WORKING PAPER
ALFRED P. SLOAN SCHOOL OF MANAGEMENT

Crime and Capital Punishment:
Some Recent Studies

Arnold Barnett

WP 921-77 April 1977

MASSACHUSETTS
INSTITUTE OF TECHNOLOGY
50 MEMORIAL DRIVE
CAMBRIDGE, MASSACHUSETTS 02139
Crime and Capital Punishment: Some Recent Studies

Arnold Barnett

WP 921-77 April 1977
Crime and Capital Punishment:
Some Recent Studies

Abstract

Some of the recent papers on the deterrent effect of the death penalty are discussed, and the need for further work suggested.

Arnold Barnett
Sloan School of Management
Massachusetts Institute of Technology

In November, 1976, the Commonwealth of Massachusetts, arguably the most liberal state in the nation, put a gun-control proposition on its ballot. The proposal was defeated by what *Newsweek* called "a stunning 3-1 margin." Presumably a question on the same ballot about restoring the death penalty would not have fared so badly, for a capital punishment bill had sailed through the Massachusetts Legislature, which failed later by only one vote to override the Governor's veto. In December, 1976, New Jersey became the 35th state since 1972 to pass new laws that provided capital sanctions in certain homicide cases.

The events in Massachusetts and the rest of the nation make clear that, rightly or wrongly, the restoration of capital punishment has become the linchpin of the American people's response to the doubling of murder rates in the last decade. Gun control measures, attempts to reduce media violence, and more general social and economic reforms have been accorded a much lower priority in the national consensus. It seems probable that, first erratically, then more regularly, executions will take place in America throughout the foreseeable future.

It seems fair to say that America's historical experience with the death penalty -- which includes thousands of executions earlier this century -- has had minimal net impact on the evolving national decision to restore capital punishment. The Supreme Court, for example, in its 1976 decision authorizing the resumption of executions, stressed that it had reached no conclusions on whether past executions had served to deter murders. It would be one thing if this situation arose because those who had studied past murder and execution data found them inherently inconclusive on deterrence questions. But on the contrary, most researchers seem to find in such data strong implications on
the subject; the problem is intense disagreement on what these implications are. The effect has been the creation in the public mind of what Supreme Court Chief Justice Warren E. Burger called "an empirical stalemate," in which opposing studies "cancel" each other out and conjectures become the bases of public views.

This spectacle has doubtless amused those who believe "you can prove anything with statistics," and that statisticians should be ignored if not jailed. But others will find in the situation short-term and longer-term reasons to lament. The recent Supreme Court decision authorized the resumption of executions only under certain conditions; most states must apparently revise their laws to comply with the Court's ruling. This means that once again, state legislatures will be considering the death penalty issue; as things stand now, they will do with little effective guidance provided by the past. Looking ahead, it seems inevitable that, ten or twenty years hence, attempts will be made to evaluate the effects of death penalty statutes passed now. What reason is there to believe that the statistical confusion about past data will not reoccur for future data? Will it perpetually be the case that judgements on this issue will be based less on science than on "hunches"?

In this paper, we discuss in detail two of the most prominent recent studies on the deterrent value of capital sentences. One of these studies suggested that executions may act strongly to deter murders; the other discerned no deterrent effects. We argue that neither study offers any strong reasons for confidence in its conclusions. Then we discuss some considerations relevant to a reanalysis of data now available, a reanalysis that might save such data from the de facto irrelevance accorded to it by recent controversies.
The Work of Ehrlich

Before 1970 there were apparently no extensive statistical studies on the deterrent effect of capital punishment, though some well-publicized rudimentary work was done by Sellin [1]. In a 1967 book, he compared the murder rates over several decades in clusters of adjacent American states which had different penalties for convicted murderers; the graphs he prepared led him to conclude that executions exert no influence on murder levels. As Ehrlich [2] and others have pointed out, Sellin effectively assumed that no differences (e.g. demographic, economic) other than capital-offense statutes affect the relative murder rates in neighboring states. Furthermore, he considered only the legal status of the executions in a given state, which missed the point that many states that had death penalty laws on their books almost never used them. Also, Sellin's comparisons over time and space were not made with great precision. Because of such limitations, it would probably be inappropriate to give too much weight to Sellin's conclusions; his work does establish, however, that the effects of executions are not so gargantuan as to make formal analysis of the data superfluous.

