is based on some of the same dubious assumptions that have been used to estimate the number of offenders: that capture samples are independent and that every individual has the same probability of capture. If n_{12} individuals are captured in each of two samples, then the population estimate is $n_1 n_2 / n_{12}$.

Among the dynamic models they review are the *open-population multiple capture* model, the *truncated Poisson estimation* model, *Markov open-population* models, and *system dynamics* models. The open-population multiple capture model is similar to the closed-population model, except that the population itself is not considered fixed in size but is integrated with a demographic birth and death process in which the population itself is changing its size and character.

The truncated Poisson model is based on the method used by Greene and Stollmack (1981) to estimate offender populations. In a given period of time some narcotics users are not arrested, while some are arrested once, some twice, some three times, etc. Based on the statistics of those arrested at least once, one can fit a Poisson distribution to the data and estimate the number of users who have avoided arrest. This method also is based on questionable assumptions, those of independence of arrests and homogeneity of users' arrest rates.

Markov open-population models consider the fact that individuals move into and out of different states; in some of those states they are at risk of capture, while in other states they are not. The model is calibrated using the statistics of known individuals and makes the implicit assumption that all individuals have the same transition probabilities.

System dynamics models have the advantage of incorporating feedback relationships among the variables, e.g., including the effect of increased enforcement on increasing arrest rates initially, on increasing prices, and on eventually reducing narcotics use (and therefore arrests) due to the increased prices. The difficulty with this method is that there are a great number of parameters to be estimated, and therefore it is difficult to ensure that the structure posited for the feedback relationships is correct, because so many different model structures could produce similar results.

9. PATTERNS IN CRIME

Operations research has also been used to study patterns of crime occurrences. This section describes recent efforts in the analysis of crime reporting patterns, geographical patterns, and enterprise-based crime.

9.1. Crime Reporting Patterns

Crime statistics in the United States are subject to errors, and sometimes the errors can be substantial. This is especially true of police-collected crime statistics, which are subject to changes in police policies that may have the inadvertent (or intentional) effect of changing reporting behavior. Consider what can happen when a police official, concerned about the amount of time spent by patrol officers writing burglary and larceny reports, decides to restrict them to taking reports of only those thefts that exceed $200; other victims must go to the police station to file reports. Although this policy may be a direct result of an *increased* crime rate, its effect may be to *reduce* the reported crime rate dramatically: making victims of less costly crimes go into the police station to report a crime may just convince them to forgo reporting it, especially if they are not insured or if their insurance deductible is $200. (This scenario is not hypothetical; it occurred in one large city in the 1970s.)

Police departments have also been found to have reduced the number of reported crimes by the simple expedient of unfounding them: after the initial report is filed by the patrol officer, a follow-up investigation is conducted by a detective. If, based on his/her judgment (or based on pressure from above), the detective feels that no crime took place or is unable to locate the victim or witnesses, that crime can be "unfounded" or treated as if it did not occur.

These irregularities are doubtless rarer now, and an independent set of crime data is now available at the national level. It was to provide a set of data not subject to police reporting biases that the Justice Department funded the Census Bureau to begin to collect victimization data from the general public. Reconciling the two data sets requires knowledge of the different methods used to collect them. Comparing the two measurement

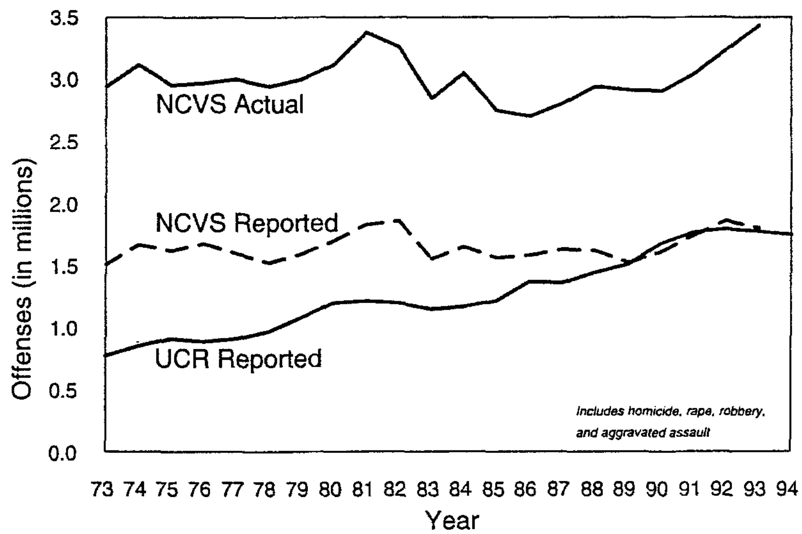


Fig. 15. Measures of violent crime. (Sources: Bureau of Justice Statistics, National Crime Victimization Survey, and Federal Bureau of Investigation, Uniform Crime Reports.)

processes, Maltz (1975) showed how some of the differences in methodologies can lead to incorrect conclusions. Eck and Riccio (1979) also compared the two data sets, attempting to reconcile them by modeling the extent to which the extent of nonreporting (to the police) accounted for the discrepancies between them.

The discrepancy between these two crime statistics series has now all but diminished, at least for violent crime. Figure 15 shows that the violent crime levels reported by the police to the FBI's Uniform Crime Reporting Program now reflect the same levels reported by victims on the National Crime Survey, for victimizations that victims subsequently reported to the police.[32]

9.2. Geographical Patterns of Crime

The study of geographical patterns has played a prominent role in criminology since Shaw and McKay (1969) first studied them in Chicago 60 years ago, showing the way juvenile delinquency rates seemed to spread

[32]Because the two series have different counting rules, adjustments have to be made to reconcile them. For example, commercial robberies were excluded from the UCR figures, since the National Crime Victimization Survey does not count them; victimizations occurring to individuals under 12 were excluded for the same reason; and homicides were added to the NCVS data.

out from the center of the city in concentric rings. Not only social scientists but also police departments have studied crime geographically, using pin maps to spot crime patterns. The problem with using such techniques in the past, for both the researcher and the police, was that they were very labor-intensive. Locating a few hundred events on a map is an arduous task. Even more difficult is attempting to determine what each pin represents. Although using different-colored pins to represent different types of offenses can help, it can only do so much: on the next day, a person would have a hard time remembering whether a particular pin represents a nighttime robbery of an elderly latina on a Wednesday or the daytime robbery of a young African-American male last Friday.

