PUNISHMENT AND CRIME:
A CRITIQUE OF CURRENT FINDINGS
CONCERNING THE PREVENTIVE
EFFECTS OF PUNISHMENT*

PHILIP J. COOK†

INTRODUCTION

Criminal justice policy, whether formulated by a police chief, state legislature, or the United States Supreme Court, is heavily influenced by policymakers' opinions concerning the degree to which crime is responsive to criminal sanctions. Until recently these opinions were uninformed by scientific research findings on the relationship between sanctions and crime rates, simply because little systematic empirical research had been performed. During the last three decades, however, this situation has changed rather dramatically; there have been literally hundreds of empirical studies on the effects of criminal corrections programs on recidivism rates; more recently, there have been scores of empirical studies dealing with the deterrence and incapacitation effects of punishment. The evaluation literature on correctional programs has begun to play a major role in the evolution of correctional policy, and the recent research on deterrence is receiving considerable attention; indeed, several sophisticated empirical studies of the deterrent effect of capital punishment were submitted in briefs to the United States Supreme Court during its recent hearings on this issue.²

My purpose in this review is to explain what questions have been addressed in the recent research on the preventive effects of criminal sanctions, indicate what techniques are being used to answer them, and give my interpretation of what we have learned and can hope to learn through these techniques. Beyond just summarizing current results, then, I hope to communicate a sense of how much confidence should be placed in these results. The first section presents a brief overview of the literature on the influence of

---

* This research was supported in part by a grant from the United States Department of Justice. Edward D. Jones III, Daniel Nagin, James Q. Wilson, and Ann Witte offered a number of useful suggestions on an earlier draft.
† Assistant Professor of Policy Sciences and Economics, Duke University, and Economic Consultant to the Office of Policy and Planning, United States Department of Justice.
1. See Plattner, The Rehabilitation of Punishment, 44 PUB. INTEREST 104 (Summer 1976).
criminal sanctions on the number of crimes committed by convicts. Incapacitation, rehabilitation, and "dehabilitation" effects are discussed in the context of a supply and demand model of crime. The second section presents a critique of theoretical perspectives on the general preventive consequences of criminal sanctions, with emphasis on the recent "economic" theories of criminal behavior and their relationship to more traditional formulations. Sections three and four evaluate the two major sources of evidence on the deterrence effect of criminal sanctions—correlational analysis and the analysis of "natural" experiments—and summarize the findings of these studies. The final section presents my conclusions concerning what we now know about the preventive effects of punishment and the methods we have for gaining further information.

I

THE SPECIAL EFFECTS OF PUNISHMENT

The preventive effects of punishment can be classified as either "general" or "special." The special effect of a punishment refers to its effect on the amount of crime committed by the particular convict who is punished; the general effect (discussed in subsequent sections) refers to its effect on everyone else. A special effect could be achieved through any of several causal mechanisms. A prison term, for example, will incapacitate the prisoner by preventing him from committing crimes against the free community for the duration of the term (though it seems that crime rates within prisons are high). Following release, he may be less prone to commit crimes because of some rehabilitation or reformation process which occurred within prison, or because his personal experience with prison has persuaded him that crime does not pay (a mechanism known as "special deterrence"). It has also been argued that he could be more crime prone following release due to a reduction in his legitimate opportunities (stigmatization, depreciation of job skills), the acquisition of criminal norms ("prisonization"), or an enhancement of his crime-related skills. Other types of punishments could also produce some of these effects.

Of these possible effects, incapacitation is the least controversial. Unfortunately, very little research has been done which attempts to estimate the magnitude of the incapacitation effect.3

A great many evaluations of correctional programs designed to rehabilitate convicts have been conducted in recent years, with disappointing results. Robert Martinson summarized his now famous review of 231 correctional evaluation studies conducted in the postwar era with this statement: "With few and isolated exceptions, the rehabilitative efforts that have been reported so far have had

The quest for a correctional technique which would be effective in reducing recidivism rates was marked by an extraordinary period of massive experimentation during the 1960s, much of it conducted within the California penal system. Some techniques—intensive parole supervision, community treatment of juveniles, group therapy within prison—were evaluated through the use of true scientific experiments with subjects randomly assigned to treatment and control groups. The failure of these studies to demonstrate efficacy in reducing recidivism rates has been a major impetus in the current trend away from indeterminant sentencing and the "rehabilitative ideal."5

It is safe to conclude that correctional rehabilitation programs, taken collectively, have had a small effect on crime rates in the past, and that a number of notable programs failed completely. But as long as it can be shown that one or more existing, practical rehabilitation strategies can produce a positive effect on convicts' behavior in the community, then rehabilitation remains a viable objective for the correctional system. Ann Witte's evaluation of the North Carolina work release program, reported in this symposium, gives persuasive evidence that at least one such strategy does exist and deserves consideration by every state correctional system.6

The various hypotheses which suggest that the imposition of criminal sanctions may have the effects of increasing the post-release criminal activity of convicts have not been so thoroughly tested, but available evidence places them in doubt. A fairly straightforward prediction of the dehabilitation arguments is that longer prison sentences will be associated with a higher post-release recidivism rate, certeris paribus; both the "school of crime"7 and the prisonization8 hypotheses postulate that a dysfunctional learning process oc-

4. Martinson, What Works—Questions and Answers about Prison Reform, 35 PUB. INTEREST 22 (Spring 1974). Martinson's conclusion is based on a subjective weighing of the evidence he reviewed; others who have evaluated this evidence believe that his conclusion is overly negative. See, e.g., Palmer, Martinson Revisited, 12 J. RES. CRIME & DELINQUENCY 139 (1975); Halleck & Witte, Is Rehabilitation Dead?—CRIME AND DELINQUENCY—(forthcoming) 1977.

5. The failure of these studies to demonstrate rehabilitation has been thoroughly documented. See G. Kassenbaum, D. Ward, & D. Wilner, Prison Treatment and Parole Survival: An Empirical Assessment (1971); P. Lerman, Community Treatment and Social Control: A Critical Analysis of Juvenile Correctional Policy (1975); D. Lipton, R. Martinson, & J. Wilks, The Effectiveness of Correctional Treatment: A Survey of Treatment Evaluation Studies (1975); Ward, Evaluation Research for Corrections in Prisoners in America 184 (L. Ohlin ed. 1973); Robinson & Smith, The Effectiveness of Correctional Programs, 17 CRIME & DELINQUENCY 67-80 (1971). The impact of these findings on correctional policy is analyzed in Plattner, supra note 1, at 104-10.


8. An inmate is said to be "prisonized" to the extent that he conforms to inmate rather than staff norms. See D. Clemmer, The Prison Community 299-315 (1940). An extensive review of the sociological literature on prisonization is presented in G. Hawkins, The Prison: Policy and Practice 60 (1976). Hawkins notes that empirical tests of the prisoniza-
curs in prison, and it would presumably be true that longer prison terms produce more thorough learning.

The most thorough and persuasive study of the relationship between time served and post-release recidivism was conducted by the National Council on Crime and Delinquency (NCCD).\(^9\) The data base for this study consisted of a nationwide sample of 104,182 male felons paroled for the first time on their prison sentences sometime during the years 1965-1970. This sample was divided into subsamples defined by two dimensions: crime for which they were imprisoned, and whether or not they had prior convictions. Each subsample was broken into pentiles according to length of time served before release, and the first year parole success rate was calculated for each pentile.\(^{10}\)

For the subsamples of violent offenders—those convicted of robbery, assault, rape, manslaughter, or homicide—there is virtually no difference in success rates across pentiles: Taking all violent crime categories together, the success rate of those with no prior sentences drops from 88.8 per cent for the first pentile (least time served) to 86.7 percent for the fifth pentile (longest time served).\(^{11}\) For those with prior sentences, the drop in success rate from the first to the fifth pentile is even less—only 1.3 percent.\(^{12}\) For crimes of theft, there is a more substantial drop in success rates as time served increases; for example, the drop from the first to the fifth pentile for those with prior sentences is four percentage points (70.2 per cent to 66.3 percent).\(^{13}\)

Thus, while there does appear to be a relationship between parole success and time served, it is very weak. Furthermore, this relationship may reflect factors other than dehabilitation. For example, those who served relatively long terms for a given crime were no doubt identified as poor risks by their judges and parole boards, and were therefore systematically different than those who served relatively short terms even before beginning their prison terms.

---


10. "Success" in this case was defined to mean "continued on parole," i.e., not returned to prison or absconded. The parole success rates were adjusted by a statistical procedure for the effects of the age of the parolees, to eliminate that source of variation in success rates.

11. D. GOTTFREDSON, M. NEITHERCUTT, J. NUFFIELD, & V. O'LEARY, supra note 9, at 23.

12. Id. at 23.

13. Id. at 21.
Other studies\textsuperscript{14} have had findings similar to those of the NCCD project. It would appear that, if there is a dehabilitation effect for adult prisoners, it is weak or masked by other effects—though it is hard to imagine what they might be. Furthermore, there is strong evidence that incarceration does not influence the criminality of convicted juvenile delinquents in a measurable way: An analysis of the California Community Treatment Project data demonstrated that juveniles placed on probation immediately following conviction had at least as high a crime rate as a group of juveniles who had been released from a reformatory.\textsuperscript{15} Since, in other respects, the two groups were supposed to be identical, having been sorted between reformatory and probation by a random process, this result is strong evidence against dehabilitation for juveniles.

It has been argued that the stigma of a criminal record will reduce the legitimate opportunities available to a convict and thereby increase his propensity to commit crimes.\textsuperscript{16} This argument differs from the dehabilitation arguments in that it does not necessarily imply any relationship between time served and post-release success—the conviction itself confers the stigma, and it is doubtful that the length of the prison sentence would make much difference. The evidence for the “stigma effect” is indirect (indeed, it is hard to imagine how a direct empirical test could be conducted), consisting of interviews with employers and other data which suggest that many jobs are closed to ex-convicts, either de facto or de jure.\textsuperscript{17} As I have argued elsewhere,\textsuperscript{18} however, the job categories which exclude ex-convicts tend to be jobs with a relatively high salary and degree of responsibility—jobs for which most ex-convicts would not be eligible even if they lacked a criminal record, simply because most of them lack the education and skill requirements for such jobs. Parolees are successful for the most part in finding jobs; although these jobs tend to be menial and poorly paid, they do not appear to be substantially inferior to jobs which nonconvicts with similar education, skill, and experience are able to obtain. While this type of argument does not settle the issue of whether the stigma effect may be important, it does cast doubt on its importance.