In the early 1970's Isaac Ehrlich [2] performed a study on capital punishment that has been quoted and debated extensively. He recognized that questions about the impact of the death penalty could only be answered in the context of a comprehensive theory to explain the evolution over time of homicide rates. (In this paper, we mean by homicide "murder and nonnegligent manslaughter" as defined by the FBI, and use the terms murder and homicide interchangeably.) Ehrlich identified several variables that he believed possibly could influence such rates:

\[ P(a) = \text{Probability of arrest for a randomly-chosen homicide} \]
\[ p(c|a) = \text{Probability of conviction for homicide given arrest for homicide.} \]
\[ p(e|c) = \text{Probability of execution given conviction for homicide} \]

\[ L = \text{Fraction of civilian population in the labor force} \]

\[ U = \text{Unemployment rate} \]

\[ A = \text{Fraction of residential population in age group 14-24} \]

\[ Y_p = \text{Friedman's estimate of (real) permanent income per capita in dollars} \]

\[ NW = \text{Fraction of nonwhites in the civilian population} \]

\[ x_p = \text{Per capita (real) expenditures on police} \]

\[ x_g = \text{Per capita (real) expenditures (excluding national defense) at all levels of government.} \]

Of the three variables above that directly concern the punishment for murder (i.e. \( p(a), p(c|a), p(e|c) \)), only the last relates to capital punishment; the point being that one should attempt to separate the deterrent effect of executions (if any) from the deterrent effects of arrest, conviction and incarceration. Ehrlich hypothesized that the mathematic form of the relationship between murder rates and their determinants was that of the Cobb-Douglass econometric model:

\[ M = C (p(a))^{\alpha_1} (p(c|a))^{\alpha_2} (p(e|c))^{\alpha_3} U^{\beta_1} L^{\beta_2} \cdots \varepsilon \quad (1) \]

where \( M \) is the murder rate; \( C, \alpha_1, \alpha_3, \beta_1, \ldots \beta_k \) are constants, and \( \varepsilon \) a "random fluctuations" factor.

As is standard in econometric regression analysis, Ehrlich chose as numerical estimates for the constants in his model those values that achieved the greatest agreement between Equation (1) and the data he considered, which was aggregate data for the continental United States from 1933-70. Recognizing that there were no executions (i.e. \( p(e|c) = 0 \)) in the last three of those years, Ehrlich adopted a convention to avoid peculiar effects in Equation (1): he approximated \( p(e|c) \) for year \( b \) by \( 1/N_b \), where \( N_b \) was the number of convictions for willful homicide the previous year.
On the basis of economic utility analysis, Ehrlich had argued that, if there are rational murderers who consider in advance the risks their acts entail, then $\alpha_1$, $\alpha_2$, $\alpha_3$, the "elasticities" of murder rates with the three punishment variables, should all be negative and in order of decreasing magnitude. (His argument is summarized briefly in Appendix A.) His actual estimates, based on regression on his time-series data, were $\alpha_1 \approx -1.3$, $\alpha_2 \approx -.4$, $\alpha_3 \approx -.06$; Ehrlich considered this outcome strong evidence in support of his "rational-murderer" theory. Although his value of $\alpha_3$ (the elasticity of murder rates with execution rates for convicted killers) is close to 0, the difference is statistically significant and in fact implies a strong deterrent effect: every execution in the period studied prevented "on the average" 7 or 8 murders if the results are taken literally. Ehrlich has also performed similar regression analyses with cross-state data for certain years that apparently corroborate his original findings. He is almost surely the academic cited most often by the supporters of capital punishment.

Ehrlich's research contains subtle insights, careful distinctions, and extensive attempts to avoid biased results. But questions can be raised that cast doubt on the value of his conclusions. The next section of this paper describes a sampling of potential problems in Ehrlich's work that trouble this author; they concern definitions, assumptions, interpretations of results and "quality control" standards for the work. (These issues have apparently not been raised in the same form before.) Other of this author's concerns about Ehrlich's approach appear later in the paper. We restrict our attention to his initial study [2] although most of the comments also apply (or have analogies for) his more recent work.
Some Problems in Ehrlich's Work

We discuss below several potential problems in Ehrlich's work; others are discussed later in the paper. We will not consider the issues both obvious and subtle that can arise because of data inaccuracies; Nagin [4] has written an instructive paper on this matter. Rather we will concentrate on difficulties that would exist even if all data used were perfect. The list below is not intended to be exhaustive.

Obviously a potential murderer might not consider such quantities as \( p(a), p(c|a), \) and \( p(e|c) \) (defined earlier) as indicative of the risks he personally faces if he commits murder; he might, for instance, think himself craftier than the average killer and thus far less likely to get caught. But criticism of Ehrlich's work on this ground seems unwarranted, for it is not unreasonable to assume, as Ehrlich explicitly does, that a rational murderer's assessment of his own risk is monotonically related to appropriate aggregate probabilities. Yet there remain some apparent problems in Ehrlich's choice of risk indices in his model, as we discuss below.