These problems have been all but eliminated by two technological facts. First, more and more police departments have computerized data available for analysis; and second, mapping software is now available that permits researchers and/or the police to plot selected data (e.g., nighttime robberies of elderly women) automatically, making it possible to infer different kinds of patterns from the data. Maltz *et al.* (1990) provide an overview of the potential uses of such patterning for police purposes. Block (1989) describes how the STAC (Spatial and Temporal Analysis of Crime) algorithm, developed by the Illinois Criminal Justice Information Authority for use by police departments, can be used to find crime "hot spots." Harries (1990; see also Harries, 1980) shows how mapping data can assist the police in other ways. And Rengert (1992; see also Rengert and Wasilchick, 1994) viewed the "journey to crime" as a set of transportation problems the offender must solve, depending on his/her anchor points—home, work, drug purchase location(s)—and the chances of recognition, for personal crimes.

Such data lend themselves to geographic models that have been developed by geographers for other purposes. Brantingham and Brantingham (1984) describe a number of techniques and models that can be used to discern geographic patterns of crime. They show how patterns change as the unit of geographical analysis changes (Brantingham *et al.*, 1976). In addition, they discuss the utility of distance-decay modeling (Capone and Nichols, 1976), gravity potential modeling (Smith, 1976; Rossmo, 1990), graph theory; and diffusion modeling to the analysis of such patterns. A recent summary of some recent research on crime mapping is given by Block *et al.* (1995).

9.3. Crime as a Business Enterprise

Many crimes are not committed with the relative spontaneity that marks the common burglary or street robbery, but are organized and planned like business enterprises. Modeling legal enterprises has been a strong element of traditional OR, and it is being applied as well to illegal enterprises such as traditional organized crime or drug markets or to legal enterprises operating illegally, as in white-collar crime.

9.3.1. *White-Collar and Organized Crime*

Maltz (1976) developed a model of the cigarette smuggling process to determine the effect of differential tax rates on the cigarette smuggling process. The model assumed that the amount of smuggling from state i to state j was based on the population of state j, the tax difference (smuggling occurs only if $t_j - t_i$ is positive), and the distance between their borders. This model was later applied (Maltz, 1981) to estimating the effect of a cigarette tax increase; since it estimated both the increase in revenue and the increase in smuggling, it showed how the state could in essence set the level of cigarette smuggling that would be generated by adjusting its tax rate.

The stochastic properties of bidding data, from road construction work, were analyzed by Maltz and Pollock (1980a) to infer collusion among bidders. Although their study developed no models, it presented the data in a way (by displaying the cumulative distribution of bids) that graphically depicted clear evidence of collusion.

Maltz (1990) employed graph theoretic concepts in developing measures to evaluate organized crime enforcement efforts. Since the essence of organized crime is that individual crimes are in some way connected, he suggested that the extent of their connectivity be used to determine if there is sufficient evidence from separate crimes to warrant labeling the crimes as part of an organized crime enterprise.

9.3.2. *Drug Markets and Enforcement Policies*

A number of aspects of drug markets have been investigated as enterprises, using various modeling techniques. Basing his model on reasonable assumptions about consumption, costs, and penalties, Caulkins (1993a) developed a model of drug users' purchasing patterns to consider the efficacy of "zero tolerance" to drug crimes, a policy that enforces the antidrug laws irrespective of the quantity of drugs involved in the crime. Although this policy may be predicated on minimizing the number of users, he shows that a punishment policy that is proportional to quantity reduces overall drug consumption.

Using the approach initiated by Becker (1967), Caulkins (1993b; Baveja *et al.*, 1993; Naik *et al.*, 1995) studied the relationship between street drug markets and police enforcement policies, in particular "crackdowns." Based on reasonable assumptions about the nature of street markets, they found that the best policy would be one that focused police efforts on causing individual markets to collapse rather than spreading their resources among many markets.

Modeling such interactions between police policies and drug market size is of great benefit, because all assumptions are explicit in the model formulation. Different police departments might have different results. For example, in some police departments suppressing them entirely is tantamount to killing the goose that lays the golden eggs. One police official (Frank Gajewski, personal communication) likens drug markets to vineyards that are harvested for arrests (and the consequent overtime for court appearances), the harvesters ensuring that the roots are still firmly in place.

Another advantage of mathematical models, and the explicitness with which they treat assumptions, is that they can be used to point out deficiencies in, and new uses for, data. Caulkins and Padman (1993) have shown that a log-linear model provides a reasonably good representation of the variation [due to purity, quantity sold, and geography (see also Caulkins, 1995)] in the cost of illegal drugs. This model was then used (Caulkins, 1994) to develop a time series of the price of cocaine, a necessary first step in evaluating the effectiveness of drug control efforts. In a parallel effort, Everingham and Rydell (1994) created a model of the demand for cocaine, which Rydell and Everingham (1994) then used as the basis for modeling the relative cost-effectiveness of different cocaine control programs. Their estimates show that treatment (one kind of demand control) is considerably more effective than interdiction and enforcement (supply control) programs.

10. CONCLUSION

This paper has reviewed the many applications of operations research to problems of crime and justice. As has been shown, the applications have been both considerable and successful. The systematic, quantitative thinking exemplified by these approaches to problem-solving has been very effective in helping to improve the administration of criminal justice, and it is likely that this effort will continue.

As more and more of the data problems are overcome in criminal justice, we will be awash in an even greater "avalanche of printed numbers" (Hacking, 1990), but numbers that will be far more amenable to analysis than in the past.

Looking to the future, one can foresee the need to develop better means of handling the streams of data. Mathematical models are one way of coming to grips with them. It might also be expected that, although analytic models will continue to be developed to provide the insight into crime and justice problems, more efforts will be devoted to understanding the data using maps and other graphical or pictorial techniques. The development of new analytic models might then be based on insights gained from being able to visually infer patterns in the data.

In addition, it is likely that more realistic models of offender behavior will be developed, recognizing that different offenders can be subject to different causal mechanisms (e.g., Moffitt, 1993). With the increases in data that are inevitable, it may be possible to develop multiple models, models that would reflect more closely the multiplicity of human behaviors that manifest themselves criminally.

Patterns of offending also reflect responses to criminal justice actions. Crime displacement (Maltz, 1972; Barr and Pease, 1990) may result from enforcement activity. More generally, it may be possible to model the decisions of victims and offenders; Felson (1994) and Cornish and Clarke (1986) have developed the "routine activities" and "rational offender" approaches, respectively, which have utility in crime-prevention research.