The extensive literature, summarized above,\textsuperscript{19} on whether incarceration

\textsuperscript{15} Lerman, \textit{supra} note 5, at 58-67.
\textsuperscript{17} Id.; R. Taggart III, \textit{The Prison of Unemployment: Manpower Programs for Offenders} (1972); Friedman & Pappas, \textit{The Training and Employment of Offenders} in \textit{President’s Commission on Law Enforcement and the Administration of Justice, Report}, 1967.
\textsuperscript{18} Cook, \textit{The Correctional Carrot: Better Jobs for Parolees}, \textit{1 Pol’y Analysis} 11 (1975).
\textsuperscript{19} My summary is not comprehensive, and slights, among others, the extensive body of so-
and other criminal sanctions increase or reduce the amount of crime committed by convicts, has for the most part failed to pose the larger question: If a criminal's career is modified by a spell of incarceration (through some combination of incapacitation and rehabilitation effects, say), does this change produce a corresponding change in the overall crime rate? Ernest van den Haag has recently pointed out that, even abstracting from general deterrence and other general preventive effects, we cannot be sure that a change in an individual convict's behavior is reflected in an equal net change in the total amount of crime. His argument is that the amount of some types of crime may be limited by the number of profitable opportunities to commit the crime, rather than by the number of people who are prone to commit the crime. "The crime rate is fairly independent of the incapacitation of actual offenders as long as an unlimited number of potential offenders is willing to replace them as soon as net benefits suffice—when the supply of offenders is elastic."21

His argument is clearly valid for crimes which involve the production and sale of illicit commodities. The sudden incapacitation or rehabilitation of 20 per cent of the prostitutes, numbers runners, and illicit drug dealers in New York City may cause a temporary disruption in these activities. But we would expect that eventually they would be almost entirely replaced and/or that the remaining people in these occupations would step up their level of activity to make up the deficit because this is the normal supply response to the initial increase in price that would result from the withdrawal of some suppliers. This reasoning has motivated a recommendation that law enforcement efforts to reduce heroin use be redirected to focus on the demand side rather than the supply hierarchy. "[T]he key element in the heroin market will not be the poppy grower, the heroin smuggler, or the drug dealer. There are any number of alternative ways to perform each of these functions. The indispensable element is the heroin user."22

This argument can be applied to any type of crime where the supply and quality of criminal opportunities is responsive to the amount (supply) of crime. For crimes of theft, for example, the quality of opportunities available is presumably a function of the effort that potential targets devote to protecting their property, and this amount of effort will in turn depend on the risk of victimization. To the extent that businesses and households buy locks and

---

21. Id. at 55. Though Van den Haag does not mention it, his point applies almost as well to rehabilitation, dehabilitation, and stigmatization as it does to incapacitation.
alarms, guard their property, keep valuables in the bank, minimize the amount of cash they have on hand, and so on, the quality of criminal opportunities will be reduced, because more effort and skill will be required to steal a given amount of property. Casual observation suggests that the care exercised by potential victims is positively related to the risk of victimization; the one empirical study of which I am aware offers support for this notion. The quality of opportunities for the theft of merchandise, as opposed to cash, also adjusts to the quantity of theft through the forces of supply and demand in the black market.

The complete argument is illustrated by Figure I for the case of theft. Abstracted from deterrence and other general preventive effects of punish-

![Figure I](image)

23. The amount which businesses and households invest in self-protection is quite possibly as large as the amount spent in public protection activities—police, courts, and corrections. One estimate placed a lower bound on private expenditures in 1970 at $3.3 billion (two-thirds of the public outlay), but many of the minor expenses and inconveniences of defensive living were omitted from this estimate. See Bartel, An Analysis of Firm Demand for Protection Against Crime, 4 J. LEGAL STUD. 443-78 (1975).

24. Id. Bartel found that the probability that a business employed a guard increased with the victimization rate of similar businesses in the same geographic area.
ment, the equilibrium level of theft in a particular jurisdiction is determined by the intersection of two schedules: (1) the theft supply schedule $S$ which relates the theft rate to the average wage $w$ of time spent in theft activity. The amount of theft is assumed to increase with $w$, other things being equal; (2) the theft opportunity schedule $D$ which relates the average wage $w$ to the theft victimization rate. As the probability of being victimized increases, potential victims are assumed to reduce the amount of cash and other valuables they keep on hand and to purchase locks, dogs, safes and other means of protecting what they do keep—thus reducing the average reward to theft activities.

If 20 per cent of thieves are suddenly incapacitated, the theft supply schedule shifts from $S$ to $.8S$, causing (I assume) an initial 20 per cent reduction in the theft rate. However, the resulting reduction in the probability of being victimized causes potential victims to devote less effort to defending their property, which in turn stimulates theft activity among those thieves who are not incapacitated. The new equilibrium will involve a reduction of less than 20 per cent in the theft rate. Manipulation of the $S$ and $D$ curves demonstrates that if either $D$ is vertical (theft opportunities unresponsive to the victimization rate) or $S$ is horizontal (thieves are unresponsive to the average reward to theft activity), the replacement effect will be zero—the incapacitation effect will be fully reflected in the crime rate. Conversely, incapacitation will have no effect on the theft rate if $S$ is vertical or $D$ horizontal.

One caveat should be noted. It is possible that individual self-protection measures will improve rather than reduce the quality of crime opportunities. If, for example, fear of mugging keeps most people off the street at night, those people who still choose to walk are not protected by the nearby presence of fellow citizens. In this circumstance, robbery may be a more attractive crime than if most people did not choose to protect themselves.

What can be concluded from this discussion of the direct effects of punishment on convicts? First, currently available evidence suggests that rehabilitation, prisonization, stigmatization, and other such hypothesized effects of the criminal corrections system appear to be weak in practice, though the hope remains that moderately effective rehabilitation methods do exist and could be implemented on a wider scale. Second, while it can plausibly be argued that the total incapacitation effect may be large, there are no good estimates of its magnitude currently available. Third, incapacitation, rehabilitation, and other special effects of punishment will not be fully reflected in the crime rate if a replacement process similar to a market adjustment process is operative. While it may be true (as Van den Haag argues) that aberrant crimes committed by psychopaths can be reduced on a one-for-one basis by incapacitation, other types of crime which are motivated by a nondeviant

25. Witte, supra note 6.
quest for pecuniary enrichment may not be reduced substantially in the long run by incapacitation alone.

II

General Preventive Effects

The above discussion of the direct effects of criminal sanctions on the amount of crime committed by those who are subjected to such sanctions serves to clear the decks for an analysis of the general preventive effects of criminal sanctions: those "restraining influences emanating from the criminal law and the legal machinery" which operate through the threat of punishment rather than through its direct effects on those convicts who are subjected to it. Hypothetical general preventive mechanisms include the Bentham-Beccaria notion of simple deterrence, in addition to more subtle mechanisms which have a psychological or sociological basis. Zimring and Hawkins assert, for example, that the threat of punishment can be a teacher of right and wrong, a stimulant to habitual law-abiding conduct, a mechanism for building respect for the law, and a rationale for conformity.

Analysis of these general preventive mechanisms by legal scholars and social scientists forms one evidentiary basis for criminal justice policy. Ideally, policymakers would want to be informed not only of the qualitative effects of changes in the threat level (will an increase in the average severity of sanction for drunk driving reduce the incidence of this crime?), but also the magnitude of these effects. Increasing the certainty and severity of punishment typically imposes increased costs on the public fisc and, of course, on violators. These costs must be justified by a belief that the benefits (in the form of reduced crime rates) are sufficiently large that the change is worthwhile.

Furthermore, a complete evidentiary basis for policy would provide information on how changes in the threat of punishment affect different groups in the population (juveniles versus adults, poor versus middle class, those who already have a criminal record versus those who don't), and on the relative effectiveness of threats in restraining different types of crime. Other information requirements include the lags in the impact of changes in threat level, and the degree to which the process by which legal sanctions are imposed has influence on their effectiveness—as would be salient, for example, if the strength of the preventive effects were influenced by the perceived legitimacy.

28. Thus the appropriate policy question in the literature on the deterrent effect of capital punishment is not "Does capital punishment deter homicide?" but rather (perhaps) "Does capital punishment deter homicide by enough to justify incurring the chance of errors or injustice in its imposition?"
of the process or the ability of the process to communicate an immediate sense of threat to potential criminals.

It should be clear to anyone familiar with the literature that these evidentiary requirements of informed criminal justice policymaking are far from being met by the current state of the art. Until quite recently, the discussion of general preventive mechanisms has been based largely on a priori reasoning, casual observation, and indirect evidence from psychology and the social sciences. Mainly during the last decade, this type of analysis has been supplemented by systematic empirical studies of the relationship between threat levels and crime rates. This recent literature is reviewed in the next two sections, but my conclusion can be stated now: There is strong evidence from some of these studies that an increase in the threat of punishment can reduce the amount of some crimes in some circumstances, but very little evidence about the long-term effects, the typical magnitudes involved, the relevance of process, the responsiveness of juveniles to threats, or the extent to which some important crime categories (such as murder) are responsive to changes in the threat level. Furthermore, the intrinsic difficulties of this empirical work made it highly unlikely that anything like a complete scientific basis for criminal justice policy will be produced in the foreseeable future. Under these circumstances, a priori arguments and indirect evidence on human nature and criminal behavior remain an important source of information. A case in point is the recent majority opinion in *Gregg v. Georgia*, which noted the inconclusiveness of statistical studies of the deterrent effect of the death penalty, but then went on to assert:

> We may . . . assume safely that there are murderers, such as those who act in passion, for whom the threat of death has little or no deterrent effect. But for many others, the death penalty undoubtedly is a significant deterrent. There are carefully contemplated murders, such as murder for hire, where the possible penalty of death may well enter into the cold calculus that precedes the decision to act. And there are some categories of murder, such as murder by a life prisoner, where other sanctions may not be adequate.

The remainder of this section will present a brief characterization of deterrence theory as developed in recent years by economists. The subsequent section will consider other general preventive mechanisms and contrast the implications of such mechanisms with the implications of the economic theory. The brief characterization of the economic theory presented here gives the main ideas without reproducing the highly mathematical method of argumentation.

---

It is no surprise that this recent development in the economics literature harks back to Bentham: "[T]he profit of the crime is the force which urges a man to delinquency: the pain of the punishment is the force employed to restrain him from it. If the first of these forces be the greater the crime will be committed; if the second, the crime will not be committed."\(^3\) The threat of punishment is in effect a government-imposed tax on criminal activity; the higher this tax, the fewer the criminal opportunities which will be deemed worthwhile by potential criminals.

There are of course several complications that have to be taken into account. First, the "tax" is imposed with some probability less than one, and its magnitude (the severity of sentence) is also highly uncertain in practice. Thus the decision whether to commit a crime is analogous to a decision whether to gamble, play the stock market, and so on; it depends on the individual's perception of the risks and the degree to which he is averse to taking risks. Second, information on the true magnitude of the "tax" (i.e., the actual probability and severity of punishment for a particular crime) is communicated in a very haphazard fashion. Perceptions of this threat may differ widely among potential criminals, and may be only loosely related to the true threat level. It may be important in some circumstances, then, to maintain a distinction between perceived and actual threat levels in developing the implications of the theory (though this has not been done).\(^3\) Third, the severity of a given punishment (e.g., one year in prison) will be measured differently by people with differing tastes and circumstances; the person with a high standard of living and a good reputation has much more to lose from incurring this sentence than does a reprobate. Fourth, for crimes which require a substantial amount of time to plan and execute, the opportunity cost of a man's time (legitimate wage rate) will influence his perception of profitability.

For these four reasons, then, a given criminal opportunity will look more "profitable" to some than to others even if we abstract from interpersonal differences in moral outlook. But for all rational potential criminals, it will be true that any of the following changes should have a deterrent effect: an increase in the perceived probability or severity of punishment, an increase in the effort required to commit a crime, a reduction in the payoff to crime, or

---

\(^3\) 7 J. Bentham Works 399 (1843), quoted in F. Zimring & G. Hawkins, supra note 27 at 75.