(1) A potential murderer who is rational would presumably care immensely about the risks he is taking, as Ehrlich believes. But it is not clear exactly how these risks enter his "calculations." Does he, for example, ask himself the probability he'll be arrested if he commits murder or does he instead concern himself with the probability he won't be arrested for the killing, and thus will avoid all chance of punishment? The answer isn't obvious; it seems to depend on whether the person in question is an optimist (and starts with the feeling he'll "get away with it") or a pessimist (and starts off assuming that it's hopeless). One could argue that it should not matter at all, that since an 80% probability of arrest and a 20% chance of avoiding arrest mean the same thing, a rational person would make the same decision whichever number enters his thinking.
Two things are suggested by these comments: (1) Perhaps a model relating murder levels to arrest chances should have the property that the overall results are unaffected whether \(p(a)\) or \(1-p(a)\) is used as the arrest-risk variable; and (2) for a model lacking the above property, one should at least think carefully about which of these variables should be included. Ehrlich's model is greatly affected by the choice of arrest-risk index; the use of \(1-p(a)\) rather than \(p(a)\) in equation (1) would change not only the value of \(\alpha_1\), but all the other parameters in the equation, even, for instance, the estimated value of \(\beta_1\), the elasticity of murder rates with unemployment levels. More important, the accuracy of the model in explaining the data could theoretically be highly dependent on which arrest-risk variable is used. Suppose for example that, all other factors held constant, murder rates varied directly with the probability of avoiding arrest. Since there is no value of \(\alpha\) for which, for all \(p(a)\) from 0 to 1, \(p(a)^\alpha = (1-p(a))\), Ehrlich's model could not possibly capture the true effect precisely; in fact, it could only approximate it very crudely.

We are not trying to make the obvious point that some functions of \(p(a)\) are not well approximated by those of the form \(p(a)^\alpha\). Rather we are making the further point that among those functions poorly approximated are those of the form \((1-p(a))^\alpha\), a family that common sense suggests as likely as the one Ehrlich uses to contain the true description of how arrest risks influence murder rates. It is unsettling that when it seems important which of two equivalent -- and apparently equally plausible -- measures of arrest risk is used in Ehrlich's model, he apparently made a choice arbitrarily. And this problem, of course, affects the other punishment variables besides \(p(a)\).
Ehrlich's model assumes that the way the variables considered influence murder rates is unchanging over time. For some variables there seems no overwhelming reason to argue with this assumption, but for others the matter appears less clear. The unemployment rate U, for example, is included in the model partly because of the belief that some unemployed people, financially desperate, will commit property crimes (e.g. armed robbery) of which homicide is a by-product. One might conjecture that between 1933 and 1969 -- when social legislation and the growth of union power reduced considerably the financial hardships of unemployment -- the impact of U on murder levels might have decreased. Similarly, NW, the fraction of nonwhite Americans, presumably enters the model because of the view that even beyond its economic effects, racial discrimination is conductive to violence. But one might speculate that civil rights legislation, court decisions, and changing public attitudes between the 30's and the 60's lessened the role of race as a determinant of murder levels. (Ehrlich eventually found NW an insignificant factor in national homicide rates, but this does not lessen the intellectual case against his treating its influences as time-invariant.)

(3) It has been pointed out elsewhere (e.g. [5]) that Ehrlich may have ignored certain variables that exert influence on murder rates (e.g. the mean sentence length for those convicted of murder but not executed). But questions can be raised about some of the variables he did include, such as $p(c|a)$, the probability someone arrested for homicide is actually convicted of homicide. As used in the model, this variable obliterates the distinction between full acquittal and conviction on a lesser charge. Even ignoring this, there is the problem that it is not $p(c|a)$ that should concern potential killers but rather $p(c|a+g)$, the conviction probability given that a suspect is arrested for
murder and that he is actually guilty. (It would seem of little comfort to a potential murderer when an obviously innocent person is acquitted.) Clearly it would be improper to assume that \( p(c|a) = p(c|a+g) \); such an assumption would imply juries and judges completely incapable of discerning the guilt or innocence of homicide defendants brought before them. However, Ehrlich's use of \( p(c|a) \) would continue to be defensible if one could assume such a strong relationship between \( p(c|a) \) and \( p(c|a+g) \) that the former could serve as surrogate for the latter. The problem is that one can imagine circumstances where this might not be the case.