The implicit assumption in many studies of crime (and police deployment in response to crime) is that crimes and calls for service are stochastic processes. Although this model may serve adequately for many purposes, the validity of this assumption should be ascertained. As with organized crimes, it may well be that street crimes are interconnected in ways that would make them more amenable to prevention and/or deterrence. Investigating the "fine structure" of criminal activity to find and model patterns may yield benefits in terms of the understanding of crime and its prevention that are not possible using more traditional methods of analysis.

These other perspectives are bound to enrich criminological research. New insights, coupled with the expectation of increasing amounts of data, will require new analytic tools and multidisciplinary approaches and models, to help us advance our understanding of one of the major social problems of our time.

ACKNOWLEDGMENTS

This paper was completed while I was a Fulbright Scholar at El Colegio de Michoacán, Zamora, Michoacán, Mexico. I am indebted to Arnold Barnett, Jacqueline Cohen, David Farrington, Joseph Kadane, Richard Larson, Stephen Pollock, and Roger Tarling, for their advice and comments on an earlier version of this paper (Maltz, 1994b). Any mistakes and omissions, however, remain mine.

REFERENCES

Ahn, C. W., Blumstein, A., and Schervish, M. (1990). Estimation of arrest careers using hierarchical stochastic models. *J. Quant. Criminol.* 6: 131–152.

Avi-Itzhak, B., and Shinnar, R. (1973). Quantitative models in crime control. *J. Crim. Just.* 1(3): 185–217.

Barlow, R., and Proschan, F. (1975). *Statistical Theory of Reliability and Life Testing: Probability Models*, Holt, Rinehart and Winston, New York.

Barnett, A. (1978). Crime and capital punishment: Some recent studies. *J. Crim. Just.* 6: 291–303.

Barnett, A. (1981a). The deterrent effect of capital punishment: A test of some recent studies. *Operat. Res.* 29: 346–370.

Barnett, A. (1981b). Further standards of accountability in deterrence research. In Fox, J. A. (ed.), *Methods in Quantitative Criminology*, Academic Press, New York, pp. 127–145.

Barnett, A. (1982). Learning to live with homicide. *J. Crim. Just.* 10(1): 69–72.

Barnett, A. (1983a). The capital punishment controversy: Part I. *Interfaces* 13(3): 24–28.

Barnett, A. (1983b). The capital punishment controversy: Part II. *Interfaces* 13(5): 35–39.

Barnett, A. (1987). Prison populations: A projection model. *Operat. Res.* 35(1): 18–34.

Barnett, A., and Lofaso, A. J. (1985). Selective incapacitation and the Philadelphia Cohort data. *J. Quant. Criminol.* (1): 3–36.

Barnett, A., and Lofaso, A. J. (1986). On the optimal allocation of prison space. In Swersey, A. J. (ed.), *Delivery of Urban Services*, TIMS Series in the Management Sciences, Vol. 22, Elsevier North-Holland, Amsterdam.

Barnett, A., and Schwartz, E. (1989). Urban homicide: Still the same. *J. Quant. Criminol.* 5: 83–100.

Barnett, A., Kleitman, D., and Larson, R. C. (1975). On urban homicide: A statistical analysis. *J. Crim. Just.* 3: 85–110.

Barnett, A., Essenfeld, E., and Kleitman, D. (1980). Urban homicide: Some recent developments. *J. Crim. Just.* 8: 379–385.

Barnett, A., Blumstein, A., and Farrington, D. (1987). Probabilistic models of youthful criminal careers. *Criminology* 25: 83–108.

Barnett, A., Blumstein, A., and Farrington, D. (1989). A prospective test of a criminal career model. *Criminology* 27: 373–385.

Barnett, A., Blumstein, A., Cohen, J., and Farrington, D. (1992). Not all criminal career models are equally valid. *Criminology* 30: 133–140.

Barr, R., and Pease, K. (1990). Crime placement, displacement, and deflection. In Tonry, N., and Morris, M. (eds.), *Crime and Justice: A Review of Research, Vol. 12*, University of Chicago Press, Chicago.

Barton, R. R., and Turnbull, B. (1979). Evaluation of recidivism data: Use of failure rate regression models. *Evaluat. Q.* 3: 629–641.

Barton, R. R., and Turnbull, B. (1981). A failure rate regression model for the study of recidivism. In Fox, J. A. (ed.), *Models in Quantitative Criminology*, Academic Press, New York, pp. 81–101.

Baveja, A., Batta, R., Caulkins, J. P., and Karwan, M. H. (1993). Modeling the response of illicit drug markets to local enforcement. *J. Socio-Econ. Plan. Sci.* 27(2): 73–89.

Becker, G. S. (1967). Crime and punishment: An economic approach. *J. Pol. Econ.* 78: 526–536.

Belkin, J., Blumstein, A., Glass, W., and Lettre, M. (1972). JUSSIM, an interactive computer program and its uses in criminal justice planning. In Cooper, G. (ed.), *Proceedings of the International Symposium on Criminal Justice Information and Statistics Systems*, Project SEARCH, Sacramento, CA, pp. 467–477.

Belkin, J., Blumstein, A., and Glass, W. (1973). Recidivism as a feedback process: An analytical model and empirical validation. *J. Crim. Jus.* 1: 7–26.

Berecochea, J. E., Himelson, A. N., and Miller, D. E. (1972). The risk of failure during the early parole period: A methodological note. *J. Crim. Law Criminol. Police Sci.* 63(1): 93–97.

Berk, R. A., Rauma, D., Messinger, S. L., and Cooley, T. F. (1981). A test of the stability of punishment hypothesis: The case of California, 1851–1970. *Am. Sociol. Rev.* 45: 805–829.

Berk, R. A., Messinger, S. L., Rauma, D., and Berecochea, J. E. (1983). Prisons as self-regulating systems: A comparison of historical patterns in California for male and female offenders. *Law Soc. Rev.* 17(4): 547–586.

Beyleveld, D. (1982). Ehrlich's analysis of deterrence. *Br. J. Criminol.* 22: 101–123.

Bloch, A. (1977). *Murphy's Law and Other Reasons Why Things Go Wrong*, Price/Stern/Sloan, Los Angeles.