\(^3\) One possibly important mechanism by which an individual's perceptions of the threat level may be modified is by actually being subjected to a criminal sanction. Actually experiencing a prison sentence may increase the criminal's perception of the severity of this punishment, and persuade him that "crime does not pay." It is my impression that it is this possibility of a punishment-induced change in perception of the legal threat to crime that distinguishes the term "special deterrence" from reformation or rehabilitation. It is difficult to distinguish between these mechanisms empirically.
an improvement in legitimate economic opportunities.\textsuperscript{33}

These qualitative implications follow from a theoretical analysis of rational behavior under uncertainty which has also been applied to a number of other economic choices; the approach assumes, then, that in its essential aspects the crime choice is not unique, and that there is no need to presuppose a distinctive set of motivations in analyzing criminal behavior. Even "deviant" crimes of sex and violence and noninstrumental crimes (illicit drug use) are readily analyzed within this framework, though the notion of how to define the "payoff" to such crimes becomes a bit murky. The implications of the theory with respect to deterrence are not changed, but are of course restricted to that segment of the population which derives enjoyment from such activities.\textsuperscript{34}

One interesting application of the theory is to the old question of whether changes in the certainty of punishment can be expected to have a greater effect than changes in the severity of punishment. First, the question has been restated more precisely: Is the elasticity of the crime rate with respect to a change in the probability of punishment greater than the elasticity with respect to a change in the severity of punishment?\textsuperscript{35} For example, would a 20 per cent increase in the probability of being imprisoned for committing a burglary (an increase from, say, a probability of .05 to .06) cause a greater reduction in the burglary rate than a 20 per cent increase in the average length of the prison term (an increase, say, from ten months to twelve months)? The answer which follows from the theory in one important circumstance is clearly no. If the typical severity of punishment for a particular crime is insufficient to make that crime unprofitable, and the punishment were to be applied with certainty, then changes in the actual probability will not influence the amount of that crime. Increases in severity, on the other hand, may have an effect. This point is salient to those large urban jurisdictions that are increasingly following the practice of giving probation to convicted robbers and burglars who lack a felony record ("the first one is free").

In cases where severity is sufficiently great, on the other hand, the theoretical answer conforms with the conventional wisdom in this matter. If potential criminals "discount" the future, then the threat of a two-year prison sentence with probability .10 will be a less effective deterrent than the threat of a one-
year sentence with probability .20; the elasticity of crime with respect to probability will be less than the elasticity with respect to severity.  

The economists' work on the general deterrence mechanism has been useful in developing the implications of a long neglected notion in criminology literature—that criminals can be viewed as rational decision makers intent on furthering their personal welfare in an environment which provides crime opportunities coupled with sanction threats. While this approach has proved productive, its focus on rational decision making neglects socialization processes which may also be important in determining crime rates. The next section argues that the economic model in some cases yields different predictions than the "socialization model," suggesting that the economists' focus on general deterrence may be excessively narrow.

A. Other General Preventive Mechanisms

Other general preventive mechanisms have been discussed at length in a number of recent publications, and there is no need to restate the arguments here. It is interesting, however, to note the potential conflicts between the predictions which follow from consideration of these mechanisms and the apparent implications of the simple deterrence formulation.

1. Lags in the Effect

While the simple deterrence theory presumes that changes in the threat level will have their full impact on crime rates the moment these changes are perceived by potential criminals, effects which operate through other mechanisms may not have their full impact for some years. It has often been asserted, for example, that law-abiding behavior is habitual for many people and that force of habit precludes them from taking advantage of profitable criminal opportunities even if they are virtually risk-free. If an effective criminal justice system is one stimulant to individuals initially acquiring this habit, then the possibility arises that an era of mild and infrequently applied criminal sanctions may produce a whole generation of people who fail to see any

36. Id. at 489. The authors argue, using the same assumptions, that increasing the variability of sentences reduces the deterrent effect of a given average length of sentence—an argument for determinate sentencing.


38. It should be noted that additional lags may occur if the supply of criminal opportunities is slow to respond to initial changes in the crime rate caused by the change in the threat level. For example, homeowners may respond only slowly to a reduction in the burglary rate. Eventually they would be expected to reduce their expenditures on self-protection, thus increasing the quality of burglary opportunities, but it is doubtful that this response would occur as quickly as the initial deterrence effect. If these suppositions are correct, then the initial drop in the burglary rate due to the increase in sanction threat will be followed by a partial remission as burglars find that opportunities have become more attractive.
rational reason to acquire the habit of law-abiding behavior. The ramifications of this failure could be reflected in the crime rate for years, even if a subsequent increase in the effectiveness of the system were achieved. Similarly, if criminal law and the effectiveness with which it is enforced shape the framework for the moral education of youth, then a reduction in the effectiveness of law enforcement may contribute to the production of an amoral, crime-prone generation: "It would be difficult to teach honesty, nonviolence, and similar positive values in a society where these rules were openly and commonly broken without punishment." In both the mechanisms of habit-formation and moral education, then, an initial failure in deterrence produces long-term consequences for the crime rate.

One implication of these hypothesized lags in the general preventive effect of law enforcement is that the full magnitude of a change in the threat level cannot be assessed by observing its apparent short-term influence on the appropriate crime rate. Indeed, the virtual impossibility of measuring directly the full, long-run effect suggests once again the importance of a priori reasoning and indirect evidence. The notion that "crime breeds crime"—especially if the crime is uncompensated by punishment—is a powerful one and should be given more serious attention by scholars.

2. The Importance of Severity in Punishment

The legislated (and actual) severity of penalty for a particular offense may influence the public's feeling for the seriousness or moral repugnance of this offense. The famous quote from Stephen expresses the point well:

Some men, probably, abstain from murder because they fear that, if they committed murder, they would be hung. Hundreds of thousands abstain from it because they regard it with horror. One great reason why they regard it with horror is, that murderers are hung with the hearty approbation of all reasonable men.

Murderers are no longer hung; in fact, the median time served in prison by those convicted of homicide before their release on parole is less than five years. Andenaes notes that "in the long run, such a reduction in penalty might also reduce the inhibitions against committing murder in situations

39. J. ANDENAES, supra note 26, at 124. An interesting bit of evidence is that Reiss and Rhodes, in their study of juvenile delinquency in Nashville, found that the delinquency rate of upper class boys was heavily influenced by the delinquency rate of the neighborhood in which they lived. Reiss & Rhodes, The Distribution of Juvenile Delinquency in the Social Class Structure, 26 AM. SOC. REV. 720 (1961).


41. D. GOTTFREDSON, M. NEITHERCUTT, J. NUDDIELD, & V. O'LEARY, supra note 9, at 17. It was recently reported that twenty per cent of those convicted of homicide in New York City were freed on probation. An additional fifty-nine per cent of those convicted of homicides received prison sentences that would make them eligible for parole in three years or less. N.Y. Times, January 27, 1975, at 1, col. 3.
where murder seems a tempting escape from a situation of emotional conflict.\(^4\)

The issue for our purposes is whether considering the moralizing influence of a severe penalty might change the conclusions concerning the relative magnitudes of the responsiveness of crime to changes in the probability and severity of punishment which follow from an analysis of simple deterrence. In the long run, for those crimes in which “moral inhibition” plays an important role, maintaining high severity may be more important than a high probability of punishment.

3. Other Characteristics of the Law Enforcement Process

I have argued that the moralizing influence of law enforcement may influence the public’s attitudes towards illegal activities and hence the propensity to engage in such activity. A related but logically distinct mechanism by which law enforcement activities may inhibit crime is through the influence of these activities on the public’s respect for the law as such. If the process by which laws are made and enforced is generally viewed as legitimate and just, then many will obey the law simply out of a sense of obligation to honor its authority. The economist’s term for this is “honesty preference,” which will “bias” people’s decision in favor of law-abiding behavior.\(^4\) Behaviorally, this preference would mean that an individual would abstain from illegal behavior unless such behavior was perceived as sufficiently profitable (net of the threat of punishment and other considerations) to overcome the individual’s preference for doing things legally.

The importance of honesty preference in inhibiting criminality is well expressed by Andenaes:\(^4\)

A certain degree of respect for the formal law is probably essential for the smooth functioning of society. Where it is lacking, law enforcement agencies play a role similar to that of an occupying army in foreign territory, a comparison which has been made often enough with regard to the law enforcement agencies in the ghettos of American cities. Experience shows that an occupying army may be able to create a high degree of order and compliance if it is prepared to use sufficiently harsh terror methods. But such a reign of terror is fundamentally different from the system we are used to in democratic societies.

The arguable importance of respect for the law adds a new dimension to the relationship between law enforcement efforts and the degree of com-

\(^{42}\) J. ANDENAES, supra note 6, at 59. The “reduction in penalty” which he is hypothesizing in this paragraph is a reduction to “three or four years imprisonment.”

\(^{43}\) Block & Heineke, supra note 30. One would suppose that the stigma of being convicted of a crime in a community which generally had respect for the law would be greater than in a community where such respect was lacking, and that community-wide respect for the law would thereby increase the deterrence value of criminal sanctions.

\(^{44}\) J. ANDENAES, supra note 26, at 116.
pliance with the law: Respect is influenced by the manner in which the law is enforced. If enforcement efforts are viewed as corrupt, arbitrary or (worse) biased, unresponsive to the needs of the community, and liable to error, then disrespect for the law and the prohibitions which the system enforces is likely to be a natural consequence.

The implication of this argument is that the general preventive effect of law enforcement efforts are not transmitted through the probability and severity of punishment alone, but also by the public’s perception of the quality of the process which produces these threats.

B. Summary

The general preventive effects of criminal sanctions may operate through several mechanisms. The simple deterrence mechanism has been formulated in terms of a rigorous model of economic choice under uncertainty which draws on and derives its credibility from the close analogy between criminal behavior and various sorts of legitimate economic behavior that are subject to uncertain payoffs. The economists have failed to note that the predictions which follow from analysis of this mechanism should not be viewed as identical to the general preventive effects of the criminal sanction. I have pointed out that some of the implications of other preventive mechanisms are not entirely consistent with the implications of a simple deterrence model.

The arguments concerning the qualitative effects of these mechanisms are persuasive to the extent that they are consistent with one’s beliefs concerning human nature, the socialization process, attitude formation, and so on. This type of a priori reasoning would ideally be tested through careful empirical research which would also provide a basis for estimating the magnitude of the predicted effects. In practice, as will be explained below, the inherent difficulties of making valid estimates of such magnitudes suggest that criminal justice decisionmakers will be forced to continue to rely mainly on “plausibility” arguments concerning these magnitudes when formulating policy. A modest but

45. The simple deterrence effect of sanctions is also weakened to the extent that potential criminals believe that the criminal justice system is prone to punish innocent people. In an extreme case, if the probability of being punished were perceived as unrelated to guilt or innocence, the law would have no deterrent effect.

46. For evidence, see Friedland, Thibaut, & Walker, Some Determinants of the Violation of Rules, 5 J. APPLIED SOC. PSYCH. 103 (1973). The authors conducted a laboratory experiment to measure the effect on rule violations of the probability and severity of sanctions and the “fairness” of the surveillance methods. Surveillance methods which were perceived as unfair generated a higher rate of rule violations than fair surveillance methods, ceteris paribus.

47. Note the error in Harris, On the Economics of Law and Order, 78 J. POL. ECON. 165 (1970). He assumes that legal safeguards for criminal suspects will increase the supply of criminal offenses, other things (including the probability and severity of punishment) being equal. The argument should be that such safeguards, if they are viewed as enhancing the legitimacy of the process, will reduce crime if they do not reduce the probability and severity of punishment.
important role for empirical research would be to limit the range of assertions concerning the preventive effects of punishment which can be viewed as plausible by policymakers.