Suppose for example that, because of a well-publicized murder wave or charges of police inefficiency, there is greater eagerness on the part of police to make arrests quickly, with a corresponding drop in the accuracy of murder charges. Then in coming months \( p(c|a) \) might drop not because \( p(c|a+g) \) has changed, but instead because the fraction of defendants who are innocent has increased. (If this situation seems farfetched, consider a Supreme Court decision on evidence that reduces false arrests, and thus raises \( p(c|a) \) while leaving \( p(c|a+g) \) constant.)

To be sure, it is possible that the problem discussed above is not very serious. Perhaps, for example, the person planning murder believes (as many citizens apparently do) that almost all of those arrested and charged with murder are actually guilty, and that \( p(c|a) \) is a useful measure of the fraction of time the police obtain enough evidence to convict. But the discussion raises the possibility that Ehrlich's choice of aggregate risk index might not correspond to that of the rational murderer allegedly being modelled. (This particular problem also has a counterpart for risk of arrest.)
(4) Ehrlich believes a multiplicative model of the form of Equation 1 describes murder rates at the state level. He also assumes that a similar equation is applicable at the aggregate national level. Taken together these assumptions are inconsistent. For example, consider a nation of two states (1 and 2) with equal population, and suppose the murder rate \( M_i \) in state 1 follows \( M_i = C_1 U_1^2 \) where \( C_1 \) is a constant and \( U_1 \) the unemployment rate in state 1. Then \( U \), the national murder unemployment rate follows \( U = (U_1 + U_2)/2 \) and \( M \), the national murder rate, follows \( M = (C_1 U_1^2 + C_2 U_2^2)/2 \). In this case, there are no values of \( C \) and \( \beta \) for which the relation \( M = CU^\beta \) will be correct (except in pathological cases). The very consistency of statewide murder rates with Equation 1 generally precludes a similar consistency for the whole nation. This issue seems particularly important because Ehrlich uses cross-state analyses to corroborate his results based on national data.

The discussion in (1) - (4) above raises questions about the structure of Ehrlich's model. Many other such issues come to mind. The mere existence of these questions need not compromise the accuracy of Ehrlich's results; they do serve, however, to stress the burden on him to present evidence that his model successfully explains how murder rates evolve. The remainder of this section discusses how limited is the evidence Ehrlich offers in this regard.

Certainly Ehrlich provides no formula that predicts perfectly the national murder rates between 1933 and 1969. However he could not reasonably be expected to do so, because of the existence of random fluctuations in each year's murder levels, quite apart from any systematic factors. Even if, for example, a person has decided to commit a murder, the time he will attempt to do so may be influenced by random factors; whether he succeeds in killing rather than maiming his victim is itself often subject to chance. Thus Ehrlich's inclusion of a random factor (for each year) in his regression equation has some justi-
fication beyond the fact that "everyone does it." The variance of these factors (which reflects the extent to which they oscillate) is itself estimated from the regression analysis.

While traditional, Ehrlich's procedure to estimate the random factor's variance is quite troublesome from the standpoint of accountability. It assures virtually be definition that the effect attributed to randomness will be large enough to explain the discrepancies between the predictions of the model and the actual data. When a pollster predicts election results with a 3% market of random error, his accuracy can be faulted if he errs by 12%. No such embarrassment could arise for a model like Ehrlich's; if it typically errs by 12%, the estimated size of random errors will not be far away.

About the only direct tests related to accuracy allowed by a model like Ehrlich's are interval consistency checks, tests to assess whether the results of the analysis do not wildly contradict the assumptions used to obtain them. But with as few data points as Ehrlich has (37, one for each year), such tests are of very limited value. Perhaps because of this, Ehrlich cautioned that the validity of his results "is conditional upon that of the entire set of assumptions underlying the econometric investigation." But this brings us back to the original issue: why should we accept the assumptions Ehrlich made, some of which, as suggested above, seem rather questionable?

Actually, there is one piece of evidence that Ehrlich believed argued strongly for the accuracy of his results. As noted earlier, he attached great significance to the fact that, as his "rational-murderer" theory had predicted, the values of $\alpha_1$, $\alpha_2$, and $\alpha_3$ calculated in his regression were negative and diminishing in magnitude ("perhaps the strongest findings of the empirical investigation.") It should be recognized that Ehrlich actually had two major
hypotheses in his investigation: (1) that predictions of utility theory, based on assumptions of rationality within a noticeable subgroup of potential murderers, are sustained by actual murder patterns, and (2) that the relationship of murder rates to the factors that determine them follows the multiplicative functional form of Equation 1. Apparently he believed that the calculated values of $\alpha_1$, $\alpha_2$, and $\alpha_3$ constitute a strong plausibility argument for both his "rational murderer" theory and the regression model used to test it.