Block, C. R. (1989). *Spatial and Temporal Analysis of Crime: Users Manual and Technical Manual*, Illinois Criminal Justice Information Authority, Chicago.

Block, C. R., and van der Werff, C. (1991). *Initiation and Continuation of a Criminal Career: Who Are the Most Active and Dangerous Offenders in the Netherlands?* Ministry of Justice, WODC, The Hague.

Block, C. R., Dabdoub, M., and Fregly, S. (eds.) (1995). *Crime Analysis Through Computer Mapping*, Police Executive Research Forum, Washington, DC.

Blumstein, A. (1983). Selective incapacitation as a means of crime control. *Am. Behav. Sci.* 27: 87–108.

Blumstein, A. (1988). Prison populations: A system out of control? In Morris, N., and Tonry, M. (eds.), *Crime and Justice: An Annual Review of Research, Vol. 10*, University of Chicago Press, Chicago, pp. 231–266.

Blumstein, A., and Cohen, J. (1973). A theory of the stability of punishment. *J. Crim. Law Criminol. Police Sci.* 63(2): 198–207.

Blumstein, A. and Cohen, J. (1979). Estimating of individual crime rates from arrest records. *J. Crim. Law Criminol.* 70(4): 561–585.

Blumstein, A., and Cohen, J. (1987). Characterizing criminal careers. *Science* 238(4818): 985–991.

Blumstein, A., and Larson, R. C. (1967). A systems approach to the study of crime and criminal justice. In Morse, P. M., and Bacon, L. (eds.), *Operations Research for Public Systems*, MIT Press, Cambridge, MA, pp. 159–180.

Blumstein, A., and Larson, R. C. (1969). Models of a total criminal justice system. *Operat. Res.* 17(2): 219–232.

Blumstein, A., and Larson, R. C. (1971). Problems in modeling and measuring recidivism. *J. Res. Crime Delinq.* 8: 124–132.

Blumstein, A., and Moitra, S. (1979). An analysis of the time series of the imprisonment rate in the states of the United States: A further test of the stability of punishment hypothesis. *J. Crim. Law Criminol.* 70(3): 376–390.

Blumstein, A., and Moitra, S. (1980). The identification of 'career criminals' from 'chronic offenders' in a cohort. *Law Policy Q.* 2: 321–334.

Blumstein, A., and Nagin, D. (1978). On the optimum use of incarceration for crime control. *Operat. Res.* 26(3): 381–403. [Reprinted (with modifications) in Fox (1981b), pp. 39–59].

Blumstein, A., Cohen, J., and Nagin, D. (1976). The dynamics of a homeostatic punishment process. *J. Crim. Law Criminol.* 67(3): 317–334.

Blumstein, A., Cohen, J., and Nagin, D. (1978). *Deterrence and Incapacitation: Estimating the Effects of Criminal Sanctions on Crime Rates*, National Academy of Sciences, Washington, DC.

Blumstein, A., Cohen, J., and Miller, H. (1980). Demographically disaggregated projections of prison populations. *J. Crim. Just.* 8(1): 1–25.

Blumstein, A., Cohen, J., Moitra, S., and Nagin, D. (1981). On testing the stability of punishment hypothesis: A reply. *J. Crim. Law Criminol.* 72(4): 1799–1808.

Blumstein, A., Cohen, J., and Hsieh, P. (1982). *The Durations of Adult Criminal Careers*, Final Report to the National Institute of Justice, Carnegie-Mellon University, Pittsburgh, PA.

Blumstein, A., Farrington, D. P., and Moitra, S. (1985). Delinquency careers: Innocents, desisters, and persisters. In Tonry, M., and Morris, N. (eds.), *Crime and Justice, Vol. 6,* University of Chicago Press, Chicago.

Blumstein, A., Cohen, J., Roth, J. A., and Visher, C. A. (eds.) (1986). *Criminal Careers and Career Criminals, Vols. I and II,* National Academy Press, Washington, DC.

Blumstein, A., Cohen, J., and Farrington, D. P. (1988a). Criminal career research: Its value for criminology. *Criminology* 26(1): 1–35.

Blumstein, A., Cohen, J., and Farrington, D. P. (1988b). Criminal career research: Further clarifications. *Criminology* 26(1): 57–74.

Blumstein, A., Cohen, J., Das, S., and Moitra, S. D., (1988c). Specialization and seriousness during adult criminal careers. *J. Quant. Criminol.* 4: 303–345.

Blumstein, A., Canela-Cacho, J. A., and Cohen, J. (1993). Filtered sampling from populations with heterogeneous event frequencies. *Manage. Sci.* 39(7): 886–899.

Bowers, W. J., and Pierce, G. L. (1975). The illusion of deterrence in Isaac Ehrlich's research on capital punishment, *Yale Law J.* 85: 187–208.

Brantingham, P., and Brantingham, P. (1984). *Patterns in Crime,* Macmillan, New York.

Brantingham, P. J., Dyreson, D. A., and Brantingham, P. L. (1976). Crime seen through a cone of resolution. *Am. Behav. Sci.* 20: 261–273.

Broadhurst, R. R., and Maller, R. (1991). Estimating the numbers of prison terms in criminal careers from the one-step probabilities of recidivism. *J. Quant. Criminol.* 7: 275–290.

Broadhurst, R. R., Maller, R., Maller, M., and Duffecy, J. (1988). Aboriginal and non-aboriginal recidivism in Western Australia: A failure-rate analysis. *J. Res. Crime and Delinq.* 25: 83–108.

Brounstein, S. H., and Kamrass, M. (1976). *Operations Research in Law Enforcement, Justice, and Societal Security,* Lexington Books, Lexington, MA.

Capone, D., and Nichols, W. J. (1976). Urban structure and criminal mobility. *Am. Behav. Sci.* 20: 199–213.

Carr-Hill, G. A., and Carr-Hill, R. A. (1972). Reconviction as a process. *Br. J. Criminol.* 12: 35–43.

Cassidy, R. G. (1985). Modelling a criminal justice system. In Farrington, D. P., and Tarling, R. (eds.), *Prediction in Criminology,* State University of New York Press, Albany, pp. 193–207.

Caulkins, J. P. (1993a). Local drug markets' response to focused police enforcement. *Operat. Res.* 41(5): 848–863.

Caulkins, J. P. (1993b). Zero tolerance: Could it increase the consumption of illegal drugs. *Manage. Sci.* 39(4): 458–476.