III

EMPirical STUDIES OF DETERRENCE I: CORRELATIONAL STUDIES

The preferred method for testing hypotheses and measuring the magnitude of causal relationships is to conduct random experiments in which the experimental treatment is administered to randomly selected subsets of the sample of experimental subjects and the remainder of the subjects serve as controls. As mentioned above, such experiments have been conducted to test the efficacy of various criminal rehabilitation methods. However, there are virtually insurmountable practical problems in conducting random experiments to measure the general preventive effects of punishment outside the laboratory. The “subjects” in such an experiment would be jurisdictions—precincts, counties, or states, for example—and the relevant criminal justice authorities would have to submit themselves to experimental control. Political, ethical, and legal problems appear to preclude such an experimental design in most cases. A rare exception to this rule is the Kansas City Preventive Patrol Experiment, which is discussed in the next section.

The principal alternatives to random experimentation are (1) studying “quasi-experiments”—usually a sudden and dramatic change in the sanction threat level, as in the case of a police strike or a major change in the law, and (2) correlational studies which analyze the relationship (if any) between threat levels and crime rates across jurisdictions or over time. Both of these methods suffer from the same basic problem. If relationships between crime rates and threat levels are found in the data, they are not necessarily the result of the general preventive causal process. Other processes (which at least in principle would be controlled in a true random experiment) are capable of generating such relationships, and it is usually difficult to make a persuasive case that some general preventive effect is responsible, in whole or in part, for patterns in the data. While this problem is especially severe for correlational studies, they do have certain advantages over studies of quasi-experiments, which account for their recent proliferation in the economics and sociology literature.

49. See note 5 supra.
50. Some interesting laboratory experiments have been conducted with human subjects to test various hypotheses relating to deterrence. See Friedland, Thibaut, & Walker, supra note 46, and the references therein. One problem with laboratory experiments of course is that it is not clear under what circumstances the results of such experiments are applicable to real life situations. See the discussion of “external validity, D. Campbell & J. Stanley, supra note 48, at 5.
I begin this review of recent empirical work with a discussion of the correlational studies, reserving my discussion of other studies for the next section.\footnote{For an excellent, comprehensive review of the findings, methods, and technical problems with these studies see Nagin, General Deterrence: A Review of the Empirical Evidence, Management Sci. (forthcoming 1977).}

The main objectives of the correlational studies have been to test the implications of the simple deterrence model and to estimate the magnitude of the deterrence effect with respect to both the probability and severity of punishment for a number of crimes. These studies typically are not concerned with the technology of producing threats, and hence are not directly relevant to policy. Knowing that an increase in the probability of imprisonment for robbery will reduce the robbery rate is helpful background information to a policymaker, but it carries no direct implication about the potential effects of varying the available criminal justice policy instruments; after all, the probability that offenders will be arrested and jailed is not a variable which can be directly manipulated by any criminal justice agency. How much can this probability of imprisonment for robbery be increased by expanding or redeploying the police force, soliciting citizen cooperation with the police, expanding the prosecutor's staff or instituting a mandatory sentencing law? The correlational studies are not designed to answer such specific questions, but are focused instead on measuring the potential efficacy of a general approach to crime control. By abstracting from the details of the technology of threat production, it is hoped that the underlying principle of general deterrence can be illuminated.\footnote{The correlational studies made the implicit assumption that the methods by which a jurisdiction produces arrests, convictions, and sentences are irrelevant; that what matters is the resulting probability and severity of punishment which results from whatever methods are adopted. This assumption is questionable, to say the least.}

A. The Approach

The correlational studies of crime are intended to answer the following question: If the probability and/or severity of punishment in some jurisdiction (state, county) for some specified crime or group of crimes were changed, what would be the direction and magnitude of the resulting change in the rate of commission of that crime in that jurisdiction? The focus of interest here is the partial effect of a change in the threat level when other factors which influence the crime rate are "controlled for" or "held constant." For example, the result may be in the form: "Other things equal, a five percent reduction in the Manhattan burglary rate will result from a 10 percent increase in the probability that a burglary will result in a prison term." If Manhattan actually did engineer such a change in this probability during the year 1977, the actual burglary rate would change by an amount which is determined by many factors (changes in economic conditions, demographic com-
position of the population, and so on,) in addition to the increase in threat level to burglars. But if the study is correct, then the actual burglary rate would be approximately five per cent lower than it would have been in the absence of the increase in the threat level.

Virtually all the correlational studies have been motivated by an interest in measuring the magnitude of the general deterrence effect. This focus has dictated both the general statistical approach and the interpretation of the results. In fact, however, it is usually not possible to measure a pure deterrence effect using this approach. If a negative (partial) relationship between sanction threat levels and crime rates is observed over time or across a number of jurisdictions at a single point in time, then the relationship may be due to deterrence. But, especially if the typical mode of punishment is incarceration, the relationship may also be due to incapacitation, rehabilitation, or some combination of effects.

It should be noted that this problem does not make the results of such studies any less interesting—estimates of the total preventive effect of punishment are at least as policy-relevant as estimates of the deterrence effect alone. But the interpretation of such results is certainly not a matter of indifference, either for scientific or policy purposes. In what follows, I will discuss the correlational studies as if they were studies of deterrence alone, since that is the approach and interpretation used by their authors. But it should be kept in mind that such an interpretation is misleading in most cases.

A recent survey of this literature explains the basic approach:

General deterrence is inherently an aggregate phenomenon since it is reflected in the behavior of the entire population. Consequently, all analyses use aggregate data on crime-commission rates, and examine the association of commission rates with various sanctions measures. These are primarily: (1) two measures of probability of apprehension (clearance rate and the ratio of arrests to reported offenses). (2) a measure of probability of imprisonment (the ratio of prison commitments to reported crimes) and (3) a measure of severity of punishment (mean or median time served).

It has been shown in a variety of contexts that there is a negative association between the rates of commission of the Index crimes and the probability of arrest or incarceration; this result obtains for both cross section studies (comparing states, counties within a state, or precincts within a city) and time series studies. These studies also show that the simple association between the severity of sentence and amount of crime is sometimes negative, but not consistently so.

53. For a detailed discussion of the issues discussed in this paragraph, see J. Gibbs, supra note 19.
54. Nagin, supra note 51.
55. Id. Early correlational studies based on comparing crime rates with the probability and average of incarceration across states include: Bean & Cushing, Criminal Homicide, Punishment, and Deterrence: Methodological and Substantive Reconsiderations, 52 Soc. Sci. Q. 277 (1971); Gibbs, Crime,
These results are interesting but by no means conclusive evidence on the deterrence hypothesis. Many factors influence the crime rate in a particular jurisdiction and year, and these must be controlled for in some fashion if the deterrence effect is to be partialled out and measured. A number of deterrence studies have attempted to control for other factors.

B. Controlling for Other Factors

The truism that observed crime rates are influenced by many factors outside the purview of the criminal justice system—socioeconomic and demographic characteristics of the jurisdiction's population, the propensity of victims to report crimes, and so on—creates two problems with the study of the deterrence mechanism. First, it obscures the underlying relationship between threat levels and crime rates by creating a great deal of "noise" in the data. This problem of "noisy" data is readily handled through elementary statistical techniques, and poses no real conceptual difficulty. It has the effect, however, of reducing the precision of estimates of the magnitude of the deterrence effect.

A second and much more serious difficulty introduced by the influence of socioeconomic variables and other factors is the possibility that some of these factors will distort the relationship between crime and criminal sanctions; i.e., that the systematic portion of the observed relationship will reflect the influence of other factors. For example, states differ with respect to the fraction of their populations which are in the crime prone ages of fifteen to seventeen; states with a relatively large youthful population will thereby tend to have a relatively high crime rate. It is also true that juveniles are rarely sent to prison; hence the fraction of a state's crimes which result in imprisonment will tend to be small in a state with a disproportionately large amount of youth crime. This "youth" variable could thus produce a negative interstate relationship between crime and imprisonment probability, even in the absence of a significant deterrence effect. And this type of distortion could be created by many other variables as well. Clearly this is a serious problem if the deterrence effect is to be isolated and measured.56

---

56. See D. CAMPBELL & J. STANLEY, supra note 48, at 64 for a discussion of the relationship between correlation and causation. A fascinating collection of examples from the medical literature and other sources which illustrates the dangers of inferring causation from correlation is Gilbert, Light, & Mosteller, Assessing Social Innovations: An Empirical Base for Policy, in BENEFIT-COST AND POLICY ANALYSIS, 1974, 3 (R. Zeckhauser, A. Harberger, R. Haveman, L. Lynn, W. Niskanen, A. Williams eds. 1975). Black concludes that the problem of controlling for other variables in assessing the deterrent effect of capital punishment from correlational studies of state data is insurmountable: "A 'scientific'—that is to say, a soundly based—conclusion is simply impossible, and no methodological path out of this tangle suggests itself." C. BLACK, CAPITAL PUNISHMENT: THE INEVITABILITY OF CAPRICE AND MISTAKE 25 (1974).
One conceptually simple solution to this problem is to compare crime rates between jurisdictions which are similar in all important respects except the probability and/or severity of punishment. Sellin attempted this approach in his classic study of the deterrent effect of capital punishment.\(^57\) In one study he selected six clusters of contiguous states, each cluster including at least one retentionist and one abolitionist state. He argued that these states, because they were contiguous, were “more nearly alike” with respect to “the nature of their populations and their economic, social, and political conditions” than would be geographically disparate states.\(^58\) For these groupings there is no clear tendency for the abolitionist states to have higher homicide rates than the retentionist states for the forty-three year period observed by Sellin, leading him to conclude that the death penalty “in law or practice—does not influence homicide death rates.”\(^59\)

Sellin’s conclusion is not persuasive, because even if each execution had the effect of preventing eight murders (as Isaac Ehrlich has estimated),\(^60\) the overall effect on the homicide rate of Sellin’s retentionist states would be so small as to be quite possibly swamped by the effects of other differences between the states in each comparison cluster. The reason is simply that the retentionist states had very low execution rates, ranging from New Hampshire, with one execution in the years 1920-1955, to Ohio, with an average of seven executions a year.\(^61\)

The second solution to the control problem in correlational studies provides the approach which has been used in all the econometric analyses of deterrence: multivariate regression analysis.\(^62\) This approach requires the development of a mathematical model that specifies the list of variables which are thought to be important in explaining variations in crime rates, and that specifies a linear equation which is hypothesized to characterize the form of the causal relationship between the explanatory variables and the dependent variable (the crime rate). The data is then used to estimate the parameters of this equation (i.e., the coefficients of the explanatory variables); the estimation procedure yields parameter values which give the closest possible “fit” between the equation and the observed values of the dependent variable.

---

58. Id. at 135.
59. Id. at 138.
62. For a thorough explanation of this technique and the problems involved in applying it to analyzing the deterrent effect of capital punishment, see L. Friedman, The Use of Multiple Regression Analysis to Test for a Deterrent Effect on Capital Punishment: Prospects and Problems (Oct. 1975) (on file at Graduate School of Public Policy, University of California, Berkeley).
The economic theory of deterrence suggests that the list of explanatory variables should include measures of the quality of legitimate and illegitimate opportunities in addition to variables which characterize the probability and severity of punishment. Other variables, reflecting cultural differences among jurisdictions, should be included in a complete specification of a crime equation. However, the number and kinds of explanatory variables which can be included in an econometric specification are limited by the available data. They are also limited by the statistical requirement that reasonably precise estimates of parameter values can be obtained only when the number of observations is substantially greater than the number of explanatory variables included in the equation.