While these results should not be dismissed, there are two reasons they should be viewed with some caution. The elasticity $\alpha_1$ of the homicide rate with arrest probability (estimated value near -1.3) is based on arrest rates between 1933 and 1969 that hovered in a small range around 90%. $\alpha_2$ (estimated as roughly -.4) by contrast is calculated with conviction probabilities that fall in the vicinity of 45%. Ehrlich's interpretation of these results -- that potential murderers are more influenced by arrest probabilities than conviction rates -- is based on the assumption that, were the arrest rate to plunge to 45% and/or the conviction rate to reach 90%, the elasticities $\alpha_1$ and $\alpha_2$ he obtained would continue to prevail. Since neither of these things happened, this assumption must be regarded as sheer speculation, untested and untestable.

Furthermore, it should be noted that the regression theory Ehrlich used not only provided estimates of $\alpha_1$ and $\alpha_2$, but also specified probability distributions for the differences between the data-based estimates and the true elasticities. It emerges that the uncertainty associated with the estimate of $\alpha_1$ is fairly high, with a standard deviation about .7. Because of this, the difference between the calculated values of $\alpha_1$ and $\alpha_2$ is not statistically significant at the standard 5% level. This means that one cannot under usual criteria reject the hypothesis that the observed ordering of $\alpha_1$ and $\alpha_2$ is a
reversal — based on random fluctuations — of the actual ordering, an ordering inconsistent with Ehrlich's theory.

Under the circumstances, it seems a bit excessive when Ehrlich states that were his regression model wrong, only a "remarkable coincidence" could explain the agreement between his theoretical predictions about the $\alpha_i$'s and his calculated results. The fact is that Ehrlich has provided little evidence that his model explains murder rate changes well enough that its implications about capital punishment should be taken seriously.
The Bowers and Pierce Study

Ehrlich's study stimulated several other regression-based analyses of the deterrent effect of the death penalty. Most of these analyses concluded against the premise that executions deter murders. Some of the studies were reported in the context of criticism of Ehrlich's work, and their results presented as refutation of his findings on capital punishment. Perhaps the most famous paper of this kind was written by Bowers and Pierce [5]; it was the one cited in The New York Times, for instance (12/10/76, p.B7) to make the point that others had reached conclusions opposed to Ehrlich's. Below we summarize and discuss some principal points of the Bowers'Pierce argument. The references at the end of the paper include several other recent papers about the death penalty's deterrent effect; for reasons of space we will not discuss them here. However, the problems in the Ehrlich and Bowers-Pierce papers generally have analogues in the others.

Bowers and Pierce title their paper "The Illusion of Deterrence in Isaac Ehrlich's Research on Capital Punishment." They argue that (i) the data used by Ehrlich, particularly for the 1930's, is of dubious accuracy, (ii) when death penalty laws greatly vary from state to state, it can be misleading to use aggregate national data, (Baldus and Cole [6] argue this point in more detail), and (iii) Ehrlich's equation (1) is an inappropriate model when \( p(e|c) \) is very low. Based on these considerations, Bowers and Pierce find greatly significant their results that (a) if the years from 1964 on are excluded from Ehrlich's analysis, his model implies no deterrent effect of executions of statistical significance and (b) a linear regression model using Ehrlich's variables of the form:

\[
M = C + a_1 p(a) + a_2 p(c|a) + a_3 p(e|c) + b_1 U + ... + \varepsilon
\] (2)

where \( C, a_1, ... b_k \) are constants estimated from data and \( \varepsilon \) is a random-error term.
yields no values of $a_3$ — the execution-risk-coefficient — consistent with significant deterrent effects, even if data from 1964 on are included.

Below we argue that, as Bowers and Pierce have presented it, their argument is ineffective.

It is certainly true that as Bowers and Pierce point out, crime data were collected far less systematically in the 1930's than at present. Perhaps this fact affected the accuracy of Ehrlich's results, though, in using FBI figures, he can hardly be accused of choosing an eccentric data source. And there is an obvious problem in Ehrlich's use of national rather than state data, even beyond the consistency issue mentioned before in point (4). Ehrlich was apparently aware of some of these issues, states that he accompanied his "nation-based" regression with a cross-state analysis for 1960, which yielded what he called "quite consistent" results. While the statement was distressingly vague, it seems inappropriate that Bowers and Pierce chose to make no mention of it, for it does attempt to address simultaneously two of their major criticisms.

Bowers and Pierce are unhappy because, under Ehrlich's multiplicative model, the murder rate responds to the percentage changes in $p(e|c)$ rather than the actual changes. They state that Ehrlich's formulation means "a difference between one and two executions per 1,000 convictions will be greater than a difference between 350 and 650 executions." Like some other critics, the authors find strange Ehrlich's procedure for years when $p(e|c)=0$, which involves literally inventing an execution. To deal with these perceived problems, Bowers and Pierce (hereafter BP) apparently prefer a linear regression model (Equation 2).