Caulkins, J. P. (1994). *Developing Price Series for Cocaine,* Report MR-317-DPRC, Rand Corp., Santa Monica, CA.

Caulkins, J. P. (1995). Domestic geographic variation in illicit drug prices. *J. Urban Econ.* 37: 38–56.

Caulkins, J. P., and Padman, R. (1993). Quantity discounts and quality premia for illicit drugs. *J. Am. Stat. Assoc.* 88(423): 748–757.

Chaiken, J. M., and Chaiken, M. R. (1982). *Varieties of Criminal Behavior,* Report R-2814-NIJ, Rand Corp., Santa Monica, CA.

Chaiken, J., Crabill, T., Holliday, L., Jaquette, D., Lawless, M., and Quade, E. (1975). *Criminal Justice Models: An Overview,* Report R-1859-DOJ, Rand Corp., Santa Monica, CA.

Chevalier, L. (1973). *Laboring Classes and Dangerous Classes in Paris During the First Half of the Nineteenth Century,* Princeton University Press, Princeton, NJ [Translation (by F. Jellinek) of *Classes Laborieuses et Classes Dangereuses à Paris Pendant la Première Moitié du XIX^e Siècle,* Librairie Plon, Paris, France, 1958.]

Christensen, R. (1967). Projected percentage of U.S. population with criminal arrest and conviction records. *Science and Technology Task Force Report*, Appendix J, pp. 216–228.

Cohen, J. (1978). The incapacitative effect of imprisonment: A critical review of the literature. In Blumstein, A., Cohen, J., and Nagin, D. (eds.), *Deterrence and Incapacitation: Estimating the Effects of Criminal Sanctions on Crime Rates*, National Academy of Sciences, Washington, DC, pp. 187–243.

Cohen, J. (1983). Incapacitation as a strategy for crime control: Possibilities and pitfalls. In Tonry, M., and Morris, N. (eds.), *Crime and Justice, Vol. 5*, University of Chicago Press, Chicago, pp. 1–84.

Cohn, E. G., and Farrington, D. P. (1990). Differences between British and American criminology: An analysis of citations. *Br. J. Criminol.* 30: 467–482.

Collins, M. F., and Wilson, R. M. (1990). Automobile theft: Estimating the size of the criminal population. *J. Quant. Criminol.* 6(4): 395–409.

Copas, J. B., and Tarling, R. (1986). Some methodological issues in making predictions. In Blumstein, A., *et al.*, (eds.), *Criminal Careers and Career Criminals, Vol. II*, National Academy Press, Washington, DC, pp. 291–313.

Copas, J. B., and Tarling, R. (1988). Stochastic models for analyzing criminal careers. *J. Quant. Criminol.* 4: 173–186.

Cornish, D. B., and Clarke, R. V. (1986). *The Reasoning Criminal*, Springer-Verlag, New York.

Cox, D. R., and Oakes, D. (1984). *Analysis of Survival Data*, Chapman and Hall, London, England.

Daston, L. (1988). *Classical Probability in the Enlightenment*, Princeton University Press, Princeton, NJ.

DeGroot, M. H., Fienberg, S. E., and Kadane, J. B. (eds.) (1986). *Statistics and the Law*, Wiley, New York.

Eck, J. E., and Riccio, L. J. (1979). Relationship between reported crime rates and victimization survey results: An empirical and analytical study. *J. Crim. Just.* 7: 293–308.

Ehrlich, I. (1975). The deterrent effect of capital punishment: A question of life and death. *Am. Econ. Rev.* 65: 397–417.

Ellerman, R., Sullo, P., and Tien, J. M. (1992). An alternative approach to modeling recidivism using quantile residual life functions. *Opera. Res.* 40: 485–504.

Everingham, S. S., and Rydell, C. P. (1994). *Modeling the Demand for Cocaine*, Report MR-332-ONDCP/A/DPRC; Rand Corp., Santa Monica, CA.

Farrington, D. P., Homel, R., Junger, M., and Wikstrom, P.-O. (1996). Quantitative criminology in the English-speaking world outside North America. *J. Quant. Criminol.* (to appear).

Farrington, D. P., and Hawkins, J. D. (1991). Predicting participation, early onset and later persistence in officially recorded offending. *Crim. Behav. Ment. Health* 1: 1–33.

Farrington, D. P., and Tarling, R. (eds.) (1995). *Prediction in Criminology*, State University of New York Press, Albany.

Felson, M. (1994). *Crime and Everyday Life: Insights and Implications for Society*, Pine Forge Press, Thousand Oaks, CA.

Finkelstein, M. O. (1978). *Quantitative Methods in Law*, Free Press, New York.

Fisher, F. M., and Nagin, D. (1978). On the feasibility of identifying the crime function in a simultaneous model of crime rates and sanction levels. In Blumstein, A., *et al.* (eds.), *Deterrence and Incapacitation: Estimating the Effects of Criminal Sanctions on Crime Rates*, National Academy of Sciences, Washington, DC, pp. 361–399.

Fox, J. A. (1978). *Forecasting Crime Data*, D. C. Heath, Lexington, MA.

Fox, J. A. (ed.) (1981a). *Methods in Quantitative Criminology*, Academic Press, New York.

Fox, J. A. (ed.) (1981b). *Models in Quantitative Criminology*, Academic Press, New York.

Gerchak, Y., and Kubat, P. (1986). Patterns and dynamics of population heterogeneity in mixtures models. *Quality Quant.* 120: 285–291.

Gigerenzer, G., Zeno, S., Porter, T., Daston, L., Beatty, J., and Kruger, L. (1989). *The Empire of Chance*, Cambridge University Press, Cambridge, England.

Gottfredson, M. R., and Gottfredson, D. M. (1987). *Decision Making in Criminal Justice: Toward the Rational Exercise of Discretion*, 2nd ed., Plenum Press, New York.

Gottfredson, M., and Hirschi, T. (1986). The true value of lambda would appear to be zero: An essay on career criminals, criminal careers, selective incapacitation, cohort studies, and related topics. *Criminology* 24(2): 213–234.

Gottfredson, M., and Hirschi, T. (1987). The methodological adequacy of longitudinal research on crime. *Criminology* 25(4): 581–614.

Gottfredson, M., and Hirschi, T. (1988). Science, public policy, and the career paradigm. *Criminology* 26(1): 37–56.