One of the better examples of the use of multivariate regression analysis to study crime is Issac Ehrlich's analysis of interstate differences in Index crime rates for 1960. His specification of the mathematical relationship between crime rates and explanatory variables includes the following explanatory variables for each type of crime: (1) $P$, the number of offenders imprisoned per offenses known; (2) $T$, the average time served by offenders in state prisons; (3) the median income of families (supposedly a measure of the quality of illicit opportunities); (4) the percentage of families below one-half the median income (a measure of the quality of legitimate opportunities); and (5) the percentage of non-whites in the population. Using a log linear form of the regression equation, he finds that there is a negative partial relationship across states between the crime rate for each type of crime and the probability $P$ and severity $T$ of punishment for that crime. (See Table I) Estimated elasticities with respect to $P$ range from -.9 (robbery) to -.1 (homicide). These estimates are of course imprecise; however, Ehrlich uses a standard statistical technique to demonstrate that one can be very confident that the "true" elasticities with respect to $P$ are negative to some degree (i.e., that the deterrence effect is "working"). The results with respect to the severity are not so clear, except for the crimes of burglary and larceny—his results do not rule out the possibility that changes in severity have no relationship to the rate of commission of the other Index crimes.

The attractiveness of this regression approach is that it appears to generate estimates of the magnitude of the deterrence effect, however imprecise, in a context which controls for the possible distorting effects of other causal factors. The obvious questions are whether Ehrlich and other scholars who have used regression analysis have actually used appropriate and accurate measures of the factors they think are important, and whether they have ac-

---

63. Ehrlich, supra note 30.

64. Once again, it should be noted that to label these estimates as estimates of a deterrent effect is not strictly correct, since logically they must be influenced by incapacitation effects. Ehrlich recognizes this fact and includes a brief discussion of its implications. Id. at 535.
**Table I**

**Estimated Elasticities of Index Crimes With Respect to Changes in the Probability and Severity of Punishment (Based on a Multivariate Regression Using 1960 Data on States)**

<table>
<thead>
<tr>
<th>Crime</th>
<th>Crime Elasticity with respect to . . .</th>
<th>P</th>
<th>T</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homicide</td>
<td>-.3</td>
<td>-.1*</td>
<td></td>
</tr>
<tr>
<td>Rape</td>
<td>-.6</td>
<td>-.2*</td>
<td></td>
</tr>
<tr>
<td>Assault</td>
<td>-.3</td>
<td>-.2*</td>
<td></td>
</tr>
<tr>
<td>Robbery</td>
<td>-.9</td>
<td>-.2*</td>
<td></td>
</tr>
<tr>
<td>Burglary</td>
<td>-.5</td>
<td>-.9</td>
<td></td>
</tr>
<tr>
<td>Larceny</td>
<td>-.1</td>
<td>-.3</td>
<td></td>
</tr>
<tr>
<td>Auto Theft</td>
<td>-.2</td>
<td>-.2*</td>
<td></td>
</tr>
</tbody>
</table>

* The standard error of estimate is more than one-half the estimated coefficient, which indicates that the probability that the true parameter is zero exceeds five per cent.


Initially succeeded in controlling for all the important factors which may distort the deterrence relationship. One does not have to be very sophisticated to find fault with Ehrlich's specification, or to suggest other factors which should be controlled for. It is certainly suggestive that published multivariate regression analyses based on a variety of data sets have for the most part found a negative partial relationship between crime rates and the probability of arrest or incarceration; however, it is not by any means conclusive evidence of a deterrence effect.

65. There are two issues here. First, the variables which Ehrlich selects to represent the factors which he thinks are important are inadequate. There is no measure of the sanction threat to statutory juveniles, for instance, even though juveniles commit about forty-five per cent of the Index crimes. His measure of illegitimate opportunities (i.e., the median income of the state) seems like a very weak proxy indeed, especially for crimes of violence. Second, there are no doubt many other factors which could be important and which are not controlled for. Differences in cultural norms among states may be important, for example, and are especially salient to the extent that such differences reflect the historical behavior and effectiveness of the state's criminal justice system. Bayley argues that Japan's low crime rate and high clearance rate are both the result of a culture which stresses compliance to authority and (hence) cooperation with police. See Learning about Crime—the Japanese Experience, 44 PUB. INTEREST 55 (Summer 1976).


67. This doubt is reinforced by the supposition that it is much easier for a scholar to get a study published in an economics journal if the empirical findings seem to support the theoretical prediction, grounded in economic theory, that the threat of punishment deters crime. This bias among editors may stimulate a certain amount of "fishing" in the data among scholars. Klein expresses this point well: "To an economist the question of whether punishment deters seems curious. We are taught from a very young age to draw demand curves sloping downward and supply curves sloping upward. And the presumption is if empirical work does not pick up these price effects, there are most likely some deficiencies in the work." Klein, *Comment*, in The
Further sources of doubt about the validity of correlational studies of deterrence are perhaps even more troublesome than the "control" problem. First, there is a possibility that the empirically observed negative relationships between crime rates and threat levels are a statistical artifact resulting from errors in the measurement of crime rates. Second, the interpretation of these negative correlations, even if they are "real," is in some doubt: There is considerable evidence for the hypothesis that crime rates influence threat levels, and that this causal process (the converse of the deterrence process) may account for all or part of the observed correlations.

C. Data Problems

The poor quality and inadequacy of published estimates of crime rates have been discussed and documented so often that there is no need to repeat the sad story here.68 The significance of this data problem for deterrence studies is twofold: first, that published categories of crime data often do not correspond to magnitudes which would be desirable for the study of deterrence; second, that the errors in measurement which certainly reduce the precision with which relationships can be estimated may also produce artificial relationships between crime and other variables of interest.

The first problem is clearly evident in studies that attempt to test the deterrent effect of capital punishment. There are no published data on the number of capital offenses committed in a jurisdiction, and all studies of which I am aware have been forced to utilize the "murder and non-negligent manslaughter" data collected by the FBI, or some similar comprehensive statistic on murder. Only a fraction of these murders are capital crimes in retentionist districts, and there is no reason to believe that this fraction is constant across jurisdictions or over time.69 Similar problems arise in the context of deterrence studies of other crimes and sanctions, the most serious of which is that data on crime rates do not permit juvenile crimes to be distinguished. Since juveniles presumably respond differently to threats and opportunities than adults, and certainly are treated differently by the criminal law, it is impossible to test the deterrent effects of sanctions without taking into account the differences between adults and juveniles.

---


69. Ehrlich, supra note 30, at 406, assumes that the number of capital murders $C$ is related to the total number of murders $M$ by the relationship

$$ C = 8e^{-2t} + 2M, $$

where $t$ represents chronological time in years and $\mu$ denotes "random errors of reporting or identifying murders."
justice system than are adults, it requires some heroic assumptions indeed to warrant lumping juvenile and adult offenses into a single explanatory equation.

The second problem, that errors in the measurement of crime rates can generate artificial negative correlations between measured crime rates and measures of the sanction threat, results from the types of variables used to measure the probability of punishment. The denominators of such ratios as the arrest rate, incarceration rate, or clearance rate are the measured crime rates, implying that this explanatory variable will be perturbed by the same errors in measurement as the dependent variable (the crime rate), but in the opposite direction. An error which exaggerates the true crime rate will result in an under-estimate of the true incarceration rate. An error which yields an underestimate of the true crime rate will yield an overestimate of the true incarceration rate. If a data set which is being used to study deterrence includes some jurisdictions whose official crime statistics exaggerate the true crime rate, and others which understate the true rate, these errors will tend to generate a negative correlation between crime and the probability of incarceration which may be erroneously interpreted as evidence of a deterrent effect.

The argument must be modified slightly to account for the fact that measured crime rates almost always understate the true rate, simply because published measures are based on “crimes reported to the police.” A substantial fraction of serious crimes (excluding only homicide and perhaps auto theft) are never reported to the police and hence cannot be included in official tabulations. If every jurisdiction’s official measure of crime understated the true rate by the same fraction, deterrence studies could safely ignore the data problem (this type of proportional error will not affect the correlation between the two variables). But the problem of an artificial “deterrence” relationship is present so long as some jurisdictions report a larger fraction of their true crime rates than others; those which report a relatively high (low) proportion of crimes will tend to have a relatively low (high) computed incarceration rate.

Evidence that the ratio of the true crime rate to the reported crime rate differs widely among jurisdictions can be deduced from comparing the crime rates calculated from some recent victimization surveys of major American cities with the official crime rates for these cities. This comparison is made

---

70. See generally U.S. DEP'T OF JUSTICE, FEDERAL BUREAU OF INVESTIGATION, UNIFORM CRIME REPORT (published annually).

in Table II for the crime of burglary. While it is subject to certain kinds of error, the victimization survey data is almost certainly a more consistent and accurate measure of the true burglary rates in these cities.\textsuperscript{72} The considerable

<table>
<thead>
<tr>
<th>City</th>
<th>(1) Burglary Rate UCR</th>
<th>(2) Burglary Rate NCP</th>
<th>(3) UCR/NCP</th>
<th>(4) Clearance Rate UCR</th>
<th>(5) Clearance Rate NCP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chicago</td>
<td>15</td>
<td>68</td>
<td>.22</td>
<td>.29</td>
<td>.065</td>
</tr>
<tr>
<td>Detroit</td>
<td>41</td>
<td>106</td>
<td>.39</td>
<td>.08</td>
<td>.031</td>
</tr>
<tr>
<td>Los Angeles</td>
<td>34</td>
<td>94</td>
<td>.37</td>
<td>.20</td>
<td>.073</td>
</tr>
<tr>
<td>New York</td>
<td>26</td>
<td>71</td>
<td>.37</td>
<td>.19</td>
<td>.071</td>
</tr>
<tr>
<td>Philadelphia</td>
<td>15</td>
<td>73</td>
<td>.21</td>
<td>.28</td>
<td>.058</td>
</tr>
<tr>
<td>Atlanta</td>
<td>43</td>
<td>119</td>
<td>.36</td>
<td>.24</td>
<td>.085</td>
</tr>
<tr>
<td>Baltimore</td>
<td>26</td>
<td>81</td>
<td>.32</td>
<td>.28</td>
<td>.090</td>
</tr>
<tr>
<td>Cleveland</td>
<td>20</td>
<td>78</td>
<td>.26</td>
<td>.11</td>
<td>.029</td>
</tr>
<tr>
<td>Dallas</td>
<td>35</td>
<td>94</td>
<td>.37</td>
<td>.24</td>
<td>.084</td>
</tr>
<tr>
<td>Denver</td>
<td>41</td>
<td>104</td>
<td>.40</td>
<td>.25</td>
<td>.101</td>
</tr>
<tr>
<td>Newark</td>
<td>47</td>
<td>107</td>
<td>.44</td>
<td>.15</td>
<td>.065</td>
</tr>
<tr>
<td>Portland</td>
<td>37</td>
<td>100</td>
<td>.37</td>
<td>.12</td>
<td>.045</td>
</tr>
<tr>
<td>St. Louis</td>
<td>42</td>
<td>89</td>
<td>.47</td>
<td>.32</td>
<td>.150</td>
</tr>
<tr>
<td>Boston</td>
<td>31</td>
<td>109</td>
<td>.29</td>
<td>.13</td>
<td>.037</td>
</tr>
<tr>
<td>Buffalo</td>
<td>18</td>
<td>64</td>
<td>.28</td>
<td>.32</td>
<td>.087</td>
</tr>
<tr>
<td>Cincinnati</td>
<td>32</td>
<td>118</td>
<td>.27</td>
<td>.41</td>
<td>.112</td>
</tr>
<tr>
<td>Houston</td>
<td>31</td>
<td>96</td>
<td>.32</td>
<td>.12</td>
<td>.038</td>
</tr>
<tr>
<td>Miami</td>
<td>36</td>
<td>68</td>
<td>.53</td>
<td>.10</td>
<td>.054</td>
</tr>
<tr>
<td>Milwaukee</td>
<td>11</td>
<td>83</td>
<td>.13</td>
<td>.44</td>
<td>.056</td>
</tr>
<tr>
<td>Minneapolis</td>
<td>34</td>
<td>112</td>
<td>.30</td>
<td>.12</td>
<td>.035</td>
</tr>
<tr>
<td>New Orleans</td>
<td>22</td>
<td>72</td>
<td>.30</td>
<td>.20</td>
<td>.060</td>
</tr>
<tr>
<td>Oakland</td>
<td>57</td>
<td>142</td>
<td>.40</td>
<td>.10</td>
<td>.039</td>
</tr>
<tr>
<td>Pittsburgh</td>
<td>18</td>
<td>57</td>
<td>.32</td>
<td>.17</td>
<td>.053</td>
</tr>
<tr>
<td>San Diego</td>
<td>21</td>
<td>79</td>
<td>.27</td>
<td>.41</td>
<td>.111</td>
</tr>
<tr>
<td>San Francisco</td>
<td>28</td>
<td>84</td>
<td>.33</td>
<td>.15</td>
<td>.049</td>
</tr>
<tr>
<td>Washington, D.C.</td>
<td>23</td>
<td>55</td>
<td>.42</td>
<td>.20</td>
<td>.083</td>
</tr>
</tbody>
</table>