Despite BP's observation, one could argue that when $p(e|c) > 0$, Ehrlich's assumption about its relationship to $M$ need not necessarily be implausible. It is at least conceivable that a "law of diminishing returns" prevails, and
that the incremental deterrence of a given execution is far greater when executions are rare than when they are plentiful. In their specific example, Bowers and Pierce seem to object to the suggestion that doubling \( p(e|c) \) at a very low level (.1%) has more impact than raising \( p(e|c) \) from 35% to 65%. But one might well conjecture than when \( p(e|c) \) is as high as 35%, those not deterred from killing are simply not greatly affected by the fear of their own deaths.* Thus even a substantial further rise in \( p(e|c) \) might have very little effect, so of the two effects BP compare, perhaps the relative impacts are as Ehrlich's model specifies. In the linear model BP use as an alternative, raising \( p(e|c) \) from 35% to 65% has 300 times the effect of raising \( p(e|c) \) from .1% to .2%. If this assumption is somehow more plausible than the one Ehrlich makes, it is not apparent to this author.

Ehrlich's decision about how to proceed when \( p(e|c) = 0 \) -- which he makes no attempt to justify -- does seem arbitrary and disturbing. (Presumably this issue also came up in his cross-state analysis.) But the many assumptions that Bowers and Pierce make in their linear model are not really less arbitrary simply because they are more familiar. And a linear model with some negative coefficients allows for the possibility that in a set of realizable circumstances, the homicide rate will be negative. To prevent this absurdity, Bowers and Pierce would presumably have to adapt some conventions, conventions not terribly different than Ehrlich's for \( p(e|c) = 0 \). In any case, Ehrlich showed that deleting the years when \( p(e|c) = 0 \) (i.e. 1968 on) leaves his estimates of \( \alpha_3 \) virtually unchanged.

* Assuming, of course, that the probability of conviction given guilt is itself substantial.
As noted earlier, BP found that when years after 1963 were deleted from the analysis, they were unable, using Ehrlich's model, to generate negative values of $\alpha_3$ (elasticity of $M$ with $p(e|c)$) of statistical significance. One possibility suggested by this finding is that the deterrent effect of the death penalty changed over time, with executions, in their twilight years of the late '60s, exerting greater deterrence than before. BP argue that this was not the case; thus they apparently have no quarrel with the time-independence assumptions Ehrlich made. Rather they note that $p(e|c)$ was extremely low (and often zero) in the years after 1963, which they believe makes Ehrlich's model unreliable for that period. A regression analysis from 1935-69 with a linear model (which BP, as noted, consider less vulnerable for small $p(e|c)$) did not ratify Ehrlich's conclusions about a significant deterrent effect of death sentences. Thus, BP decided, Ehrlich's result was a "statistical artifact" and an "illusion."*

We have suggested that BP may have exaggerated the weaknesses of Ehrlich's model relative to others when $p(e|c)$ is small. How much weight should be given to the fact that their linear regression results differ from Ehrlich's on the crucial issue of capital punishment? Whatever the true effect of executions, it is virtually certain that there will be some models whose coefficients, when calculated from the data, will imply the opposite effect. Thus the mere fact that two models disagree is not particularly illuminating. There do, however, exist standard methods for measuring the relative explanatory power of two different models. Had BP shown with these methods that their linear

* BP also suggest that the use of national rather than state-by-state data is particularly inappropriate for the 1960s. However, their argument, based on a simple-correlation technique, is nonrigorous and potentially misleading (see [3], p. 215).
model explained the evolution of murder rates better than Ehrlich's exponential, they would greatly have weakened Ehrlich's argument. However, no only did they not do so, but for some reason they did not even mention the existence of such methods. Indeed, Ehrlich, in a rebuttal to BP [3], claimed that comparative tests show his multiplicative format "decisively superior" to BP's.

It does not seem that Bowers and Pierce have justified their devastating claims against Ehrlich's work. Yet it seems important that BP found that, even under his own model, Ehrlich's original capital punishment finding depends very heavily on the years since 1964 (BP's pursuit of the point, however, leaves its true significance quite unclear.) Ehrlich calls this BP finding an "unfounded claim" [3, p.217] but he does not dispute the accuracy of their calculations. Nor, on intellectual grounds, can his citation of corroboratory cross-state results dispel this specific problem in his national time-series analysis, the analysis he has chosen to make the "flagship" of his reported results.
Another Approach

These comments were intended to provide an introductory -- though by no means exhaustive -- view of the "stalemate" among researchers about capital punishment. This author does not believe that any specific evidence has been presented that indicates with any conclusiveness whether executions deter murders. Nor does he believe recent researchers on the subject have given convincing reasons why their models might, even in principle, be expected to identify the true effect of executions even given perfect data. In such circumstances, it would certainly not be redundant to attempt with new approaches to assess the impact of the death penalty. Below we discuss some considerations that would seem important in any such new research.