Gottfredson, M., and Hirschi, T. (1989). A propensity-event theory of crime. In Laufer, W. S., and Adler, F. (eds.), *Advances in Criminological Theory*, Transaction, New Brunswick, NJ.

Gottfredson, M., and Hirschi, T. (1990). *A General Theory of Crime*, Stanford University Press, Stanford, CA.

Gottfredson, S. D., and Gottfredson, D. M. (1986). Accuracy of prediction models. In Blumstein, A., *et al.* (eds.), *Criminal Careers and Career Criminals, Vol. II*, National Academy Press, Washington, DC, pp. 212–290.

Greenberg, D. F. (1975). The incapacitative effect of imprisonment: Some estimates. *Law Society Rev.* 9(4): 541–580.

Greenberg, D. F. (1977). The dynamics of oscillatory punishment processes. *J. Crim. Law Criminol. Police Sci.* 68(4): 643–651.

Greenberg, D. F. (1978). Recidivism as radioactive decay. *J. Res. Crime Delinq.* 24: 124–125.

Greenberg, D. F. (1979). *Mathematical Criminology*, Rutgers University Press, New Brunswick, NJ.

Greenberg, D. F. (1991). Modeling criminal careers. *Criminology* 29(1): 17–46.

Greenberg, D. F. (1992). Comparing criminal career models. *Criminology* 30(1): 141–147.

Greene, M. A. (1984). Estimating the size of the criminal population using an open population approach. *Proc. Am. Stat. Assoc.*, Suv. Res. Methods Sect. 8–13.

Greene, M. A., and Stollmack, S. (1981). Estimating the number of criminals. In Fox, J. A. (ed.), *Models in Quantitative Criminology*, Academic Press, New York, pp. 1–24.

Greenwood, P. W., with Abrahamse, A. (1982). *Selective Incapacitation*, Report R-2815-NIJ, Rand Corp., Santa Monica, CA.

Greenwood, P., and Turner, S. (1987). *Selective Incapacitation Revisited: Why the High-Rate Offenders Are Hard to Predict*, Report R-3397-NIJ, Rand Corp., Santa Monica, CA.

Gruenwald, P. J., and West, B. (1989). Survival models of recidivism among juvenile delinquents. *J. Quant. Criminol.* 5: 215–229.

Hacking, I. (1990). *The Taming of Chance*, Cambridge University Press, Cambridge, England.

Harada, Y. (1991). Testing the "stable criminality" hypothesis with Japanese data. Paper delivered at the 50th Meeting of the American Society of Criminology, San Francisco, CA, Nov. 22.

Harries, K. D. (1980). *Crime and the Environment*, Charles C Thomas, Springfield, IL.

Harries, K. D. (1990). *Geographic Factors in Policing*, Police Foundation, Washington, DC.

Harris, C. M., and Stollmack, S. (1976). Failure-rate analysis in correctional systems. In Brounstein, S. H., and Kamrass, M. (eds.), *Operations Research in Law Enforcement, Justice, and Societal Security*, Lexington Books, Lexington, MA, pp. 143–153.

Harris, C. M., Kaylan, A. R., and Maltz, M. D. (1981). Recent advances in the statistics of recidivism measurement. In Fox, J. A. (ed.), *Models in Quantitative Criminology*, Academic Press, New York, pp. 61–79.

Hodges, J. S. (1991). Six (or so) things you can do with a bad model. *Operat. Res.* 39: 355–365.

Hoffman, P. B., and Adelberg, S. (1980). The salient factor score: A nontechnical overview. *Fed. Probat.* 44: 44–52.

Hoffman, P. B., and Beck, J. L. (1974). Parole decision-making: A salient factor score. *J. Crim. Just.* 2: 195–206.

Hser, Y.-I., Anglin, M. D., Wickens, T. D., Brecht, M.-L. and Homer, J. (1992). *Techniques for the Estimation of Illicit Drug-Use Prevalence*, Research Report NCJ 133786, National Institute of Justice, Washington, DC, May.

Hwang, S.-A. (1990). Modeling the suppression effect of correctional programs on juvenile delinquency. *J. Quant. Criminol.* 6: 377–393.

Illinois Criminal Justice Information Authority (1991). Despite improvements, many rap sheets still missing disposition information. *The Compiler*, Spring Issue, ICJIA, Chicago.

Institute for Law and Justice (1991). *CJSSIM: Criminal Justice System Simulation Model. Software and User Manual*, Institute for Law and Justice, Alexandria, VA.

Jefferys, W. H., and Berger, J. O. (1992). Ockham's razor and Bayesian analysis. *Am. Sci.* 80: 64–72.

Johnson, B. D., Goldstein, P. J., Preble, E., Schmeidler, J., Lipton, D. S., Spunt, B., and Miller, T. (1985). *Taking Care of Business: The Economics of Crime by Heroin Abusers*, Lexington Books, Lexington, MA.

Kalbfleisch, J. D., and Prentice, R. L. (1980). *The Statistical Analysis of Failure Time Data*, Wiley, New York.

Klein, L. R., Forst, B., and Filatov, V. (1978). The deterrent effect of capital punishment: An assessment of the estimates. In Blumstein, A., *et al.* (eds.), *Deterrence and Incapacitation: Estimating the Effects of Criminal Sanctions on Crime Rates*, National Academy of Sciences, Washington, DC, pp. 336–360.

Koopman, B. O. (1986). An empirical formula for visual search. *Operat. Res.* 34(3): 377–383.

Land, K. C. (1992). Models of criminal careers: Some suggestions for moving beyond the current debate. *Criminology* 30: 149–155.

Langan, P. A. (1991). America's soaring prison population. *Science* 251: 1568–1573.

Larson, R. C. (1972). *Urban Police Patrol Analysis*, MIT Press, Cambridge, MA.

Lattimore, P. K., and Baker, J. R. (1992). The impact of recidivism and capacity on prison populations: A projection model. *J. Quant. Criminol.* 8: 155–173.

Lehoczky, J. P. (1986). Random paramete stochastic-process models of criminal careers. In Blumstein, A., *et al.* (eds.), *Criminal Careers and Career Criminals*, National Academy Press, Washington, DC, pp. 380–404.

Maltz, M. D. (1972). *Evaluation of Crime Control Programs*, National Institute of Law Enforcement and Criminal Justice, U.S. Department of Justice, Washington, DC.