\textit{Definitions and Sources:}

Columns 1 and 2 are burglary rates per 100,000 residents aged twelve and over for various years: 1972 for the first five cities, September 1971—August 1972 for the next eight cities, and 1973 for the final thirteen cities. The number of burglaries was taken from various issues of the \textit{Uniform Crime Reports} (for column 1) and various Law Enforcement Assistance Administration reports on the National Crime Panel Surveys (for column 2). Central city population estimates were taken from the NCP reports.

Clearance figures are unpublished tabulations made by the FBI. Column 4 is the ratio of burglary clearances to burglaries reported in the UCR; column 5 is the ratio of the burglary clearances to burglaries reported by the NCP.

Column 3 is the ratio of column 1 to column 2 (or, equivalently, the ratio of column 5 to column 4), and represents the fraction of all burglaries which were reported to the police and recorded by them.

\textsuperscript{72} See the discussion in Skogan, \textit{Measurement Problems in Official and Survey Crime Rates}, 3 J.
differences in the ratios (column 3) are thus primarily the result of differing propensities of victims in the twenty-six cities to report burglaries to the police, coupled perhaps with differing degrees of efficiency and accuracy in the record-keeping activities of the police departments in these cities.

The possibility that intercity variability in burglary reporting rates could produce a spurious negative correlation between burglary rates and sanction threat levels was checked using the clearance rate data given in columns 4 and 5. The number of clearances for burglary in each city was available from unpublished tabulations prepared by the FBI.\textsuperscript{73} Clearance rates were calcu-

\textsuperscript{73} CRIM. JUST. 17 (1975). In intercity comparisons to test theories of crime causation, what is needed is a measure of crime rate which bears some consistent (proportional) relationship to the "true" rate. The much noted "dark figure" of crime—the large fraction of crimes which are never reported to the police, or are not recorded by the police if reported—is of no concern for this type of research so long as the underreporting is uniform across cities. For example, if 30 per cent of burglaries are reported in official police crime statistics in each city under study, then the intercity pattern of burglary will be revealed just as precisely as if 100 per cent of burglaries were reported in each city. A precise measure of this pattern is all that is needed to analyze the relative importance of various criminogenic factors.

Neither the Uniform Crime Reports nor the victimization surveys are thought to measure the true crime rates in these cities. My assertion is that the crime rates estimated from the victimization surveys, at least for the crime of burglary, bear a more consistent relationship to the true crime rates than is true for the official police statistics published in the Uniform Crime Reports.

The main sources of error in the official crime statistics are that victims do not report a large fraction of crimes to the police, and that the police in some jurisdictions do not record all the crimes which are reported to them. There is good reason to think that intercity differences in the "professionalism" of police departments make police statistics highly inconsistent proxies for the true crime rate. For example, Skogan estimates that the probability that a burglary reported to the police would be recorded in the official police statistics ranged from 79 per cent (St. Louis and Miami) down to 22 per cent (Milwaukee)—see Crime and Crime Rates, in Sample Surveys of the Victims of Crime (W. Skogan, ed. 1976).

The main source of potential error in the victimization survey estimates of the amount of crime in a city are: (1) Memory errors: Respondents are asked to relate the incidents in which they were victimized during the preceding year. They may forget some incidents which did occur during the requested time frame, and report other incidents which actually occurred more than one year before the interview; furthermore, some respondents may simply be more cooperative than others. (2) Commuter problems: Crimes actually committed in the city against nonresidents will of course be omitted from the victimization surveys, which only interview residents. (3) Sampling problems: Some groups, especially young males, tend to be hard to locate and will be underrepresented in the sample. The second problem may be particularly important for crimes such as robbery and larceny, and will cause inconsistency problems across cities to the extent that cities differ in the amount of commuting, tourism, etc. But neither the second nor third problems will be salient for the crime of burglary which I am considering. The first problem—that of memory and cooperativeness of respondents—will not cause intercity inconsistency in estimated crime rates assuming that these problems tend to be fairly uniform across cities. Inconsistency problems could arise if interviewers in some cities were better trained or more persistent than in others, or if residents of some cities tended to be more cooperative than in others.

My conclusion is that at least for the crime of burglary, there is good reason to place greater faith in the victimization survey data than in police reports.

73. "Clearances" are the sum of "clearances by arrest" and "exceptional clearances." A crime is "cleared by arrest" when at least one person is arrested, charged, and turned over to the court for prosecution. A case is "exceptionally cleared" when the identity and the location of the offender are "known" to the police, but they cannot make an arrest due to the offender's death, confession of someone who is already under custody, refusal of the victim to cooperate, and so
lated using the error prone *Uniform Crime Reports* estimate of the number of burglaries as a denominator (column 5). The correlation between the UCR reported burglary rates and the clearance rates based on these figures is -.47, apparently substantial support for the deterrence notion. However, the correlation between the burglary rate and clearance rate when National Crime Panel (NCP) data are used is only -.15, which is not significantly different from zero. The "evidence" evaporates when more accurate data are used.

The problems introduced by errors in the measure of crime rates will be more or less severe depending on the degree to which the measure of crime used in a study bears an unstable relationship to the true crime rate. The problem is not severe for a crime which is nearly always reported to the police (such as homicide or auto theft), assuming that police departments' record keeping systems are reliable; the problem is much more important for crimes such as burglary, especially when different jurisdictions are being compared. The results of the simple experiment reported above should help to deter scholars from being overly confident in estimates of the deterrent effect which are based on such unreliable data.

D. The Influence of Crime Rates and Sanction Levels

A final source of doubt concerning the evidence for deterrence generated by correlational studies is that the negative partial correlations between crime and threat levels, even if they survive the "control" problem and the errors in measurement problem, could be the result of some causal process which is unrelated to the deterrence mechanism. Two mechanisms have been suggested by which changes in the crime rate could cause changes in the opposite direction in the probability and severity of punishment, thus producing a

---

74. Let $B_{\text{UCR}}$ and $B_{\text{NCP}}$ represent the UCR and NCP burglary rates, and $C_{\text{UCR}}$ and $C_{\text{NCP}}$ represent the UCR and NCP clearance rates. The following ordinary least squares regression results were obtained from the data in Table 2:

\[
\begin{align*}
B_{\text{UCR}} &= 4.2 + 5.4 C_{\text{UCR}} \\
& (2.0) \quad R^2 = .22 \\
B_{\text{NCP}} &= 9.7 + 11.1 C_{\text{NCP}} \\
& (15.3) \quad R^2 = .02
\end{align*}
\]

The numbers in parentheses represent the estimated standard error of the coefficient estimate. It should be noted that the coefficient on the arrest or clearance probability is negative and significant when victimization survey data are used for the crime of robbery, as reported in J. Wilson and B. Boland, *Crime*, in *The Urban Predicament* (1976). (Wilson and Boland do not compare their result with what they would have gotten using UCR data.) One suspects that the victimization survey data are more error prone for robbery than for burglary (see note 72 supra), if so, then the negative correlation may be a statistical artifact.
negative correlation between crime and threat levels. The first is a production function notion. The criminal justice system has as inputs both employees and crimes, and from these inputs produces arrests, convictions, and sanctions; if the number of crimes increases without a corresponding increase in the other inputs to the system, then we would expect (due to the principle of diminishing marginal productivity) that the number of arrests and so on would increase, but by an amount which is less than proportional to the increase in crimes. Hence the increase in crime would result in a reduction in the arrest and conviction rates—at least until the jurisdiction had time to respond by increasing capacity. A clear example of this short run phenomenon is an urban riot, in which “the probability of apprehension of individual rioters, as well as of offenders committing other crimes, decreases considerably below its normal level due to the excessive load on local police units.”

In general, this “loading” mechanism would produce a negative short run relationship between crime rates and threat levels to the extent that there are unexpected changes in the crime rates.

In the long run, jurisdictions can respond to a change in the crime rate by making the appropriate change in the criminal justice system budget. There is no necessary reason in the long run why jurisdictions with large increases in crime rates should produce relatively low rates of arrests and convictions. However, casual observation would suggest that the public’s willingness to allocate resources to crime prevention is not sufficiently responsive to changes in the level of crime to maintain a constant threat level. For example, jurisdictions with the highest crime rates tend to have the least police investigatory effort per crime, the longest court backlogs, the greatest resort to plea bargaining, and so on.

One illustration of this effect is given in Figure II, which depicts an interstate comparison of Index crime rates with the number of Index crimes per police employee (“CRIMES/COP”). The curve in Figure II represents the curve which best “fits” the points plotted there. According to this functional relationship, a 10 percent increase in the Index crime rate is associated with a 5.6 per cent increase in the number of Index crimes per police employee. While it is true that relatively high crime states have a


76. Ehrlich, supra note 75, at 540.

77. One study showed that the trend in total United States expenditures for law enforcement during the 1960s was not affected by the large increases in crime rates during the 1960s. The result was a dramatic drop in clearance rates during this period. See Votey & Phillips, Police Effectiveness and the Production Function for Law Enforcement, 1 J. LEGAL STUD. 423 (1972).

78. See, for example, the descriptions of the day-to-day activities of rural and urban prosecutors in the Wall St. J., May 5, 6, 1976, at 1, col. 1. The rural prosecutor was able to devote considerable time to even very minor cases. No rigorous demonstration of these hypothesized patterns is available in the literature, however.