First and foremost is the introduction of strict standards of accountability for the models explored. As mentioned earlier, the concept of random fluctuations in murder rates -- at both state and national levels -- has grounding in reality, but it should not become an "elastic clause" that can explain away deficiencies in a model. From first principles -- unrelated to any particular models of deterrent effect -- one can attempt to estimate the magnitude of chance variations in annual murder rates, assuming that they are measured accurately. (This is akin to the pollister's estimate of the uncertainty in his results given that his sampling was genuinely random.) A first model along these lines appeared in [10]. The objective is to furnish a first "yardstick" for measuring whether a model that purports to explain the evolution of murder rates does so to an acceptable degree of precision. A random-fluctuations model is not in itself enough to assess the accuracy of a causal model; obviously, with enough arbitrary parameters, one could achieve perfect agreement between predictions and the data used to generate them. But coupled
with other techniques, independent estimates of the size of random error could provide accuracy standards far stronger than those hitherto employed in capital punishment research. The aim should not be to identify any one model as the "correct" explanation of the effect of executions; rather it is to identify the range of explanations reasonably consistent with the data at hand.

Obviously the possibility exists that no models will be found that meet standards of precision obtained independently. But even if that happened, it would simply imply that no theory examined explained the behavior of murder rates to a fully acceptable extent; thus any results obtained should be viewed with great caution. A statement to that effect -- based on specific accuracy standards -- might mean more to those studying the research than a vague disclaimer that perhaps the model used was not perfect, a disclaimer that seems related more to modesty than anything else.

A second reed is for greater flexibility than has so far been allowed in the functional form of the relationship between execution-rate variables and homicide rates. Such flexibility might increase the realism and power of deterrence studies, for the restrictions implicit in both Ehrlich's and BP's equations (and other recent work) seem fairly severe. Consider, for instance, the set S of individuals who might be deterred from murder by the threat of execution but not by threats of conviction and incarceration. What is the relationship between their decision on whether to kill and the prevailing value of \( p(e|m) \), the aggregate fraction of capital murders* whose convicted perpetrators are executed? One might conjecture that for very small values of \( p(e|m) \), the death risk associated with committing murder might not make a strong impres-

* Murders subject to the death penalty (e.g. premeditated). Presumably the very act of thinking about the murder would make the rational murderer's crime a capital one.
sion, given comparable short-term risks of death through accidents or illness. Thus potential murderers might treat very small risk-levels as effectively equivalent to zero, with the result that associated deterrence is very small. Beyond a certain "threshhold" level, however, execution might be perceived by those contemplating killing as a "clear and present danger," and thus deterrence in group S might grow rapidly with \( p(e|m) \). But one might conjecture the existence of a second "threshhold," at which \( p(e|m) \) is sufficiently high that those not yet deterred are probably undeterrable, making further increases in \( p(e|m) \) almost superfluous. These comments imply an "S-curve" relationship between \( p(e|m) \) and \( q \), the fraction of those in group S who will be deterred from murder by current execution rates (see Figure 1).

![Graph](image)

Figure 1: A Model of Deterrence vs. Execution Rate for those Undeterred by Nonlethal Sanctions (\( t_1 \) and \( t_2 \) are the two threshholds discussed)

Obviously the discussion above is speculative, and may in fact bear no relationship to the true effect of executions. But the important point is that it is impossible, using models like Ehrlich's and BP's, even to consider (except very approximately) the possibility of S-curve behavior. Cases like this raise the possibility that restrictions on the functional forms allowed can obscure even the general direction of the effect of interest.
The two issues just mentioned raise the larger issue of the appropriate role of regression analysis in capital punishment research. Buttressed by an elegant theory and vast computer packages, regression models have great appeal in criminology as elsewhere. But mathematical models of a process ideally should arise from attempts to quantify reasonable assumptions about it, rather than from an excessive concern about tractability and familiarity. It would be foolish to criticize regression-based models simply because they are familiar, but in the analysis of capital punishment data, no other sophisticated approach has really been tried. Exploring the power of alternative methods to regression might be a worthwhile objective for future death penalty research.