Maltz, M. D. (1975). Crime statistics: A mathematical perspective. *J. Crim. Just.* 3(3): 177–194.

Maltz, M. D. (1976). On the estimation of smuggling in a "gray market" commodity, *Operat. Res.* 23(6): 1156–1163.

Maltz, M. D. (1977). Crime statistics: A historical perspective. *Crime and Delinq.* 23(1): 32–40.

Maltz, M. D. (1980). Beyond suppression: More sturm and drang on the correctional front. *Crime Delinq.* 26(3): 389–397.

Maltz, M. D. (1981). Transportation modeling in analyzing an economic crime. In Fox, J. A. (ed.), *Methods in Quantitative Criminology*, Academic Press, New York, pp. 77–97.

Maltz, M. D. (1984). *Recidivism*, Academic Press, Orlando, FL.

Maltz, M. D. (1990). *Measuring the Effectiveness of Organized Crime Control Efforts* (monograph), Office of International Criminal Justice, University of Illinois at Chicago, Chicago.

Maltz, M. D. (1994a). Deviating from the mean: The declining significance of significance. *J. Res. Crime Delinq.* 31(4): 434–463.

Maltz, M. D. (1994b). Operations research in studying crime and justice: Its history and accomplishments. In Pollock, S. M., Barnett, A., and Rothkopf, M. H. (eds.), *Operations Research and the Public Sector* (Volume 6 of the Handbooks in Operations Research and Management Science, edited by G. L. Nemhauser and A. H. G. Rinnooy Kan), Elsevier North-Holland, Amsterdam, pp. 201–262.

Maltz, M. D., and McCleary, R. (1977). The mathematics of behavioral change: Recidivism and construct validity, *Eval. Q.* 1(3): 421–438.

Maltz, M. D., and Pollock, S. M. (1980a). Analyzing suspected collusion among bidders. In Geis, G. and Stotland, E. (eds.), *White-Collar Crime: Theory and Practice*, Sage: Beverly Hills, CA, pp. 174–198.

Maltz, M. D., and Pollock, S. M. (1980b). Artificial inflation of a delinquency rate by a selection artifact. *Operat. Res.* 28(3): 547–559.

Maltz, M. D., and McCleary, R. (1978). Comments on "Stability of parameter estimates in the split-population exponential distribution." *Eval. Q.* 2(4): 650–654.

Maltz, M. D., McCleary, R., and Pollock, S. M. (1979). Recidivism and likelihood functions: A reply to Stollmack. *Eval. Q.* 2(1): 124–131.

Maltz, M. D., Gordon, A. C., McDowall, D., and McCleary, R. (1980). An artifact in pretest-posttest designs: How it can mistakenly make delinquency programs look effective. *Eval. Rev.* 4(2): 225–240.

Maltz, M. D., Gordon, A. C., and Friedman, W. (1990). *Mapping Crime in Its Community Setting: Event Geography Analysis*, Springer-Verlag, New York.

Mason, W. M., and Fienberg, S. E. (eds.) (1985). *Cohort Analysis in the Social Sciences: Beyond the Identification Problem*, Springer-Verlag, New York.

McEwen, J. T., and Guynes, R. (1990). Criminal justice system simulation model (CJSSIM). Paper presented at the Annual Meeting of the American Society of Criminology, Baltimore, MD.

Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychol. Rev.* 100(4): 674–701.

Monahan, J., and Walker, L. (1984). *Social Science in Law: Cases and Materials*, Foundations Press, Mineola, NY.

Morgan, P. M. (1985). *Modelling the Criminal Justice System*, Home Office Research and Planning Unit Paper 35, Home Office, London.

Murray, C. A., and Cox, L. A., Jr. (1979). *Beyond Probation: Juvenile Corrections and the Chronic Delinquent*, Sage, Beverly Hills, CA.

Nagin, D. (1978). General deterrence: A review of the empirical evidence. In Blumstein, A., *et al.* (eds.), *Deterrence and Incapacitation: Estimating the Effects of Criminal Sanctions on Crime Rates*, National Academy of Sciences, Washington, DC, pp. 95–139.

Nagin, D. (1981). Crime rates, sanction levels and constraints on prison population. *Law Society Rev.* 12(3): 341–366. [Reprinted as Methodological issues in estimating the deterrent effect of sanctions. In Fox (1981b), pp. 121–140.

Nagin, D., and Smith, D. A. (1990). Participation in and frequency of delinquent behavior: A test for structural differences. *Criminology* 28: 335–356.

Naik, A. V., Baveja, A., Batta, R., and Caulkins, J. P. (1995). Scheduling crackdowns on illicit drug markets. *Eur. J. Operat. Res.* (in press).

New York State Identification and Intelligence System (1967). *NYSIIS: System Development Plan*, NYSIIS, Albany.

Orsagh, T. (1981). A criminometric model of the criminal justice system. In Fox, J. A. (ed.), *Models in Quantitative Criminology*, Academic Press, New York, pp. 163–187.

Partanen, J. (1969). On waiting time distributions. *Acta Sociol.* 112: 132–143.

Passell, P. (1975). The deterrent effect of the death penalty: A statistical test. *Stanford Law Rev.* 28: 61–80.

Philpotts, G. J. O., and Lancucki, L. B. (1979). *Previous Convictions, Sentence and Reconviction*, Home Office Research Study No. 53, Her Majesty's Stationery Office, London.

Pollock, S. M., and Farrell, R. L. (1984). Past intensity of a terminated poisson process. *Operat. Res. Lett.* 2: 261-263.

Pollock, S. M., and Maltz, M. D. (1994). Operations research in the public sector: An introduction and a brief history. In Pollock, S. M., Barnett, A., and Rothkopf, M. H. (eds.), *Operations Research and the Public Sector* (Volume 6 of the Handbooks in Operations Research and Management Science, edited by G. L. Nemhauser and A. H. G. Rinnooy Kan), Elsevier North-Holland, Amsterdam, pp. 1-22.

Porter, T. M. (1986). *The Rise of Statistical Thinking, 1820–1900*, Princeton University Press, Princeton, NJ.

President's Commission on Law Enforcement and Administration of Justice (1967). *The Challenge of Crime in a Free Society*, U.S. Government Printing Office, Washington, DC.

Proschan, F. (1963). Theoretical explanation of observed decreasing failure rate. *Technometrics* 5: 375–383.