79. The ordinary least squares regression line is

\[
\log \frac{\text{CRIMES}}{\text{COP}} = -1.87 + .56 \log (\text{CRIME RATE}) R^2 = .63
\]
relatively large number of police per capita, Figure II demonstrates that they have a relatively small number of police per crime. This evidence that police budgets are not very responsive to changes in the crime rate appears to provide a plausible explanation for observed negative relationships between

\[
\frac{\text{CRIMES}}{\text{COP}} = +.15 \, \text{(CRIME RATE)}^{.56}
\]

80. The interstate relationship between the number of police employees per capita and the number of crimes per capita is

\[
\frac{\text{cops}}{100,000 \, \text{Pop.}} = 93.0 + .044 \, \text{(CRIME RATE)} \quad R^2 = .54
\]
crime rates and sanction threat levels, an explanation which is clearly competitive with the "deterrence" interpretation of such observed relationships.\footnote{This type of hypothetical explanation for observed "deterrence" relationships has been used to critique Ehrlich's empirical work on the deterrent effect of capital punishment. Ehrlich, supra note 61. See Hoenack, Kudrle, & Sjoquist, Some Difficulties in the Estimation of the Deterrent Effect of Capital Punishment (1977) (unpublished manuscript on file at Institute of Policy Sciences and Public Affairs, Duke University).}

How can we know whether a negative correlation between crime and sanction threat levels, assuming it is not a statistical artifact, is the result of a deterrence process, the "public choice" mechanism, the "loading" mechanism, or all three? This is a rather deep statistical question which has been discussed extensively in the literature and which has motivated an effort to identify factors which influence criminal justice system budgets in an attempt to distinguish the budget effect from the deterrence effect. To date, this effort must be judged a failure.\footnote{Nagin, supra note 51, draws this conclusion after a careful technical critique of this literature.}

E. Conclusions

The objective of the correlational studies discussed in this section is to extract estimates of the magnitude of the simple deterrence effect for each of the Index crimes from published data on crime rates, criminal justice threat levels, and other criminogenic variables. The studies have used a variety of statistical techniques, including matching, simple correlation, multivariate regression, and more sophisticated methods. Many of these studies have reported evidence that the probability and (less consistently) the severity of punishment have the expected inhibiting influence on crime rates. In some cases the estimated magnitude of the deterrence effect is quite large. These studies are of great interest to criminologists and may serve to influence the predilections of criminal justice policymakers.

There are numerous conceptual and statistical problems with these studies, some of which were discussed above. The focus on simple deterrence has diverted attention from the long-term effects of law enforcement on crime. The accuracy of the deterrence effects estimated by the techniques is questionable due to the problems of distinguishing the deterrence process from other processes which may cause threat levels to be negatively related to crime rates, the problems introduced by inadequate and inaccurate crime statistics, and the problem of controlling for other criminogenic factors which may distort the deterrence effect.

My conclusion from reviewing these studies is that they have produced little persuasive evidence concerning the size or even the existence of the deterrence mechanism. This is not to say that we can conclude that the deterrence mechanism is not important. Indeed, Sellin's negative conclusions
cerning the marginal deterrence effect of the death penalty are no more persuasive (perhaps even less so) than analyses which appear to demonstrate the existence of a deterrence effect, either in this context or in the context of other crimes and sanctions. Rather, my conclusion is that these studies have not produced reliable evidence to date.

A number of correlational studies of the deterrent effect of capital punishment were submitted to the Supreme Court in briefs in *Gregg v. Georgia* and related cases. These studies appear to have had little influence on the majority of the Court, who, as mentioned above, resorted to common sense in assessing the deterrence issue. In my view, this was an entirely appropriate response, given the problems inherent in the approach used by these studies.

IV

**Empirical Studies of Deterrence II: Evaluating “Natural Experiments”**

The principal alternative to correlational studies is the empirical study of “quasi-experiments”: sudden, dramatic changes in the law or criminal justice policy which can be presumed to change the public's perception of the certainty and/or severity of punishment for some class of offenses. Frequently cited examples include police strikes in Liverpool (1919) and Montreal (1956) and the mass arrest of the Copenhagen police force by the Nazis in 1944. The huge increase in crime rates which followed each of these events is persuasive evidence that the threat of punishment has a substantial inhibiting effect on crime. More recently, there have been a number of studies of the effects of changes in criminal justice policy, some of which are reviewed here.

The principal advantage of studying such events, when compared with the alternative of a correlational study, is that it is often easier to make a persuasive case that observed changes in the crime rate are actually the result of the induced change in the threat level, and not the result of other factors such as those that have not been controlled for, data problems, or some causal process which is unrelated to the deterrence mechanism. Each of these problems can still be troublesome, however, and this advantage is merely one of degree.

A second advantage of this approach is that it yields information about problems involving the implementation of changes in the law or criminal justice system policy. It contributes to our knowledge concerning the technology of threat production as well as the effect of threats on behavior. There are many instances in which “get tough” programs have failed to reduce crime, not because of a failure in the deterrence mechanism but simply because the programs never succeeded in increasing the objective or perceived threat of criminal sanction.

---

83. See Brief for United States as Amicus Curiae, *supra* note 2.
84. J. *Andenaes*, *supra* note 26, at 50.
The principal disadvantage of such studies is that they lack the generality of correlational studies. First, changes in justice policy which are likely to produce unequivocal results must involve large changes of a sort that would not be feasible for most jurisdictions most of the time. Small changes in policy are likely to have small effects which would be difficult to identify and measure, given the “noisiness” and inaccuracy of crime data. For example, it would be hard to imagine that the short-term effect of the Furman decision on homicide rates could be sorted out from the many other sources of variability in homicide rates; the decision invalidated existing capital punishment laws at a time when the threat of execution was already negligible. Second, major changes in policy, when they do occur, are each limited to one time and geographical area—a particular precinct, state, or country—and the nature and magnitude of the observed effects may be heavily influenced by the idiosyncrasies of that time and place. Whether the findings serve as a valid prediction of the effects of a similar policy change in a different jurisdiction must remain questionable.

The discussion below focuses on four of the best and most interesting studies of the effects of criminal justice policy changes: the Kansas City Preventive Patrol Experiment, analyses of the effects of increases in the policing of the 20th Precinct and the subway system in New York City, and finally an evaluation of the British Road Safety Act of 1967.

A. The Kansas City Preventive Patrol Experiment

This experiment was undertaken with the primary objective of testing the efficacy of police preventive patrol in deterring crime. Police themselves, the general public, and elected officials have always believed that the presence or potential presence of police officers on patrol severely inhibits criminal activity. The deterrence effect would presumably be the result of an increased probability of an “on-scene” arrest for crimes visible from the street (e.g., auto theft, certain kinds of larceny, assaults and robberies occurring out-of-doors), and an increased probability of arrest for these and other types of crime due to a reduced police response time to citizen calls for assistance.

The experiment, which ran from October 1, 1972 to September 30, 1973, involved a contiguous fifteen-beat area of Kansas City, Missouri. Five of the beats (the control group) received the usual amount of routine preventive patrol; five were given no preventive patrol, and in the remaining five beats,
preventive patrol activity was raised to a supernormal level that was supposed to increase police visibility to two to three times its usual level. A careful evaluation of crime rates in these three areas revealed the now famous finding that no significant differences in crime rates resulted from the experimental manipulation. The reaction of James Newman, assistant chief of the Kansas City Police Department, may typify a common interpretation of this finding: “We just went out and proved that the liberal cliche is correct: crime is caused by social conditions which very frequently are beyond the control of police.”

There are two plausible explanations for the failure of intensive preventive patrol to reduce crime: either (1) crime is not responsive to an increase in the threat level (apparently the interpretation of Assistant Chief Newman), or (2) the experimental increase in preventive patrol activity in fact failed to increase the level of threat. Fortunately, the experiment generated sufficient data to allow us to choose between these possible explanations. The finding that neither police response time to citizen calls nor the arrest rate differed among the three groups of beats suggests that the second explanation is the correct one.

The question of why intensive preventive patrol activities failed to produce any measurable increase in the threat level has been recently investigated. The answer, which must be viewed as tentative until it has been subjected to criticism by other experts, is that the experimental design was flawed in the sense that it did not succeed in producing large differences among the three groups of beats in visible patrol presence or in travel time necessary to respond to calls. Police cars patrolling the perimeters of the “no preventive patrol” beats, and responding to calls in those beats, apparently produced a substantial visible patrol presence which was not much different than those beats that were receiving the normal amount of preventive patrol. The value of preventive patrol was not adequately tested by this experiment; a fortiori, it produced no evidence concerning the marginal deterrence value of increasing police manning levels in an urban area. This latter issue has been investigated in two recent studies reviewed below.

89. Id. at 5.
91. G. KELLING, T. PATE, D. DIECKMAN, & C. BROWN, supra note 87 at 23, 36.
93. The authors note that it is unfortunate that some news media reports suggested the experimental results justified reducing the level of policing in a city. “Those who drew manpower reduction conclusions from the preliminary findings assumed that if the crime prevention strategies currently being used did not work, no crime prevention strategies would work.” G. KELLING, T. PATE, D. DIECKMAN, & C. BROWN, supra note 87, at 49.
B. The Effects of Increased Police Manning

A perennial problem faced by city councils is to evaluate police department requests for increased manpower. Since a primary function of the police is to reduce crime, information on the potential productivity of additional policemen in this regard is essential to making wise decisions.

The dramatic evidence of the chaos which occurs during a police strike suggests that firing the entire police force is not a practical option for policymakers, but beyond that there is little in the way of information which is relevant to deciding the appropriate size of the police force. Two recent studies, however, are informative.

1. Subway Crime

In April 1965 New York’s Mayor Wagner ordered a large increase in the size of the Transit Authority Police Department (TAPD) in order to stem the rapidly increasing subway crime rate. The objective specified by Mayor Wagner was that every subway station and every train be policed between 8:00 p.m. and 4:00 a.m. Meeting this objective required increasing the number of TAPD patrolmen on duty during these hours from two hundred to about eight hundred. This huge increase would be expected to have a large deterrent effect on nighttime subway crime, and indeed “the number of felonies reported at night reached, at their lowest, a level only one-third as high as that attained in the year preceding the manning increase. Six years later, reported rates for nearly all crime types between 8:00 p.m. and 4:00 a.m. had not returned to their 1964 and early 1965 levels.” In particular, the number of nighttime token booth robberies fell from twenty-three during the first quarter of 1965 to a total of nine for the remaining three quarters of the year.

The sudden large increase in police manning, followed immediately by a large drop in the crime rate, is persuasive evidence that the causal mechanism was indeed deterrence. This explanation is further supported by the fact that the felony arrest rate increased substantially following the manning increase. However, two competing explanations deserve mention: the regression effect and the possibility of systematic errors in the crime data.

The “regression effect” (no relation to “multivariate regression analysis”) in this case refers to the tendency of unusual deviations from the long-run trend of a variable (e.g., the subway crime rate) to be followed by a spontane-

95. Id. at 178.
96. Id. at 182.
97. The number of felony arrests remained about constant: Coupled with the substantial drop in the number of felonies, the implication is that the arrest rate increased.
ous reversion to the mean. The year before Mayor Wagner decreed the increase in police manning of the night trains, subway felony rates had increased over 50 per cent, and this apparently unusual increase was what provoked Mayor Wagner's decision. Is it possible that the drop in felony rates which followed the increase in police manning was simply a spontaneous regression effect which would have occurred in any case? One method of controlling for this possibility is to compare the trend in nighttime subway felony rates with the trend in daytime felony subway rates, since the trends in these two crime rates are presumably influenced by the same set of criminogenic factors. The trend in daytime felony rates should serve as an indication of what would have happened to nighttime rates if there has been no increase in police presence. The finding from this comparison is equivocal: During the first year following the nighttime increase in police presence, the daytime felony rate fell almost as much as the nighttime rate, even though there were no significant changes in policing of day trains. However, in subsequent years the daytime rate rebounded to its former trend line while the nighttime rate grew much more slowly. Instead of being a regression effect, the initial drop in daytime felony rates may have been a “phantom” deterrence effect: Potential subway robbers may have been temporarily deterred from daytime crimes as well as nighttime crimes by the misperception that police manning levels had been raised throughout the twenty-four hours. Subsequent increases in the daytime felony rate would suggest that this phantom effect faded as the true pattern of police manning became evident.