The current author, with the aid of three MIT graduate students, is investigating the effect of capital punishment in accordance with the considerations above, and hopes to present detailed results in a forthcoming paper.

It should be noted that the way the death penalty was used before 1968 reduces the capacity of any model to show the ineffectiveness of executions. Values of \( p(e|m) \) have only rarely, in the past 40 years, exceeded about 2% in any American state.* One wonders if an execution risk level near 2% could really seem threatening to a potential murderer. In the context of Figure 1, the question would be: is \( t_1 > .02? \)

It is perhaps relevant in this connection that at rates prevailing in 1972, an "average" lifelong Californian would have roughly a 2% chance of eventually being killed in an automobile accident. (Lower speed limits since then -- caused by the energy crisis -- have lowered the figure to 1.5%.) Anyone familiar with the Golden State will have great difficulty believing that

* This figure refers to all murders, not just capital murders, for which the figures are a bit higher. However, someone clever enough to make the distinction in his calculations might also, one might speculate, think himself so capable that his personal risk level is below 2%.
the risks its citizens face in any way reduces their driving. One could object to the parallel being suggested here; a murderer sustains a 2% risk on the basis of one act, whereas, for any one ride, a Californian's chances of death are very small. This is true, but perhaps the longer-range behavior of Californians is instructive. Residents of Los Angeles seem to consider a year wasted if they have not voted down another rapid transit bond issue. San Francisco does, to be sure, have a new rapid transit system (BART), but its apparent impact on auto commuting has been very marginal indeed.

These comments suggest that data available now might well be inherently incapable of proving conclusively that executions can't deter murders. About the most one might be able to say is that very small execution risks yield no observable effects compared to no executions at all. Past data might, however, offer some significant evidence that the death penalty does deter killings, or, as has recently been suggested, that its prime effect is to stimulate them.

On January 17, 1977, the ten-year halt in executions in American came to an end. Whether they want it or not, researchers on the death penalty will presumably be getting a new data base relevant to the deterrence question. It would probably be desirable if now -- when charges of ex post facto criteria could not be raised -- a consensus arose among researchers on the proper ways to evaluate the forthcoming data. But somehow it seems naive to expect such a consensus. Thus it may happen that, 20 years from now, Americans will be no surer than they are today on the effects of the august decision the country is now making. We can only hope not.
If a person commits a murder, four possible consequences related to punishment can arise:

1. No arrest, thus no punishment
2. Arrest, but no conviction for homicide
3. Arrest, conviction for homicide, some nonlethal punishment
4. Arrest, conviction, execution.

Suppose the number $U_i$ is some measure of a potential murderer's happiness (utility) if outcome $i$ occurs. If the person in question is rational, Ehrlich conjectures, the numbers $U_1, U_2, U_3, U_4$ will be decreasing. (The last three could presumably be negative.) Let $p_i$ be the probability that the murder will result in outcome $i$. The $p_i$'s are related to the risk-variables $p(a), p(c|a), p(e|c)$ defined earlier by the equations:

\[
\begin{align*}
  p_1 &= 1 - p(a) \\
  p_2 &= p(a)(1 - p(c|a)) \\
  p_3 &= p(a)p(c|a)(1 - p(e|c)) \\
  p_4 &= p(a); (c|a)p(e|c)
\end{align*}
\]

These equations follow from basic laws of probability; $p_4$, for instance, is the product of the probabilities of arrest, conviction given arrest, and execution given conviction because outcome 4 requires all three of these events to occur.

The number $U$ given by $U = U_1p_1 + U_2p_2 + U_3p_3 + U_4p_4$ is a measure of the "average" unhappiness that commission of the murder will bring to its perpetrator, since it
weights the unhappiness associated with each of the four outcomes by the probability that outcome occurs.

It is an easy exercise in calculus to show that, for any particular set of values of \( p(a) \), \( p(c|a) \), \( p(e|c) \), the effect on \( U \) of raising \( p(a) \) by \( k \) percent (i.e. to \( 100+k \) percent of its original value) is greater than that of a \( k \) percent rise in \( p(c|a) \), which in turn has greater impact than a \( k \) percent rise in \( p(e|c) \). (Naturally, we consider only those \( k \) for which a \( k \) percent rise will not push any of the quantities above the value one.) Thus, Ehrlich argued, a rational person contemplating murder will be more sensitive to changes in the probability of arrest than in \( p(c|a) \), and will be affected least of all by changes in \( p(e|c) \). This belief translates into the prediction that the elasticities \( \alpha_1, \alpha_2, \alpha_3 \) in equation 1 should be diminishing in magnitude.
References and Bibliography


8. B. Forst, (Reference to be supplied).