Pullinger, H. (1985). *The Criminal Justice System Model: The Flow Model*, Home Office Research and Planning Unit Paper 36, Home Office, London.

Quetelet, L. A. J. (1835). *A Treatise on Man* (translated by S. Diamond), Scholars Facsimiles and Reprints, Gainesville, FL, 1969.

Rauma, D. (1981a). Crime and punishment reconsidered: Some comments on Blumstein's stability of punishment hypothesis. *J. Crim. Law Criminol. Police Sci.* 72(4): 1772–1798.

Rauma, D. (1981b). A concluding note on the stability of punishment: A reply to Blumstein, Cohen, Moitra, and Nagin. *J. Crim. Law Criminol. Police Sci.* 72(4): 1809–1812.

Reiss, A. J., Jr., and Farrington, D. P. (1991). Advancing knowledge about co-offending: Results from a prospective longitudinal survey of London males. *J. Crim. Law Criminol.* 82: 360–395.

Rengert, G. F. (1992). The journey to crime: Conceptual foundations and policy implications. In Evans, D. J., Fyfe, N. R., Nicholas, R., and Herbert, D. T. (eds.), *Crime, Policing and Place*, Routledge, London.

Rengert, G. F., and Wasilchick, J. (1994). Space, Time and Crime. Ethnographic Insights into Residential Burglary. Final Report submitted to U.S. Department of Justice, National Institute of Justice, and Office of Justice Programs, Washington, DC.

Rhodes, W. (1989). The criminal career: Estimates of the duration and frequency of crime commission. *J. Quant. Criminol.* 5: 3–32.

Riccio, L. J. (1974). Direct deterrence: An analysis of the effectiveness of police patrol and other crime prevention technologies. *J. Crim. Just.* 2(3): 207–217.

Riccio, L. J., and Finkelstein, R. (1985). Using police arrest data to estimate the number of burglars operating in a suburban county. *J. Crim. Just.* 13: 65–73.

Rice, S. (1984). *The Criminal Justice System Model: Magistrates Courts Sub-model*, Home Office Research and Planning Unit Paper 24, Home Office, London.

Rich, T., and Barnett, A. (1985). Model-based U.S. prison population projections. *Public Admin. Rev.* 45: 780–789.

Rolph, J. E., Chaiken, J. M., and Houchens, R. L. (1981). *Methods for Estimating Crime Rates of Individuals*, Report No. R-2730-NIJ, Rand Corp., Santa Monica, CA.

Rossmo, D. K. (1990). Fugitive migration patterns. Paper presented at the Annual Meeting of the American Society of Criminology, Baltimore, MD.

Rossmo, D. K., and Routledge, R. (1990). Estimating the size of criminal populations. *J. Quant. Criminol.* 6: 293–314.

Rydell, C. P., and Everingham, S. S. (1994). *Controlling Cocaine: Supply Versus Demand Programs*, Report MR-331-ONDCP/A/DPRC, Rand Corp., Santa Monica, CA.

Safire, W. (1995). Of nutsiness and madness. *The New York Times Magazine*, Section 6, Aug. 13.

Saks, M. J., and Baron, C. H. (1980). *The Use/Nonuse/Misuse of Applied Social Science Research in the Courts*, Abt Books, Cambridge, MA.

Schmidt, P., and Witte, A. D. (1988). *Predicting Recidivism Using Survival Methods*, Springer-Verlag, New York.

Science and Technology Task Force (1967). *Task Force Report: Science and Technology*, U.S. Government Printing Office, Washington, DC.

Shaw, C., and McKay, H. (1969). *Juvenile Delinquency and Urban Areas*, University of Chicago Press, Chicago.

Shinnar, R., and Shinnar, S. (1975). The effects of the criminal justice system on the control of crime: A quantitative approach. *Law Society Rev.* 9(4): 581–611.

Shover, N., and Thompson, C. Y. (1992). Age, differential crime expectations, and crime desistance. *Criminology* 30: 89–104.

Smith, T. S. (1976). Inverse distance variations for the flow of crime in urban areas. *Social Forces* 54: 804–815.

Stigler, S. M. (1986). *The History of Statistics: The Measurment of Uncertainty Before 1900*, Belknap Press of Harvard University Press, Cambridge, MA.

Stollmack, S. (1973). Predicting inmate population from arrest, court disposition, and recidivism rates. *J. Res. Crime Delinq.* 10: 141–162.

Stollmack, S., and Harris, C. M. (1974). Failure-rate analysis applied to recidivism data. *Operat. Res.* 22: 1192–1205.

Swersey, A. J. (1994). The deployment of police, fire and emergency medical units. In Pollock, S. M., Barnett, A., and Rothkopf, M. H. (eds.), *Operations Research and the Public Sector* (Volume 6 of the Handbooks in Operations Research and Management Science, edited by G. L. Nemhauser and A. H. G. Rinnooy Kan), Elsevier North-Holland, Amsterdam, pp. 151–200.

Tarling, R. (1986). Statistical applications in criminology. *Statistician* 35: 369–388.

Tierney, L. (1983). A selection artifact in delinquency data revisited. *Operat. Res.* 31(5): 852–865.

Tittle, C. R. (1988). Two empirical regularities (maybe) in search of an explanation: Commentary on the age/crime debate. *Criminology* 26: 75–85.

Turnbull, B. W. (1976). The empirical distribution function with arbitrarily grouped, censored and truncated data. *J. Roy. Stat. Soc. Ser. B* 38(3): 290–295.

Visher, C. A. (1986). The Rand inmate survey: A renalaysis. In Blumstein, A., *et al.* (eds.), *Criminal Careers and Career Criminals, Vol. II*, National Academy Press, Washington, DC, pp. 161–211.

Visher, C. A., and Linster, R. L. (1990). A survival model of pretrial failure. *J. Quant. Criminol.* 6: 153–184.

von Hirsch, A. (1985). *Past or Future Crimes: Deservedness and Dangerousness in the Sentencing of Criminals*, Rutgers University Press, New Brunswick, NJ.

Willmer, M. A. P. (1970). *Crime and Information Theory*, Edinburgh University Press, Edinburgh, Scotland.

Wolfgang, M. E., Figlio, R. M., and Sellin, T. (1972). *Delinquency in a Birth Cohort*, University of Chicago Press, Chicago.

Zimring, F. K., and Hawkins, G. J. (1973). *Deterrence: The Legal Threat in Crime Control*, University of Chicago Press, Chicago.