Data problems in this evaluation could stem from two sources. First, the propensity of victims to report crimes could have been affected by the increase in police presence. But if this change actually occurred, it would most likely lead to an inflation of the reported nighttime felony rates and a consequent underestimate of the true impact of the additional police. Furthermore, one would expect that 100 per cent of one type of subway crime—token booth robberies—are reported to the police under any level of police man-

98. Id. at 179.
99. A well known example is the finding that the drop in auto casualties which followed a crackdown on speeding on Connecticut highways was most likely due to the regression effect. The crackdown was initiated due to an unusual increase in road fatalities, and the subsequent drop in fatalities would be expected even in the absence of public intervention. See Campbell & Ross, The Connecticut Crackdown on Speeding: Time Series Data in Quasi-Experimental Analysis, 3 L. & Soc'y Rev. 33 (1968).
100. Chaiken, Lawless, & Stevenson, supra note 94, at 185.
101. During the first eight months of 1973 the reported hourly felony rate was three times as high during the daytime as it was during the nighttime police saturation period; in 1964, before the decision to introduce saturation police manning of the night trains, the felony rates were about the same over the course of the day. See N.Y. Times, October 10, 1973, at 49, col. 7.
ning, and the reported rate of token booth robberies fell even more during 1965 than the reported rates of other types of robberies. A second and more insidious measurement problem did occur, however: A TAPD official in charge of recording subway crimes allegedly changed the time when some nighttime robberies occurred in his official reports, with the apparent intent of producing an exaggerated impression of the effectiveness of the saturation manning program. The magnitude and duration of this intentional error in measurement are apparently insufficient to account for more than a fraction of the observed effect.103

2. Crime in the 20th Precinct in New York City104

“For the purpose of studying increased manpower as a deterrent force, on October 18, 1966, the police increased the number of patrolmen assigned to the 20th Precinct located on the West Side of Manhattan from about 200 to 300 Patrolmen (the average increase was actually about 40 percent), while manpower elsewhere in the City remained fairly constant.”105 A comparison of seasonally adjusted average weekly reported felony rates during the six-month period before the change with a thirteen-month period following the change revealed that “inside” felonies (those which would be invisible from the street) increased by eleven per week, while “outside” felonies decreased by eleven per week.106 Thus the apparent net effect was to change felony patterns while leaving the overall felony rate unchanged. However, the 40 per cent increase in police manning of the 20th Precinct occurred at a time when felony rates in other parts of the city were increasing. There is good reason to believe, then, that the increase in police manning did deter crime in the 20th Precinct, coincidentally by an amount just sufficient to cancel out changes in other criminogenic factors which tended to increase crime. It was estimated that if police manpower levels had not been increased, inside felonies would have been five per cent higher and outside felonies 36 per cent higher than they actually were reported to be during the thirteen months following the manpower increase.107 If these estimates are valid, it would appear that the deterrent effect was quite substantial.

There is no explanation in the report of why the New York Police De-

105. Id. at 2. This experiment was similar to Operation 25, in which the police force in a high crime district of Manhattan was doubled for four months in 1954. Some crime rates reportedly fell rather dramatically during this period. See Wilson, Crime and Law Enforcement, in Agenda for the Nation 179, 187 (K. Gordon ed. 1968).
106. S. Press, supra note 104 at 73 app. A.8.
107. Id. at 16. For each type of crime, the change in incidence in the 20th Precinct was compared with the change in incidence in several other precincts. The comparison precincts were chosen according to criteria of similarity to the 20th Precinct with respect to the crime rate for the crime in question, population, police manning levels and miles of highway.
partment chose the 20th Precinct as the site of this quasi-experiment. If it was chosen because it had been subject to unusually rapid increases in crime rates prior to the change, then there would be some danger that the estimated deterrent effect was actually the result of a regression effect. However, the validity of the deterrence interpretation is supported by the finding that indoor crime was affected much less than outdoor crime. This result would be expected if the primary effect of the increase in police presence was to increase the potential criminals' perception of the probability of being arrested for an outdoor crime. In fact, this perception would have been accurate: The arrest rate for inside felonies only increased from 5.6 per cent to 6.4 per cent while the arrest rate for outside felonies increased from 5.9 per cent to 9.6 per cent.108

Both the subway and the 20th Precinct studies offer persuasive evidence that increasing the level of police manning in a specified area can produce an increase in the felony arrest rate and a substantial inhibiting effect on serious crime. While these results are certainly interesting to criminologists, it should be noted that they do not provide all the information needed to judge whether the police manning increases were worthwhile. At least some of the criminals who were deterred by the extra policemen from committing offenses on the subway or in the 20th Precinct may simply have been displaced to other targets in other areas of the city. This possibility is entirely consistent with deterrence theory, and seems highly plausible. Both studies attempted to identify and measure displacement effects. Crime rates in areas contiguous to the 20th Precinct were analyzed for evidence of an unusual increase, and a relatively small effect was actually found in the case of one of these (the Central Park Precinct). Displaced subway crime is virtually impossible to detect because it would necessarily be a very small fraction of above-ground crime. The subway crime evaluation did, however, detect an interesting interaction between subway and bus robberies.109

C. The British Road Safety Act of 1967110

The British Road Safety Act of 1967 (BRSA) enacted a precise definition of drunken driving—driving with blood alcohol content in excess of .08 per cent—in order to overcome the considerable difficulties in obtaining drunken driving convictions under the earlier, vaguer law. The Act defined a wide range of circumstances under which policemen could stop a car for the purpose of administering a breathalyzer test to the driver. If the driver “failed” or refused this test, he could be arrested, given a more precise laboratory test,

108. Calculated from the statistics given in Id. at 73 app. A.8.
109. Chaiken, Lawless, & Stevenson, supra note 94, at 188.
and ultimately be subjected to a mandatory one year’s suspension of his driver’s license (in addition to other possible penalties). Following passage of the BRSA the government spent £ 350,000 in a massive publicity campaign to explain the new law to the public. This publicity effort, coupled with the many news reports generated by the passage and implementation of the law, apparently succeeded in persuading the British public that under the new Act there was a high probability that a drunk driver would be arrested and deprived of his license. Laurence Ross’s evaluation of the effects of the BRSA provides the strongest evidence I know that a moderate change in governmental policy can, under the appropriate circumstances, produce an effective deterrent to illegal activities.

While the incidence of drunken driving cannot be measured directly, the close causal link between drunkenness and road accidents suggests that the auto accident casualty rate, which is measured with a high degree of accuracy, is a good proxy for the incidence of this offense. The casualty rate is in any event an appropriate statistic with which to evaluate the effect of the Act, since its ultimate purpose was to reduce road casualties. The initial success of the BRSA in this regard is strikingly indicated by the fact that the road fatality rate for October 1967 (the first month the Act was in force) was 25 per cent below the September rate.\footnote{Id. at 32.} That this drop (and the relatively low fatality rate which obtained thereafter) was the result of the deterrence effect of the law, rather than the influence of some other factor, is supported by the following considerations. (1) The magnitude and suddenness of the reduction precludes the possibility that it was a random fluctuation.\footnote{Id. at 30.} (2) As one would expect, all of the reduction in casualties occurred during weekends. The casualty rate on weekdays, when drunk driving is rare, did not change noticeably.\footnote{Id. at 33.} (3) The road fatality rate fell more than the rate of lesser casualties, as one would expect from previous studies which showed that accidents where inebriation is a factor are more serious on the average than other accidents.\footnote{Id. at 32.} (4) Surveys taken before and after the BRSA was enacted showed an increase in the number of people who reported walking, rather than driving, to their local pub.\footnote{Id. at 64.} (5) A smaller fraction of drivers killed in auto accidents after October 1967 had high blood alcohol levels than had been true before this date.\footnote{Id. at 64.}

Unfortunately, the British public’s perception that the implementation of the BRSA would result in drunk drivers being arrested and convicted with a
high degree of certainty proved to be false. The British police were not aggressive about enforcing the law. Breathalyzer tests were administered sparingly, usually when there was evidence from his driving that a driver was inebriated. The Act did facilitate gaining convictions for drunk driving rather than plea bargained lesser charges, such as careless or reckless driving, and the average severity of punishment for those who were arrested was thereby increased. But "it appears that the Road Safety Act, contrary to popular impression, did not greatly increase the likelihood of police action in the event of drinking and driving."  

The perhaps inevitable result was that public perceptions of the legal threat to drunk driving gradually came into conformity with reality, and by 1970 the deterrent effect of the BRSA was no longer evident in British road fatality statistics. The deterrent effect was the result of an illusion produced by the initial publicity, and this illusion could not be supported long given the refusal of the police to make good the threat. In any event, the initial deterrent effect of the law has been irrefutably demonstrated, and serves as a cornerstone of the empirical deterrence literature.

CONCLUSION

The focus of this review has been on evaluating what we now know about the effects of the criminal justice system activities on the crime rate, and how we came to know it. A related subject, which has been neglected here as well as in the literature, is the effect of private protection activities on the crime rate, and the important set of issues concerning the interactions between the private and public efforts to reduce crime. The ultimate goal is to develop a complete model which incorporates both demand and supply aspects of the production of crime.

The empirical literature on the preventive effects of punishment is highly uneven. Rehabilitation and other "specific" effects of punishment lend themselves to systematic experimentation in a way that general preventive mechanisms do not, and implications of the rehabilitation and rehabilitation studies are correspondingly stronger and more reliable. The correlational studies of

117. Id. at 48.

118. Ross based his conclusions concerning the long-run effects of the BRSA on a comparison of the actual highway death rate in 1970 with an extrapolation of the pre-BRSA trend line. This procedure is not very reliable, and Ross's conclusions in this one respect should be viewed accordingly.

119. These questions are one aspect of a subdiscipline which goes under the "victimology" rubric. For recent contributions, see Bartel, supra note 23; Chaiken, Lawless, & Stevenson, supra note 94 (for an analysis of the effects of efforts to protect subway token booths and buses against robbery); A. Schneider, Evaluation of the Portland Neighborhood-Based Anti-Burglary Program (1975) (unpublished paper on file at Institute of Policy Sciences and Public Affairs, Duke University); Komesar, A Theoretical and Empirical Study of the Victims of Crime, 2 J. LEGAL STUD. 301 (1973).
the simple deterrence mechanism, while in some cases sophisticated and carefully done, have, in my judgment, contributed very little to our fund of knowledge concerning this important mechanism. What “facts” we do have about deterrence come from the study of natural or quasi-experiments.

The conclusions? First, most of the rehabilitation programs which have been studied carefully appear to have failed, but Ann Witte's article in this issue suggests that providing prisoners with a steady job and some income may be effective in reducing recidivism.120 In any event, there is little support for the notion that imprisonment interferes with rehabilitation. Second, the evidence on the effectiveness of the simple deterrence mechanism clearly precludes the flat claim that deterrence does not work. It has “worked,” sometimes with a strikingly large effect, in a number of circumstances. The evidence is very spotty, however, and we are far short of a set of reliable quantitative estimates of the responsiveness of various kinds of crime to changes in the threat level. In many cases, the short-term reduction in crime resulting from increasing the threat level may not be sufficient to justify the costs of the policy change. Third, there is almost no scientific knowledge concerning the long-term effects of punishment on the amount of crime, and yet the long-term effects (enculturation, habit formation, respect for the law) may be the most important.

There is no threat that criminology will obtain the quantitative (or even qualitative) precision of the hard sciences, although some recent research has proved valuable to policymakers and further effort may prove productive. For the foreseeable future, careful a priori reasoning, descriptive evidence on human nature and criminogenic processes, and common sense will rightfully remain the principal sources of evidence in the debate over criminal justice policy.

120. Witte, supra note 6.