MF-1

THE DETERRENT EFFECT OF CAPITAL PUNISHMENT:
II. YET ANOTHER STUDY*

ABSTRACT

Data from several recent years and 33 large American states are examined in yet another attempt to see if executions can deter murders. Since any analysis of this kind is based on numerous assumptions, we pay particular attention to the question: how can one tell if such assumptions are sufficiently accurate that the results of using them warrant serious attention? We obtain results that, as we will explain, lean slightly but not unambiguously toward the view that capital sanctions can cause reductions in homicide rates.

ARNOLD BARNETT
Sloan School of Management
Operations Research Center
Massachusetts Institute of Technology
Cambridge, Massachusetts

84551

paper was supported under Grant No. 78-NI-AX-0034, from the U.S. Depart- of Justice, Law Enforcement Assistance Administration, National Institute w Enforcement and Criminal Justice. Points of view or opinions stated in document do not necessarily represent the official position or policies e U.S. Department of Justice.

Acknowledgements

I am deeply grateful to my students Gary Goehrke, Daniel Kaplan, David Lintz, Sharon Lowenheim and William Walthall for extensive help in data gathering and statistical computing. I am likewise grateful to Daniel J. Kleitman, Richard Larson, John D. C. Little, Michael Maltz, Edwin Zedlewski, Frank Zimring and two highly conscientious referees; their astute comments greatly improved the paper. No inference should be drawn that the people just mentioned necessarily share all this paper's conclusions.



Introduction

The recent debate in the United States on whether to restore capital punishment has been dominated by the question of whether executions can deter murders. Indeed, the very existence of this debate is probably traceable to the fact that, after executions virtually ceased in the mid-1960's, homicide rates doubled within a decade. Numerous statistical studies on the deterrence question have appeared in recent years, and several have received widespread attention in the press, state legislatures, and the courts. This paper is yet another study on the deterrent effect of capital punishment.

Why another study? We might begin by noting that even when capital punishment was at its peak in the U.S., executions were rather rare compared to homicides. Thus, even if each execution had considerable deterrent power, the death penalty was probably among the smaller influences on overall murder* rates. It follows that unless a capital punishment study can accurately describe the large "extraneous" influences on homicide levels, it is unlikely to provide useful information on the question that motivated it.

In [1], we examined in detail three of the most prominent recent studies on the deterrent effect of capital sanctions. We tried with actual data to estimate the power of their underlying mathematical models and found that, quite beyond the usual effects of chance fluctuations, their predictions were subject to systematic errors that were almost surely larger than the effect they attempted to measure. This does not imply that their conclusions are necessarily wrong, but it does indicate that there is no compelling reason to believe they are right. In this circumstance, another attempt to analyze the available data is by no means redundant; that is why this paper exists.

Like earlier researchers, we recognize that one can only assess the effect of

*Throughout this paper we use the words "murder," "killing" and "homicide" interchangeably, meaning in all cases willful homicides except those by police as recorded in Vital Statistics of the United States executions in the context of a broad macroscopic analysis to explain various influences on levels of killings. We are thus forced to construct mathematical models of the evolution of homicide patterns, models that, while reflecting the influence of previous scholars, nonetheless differ from earlier ones in choices of both key variables and functional forms. A crucial question about any such model is whether it describes actual murder patterns with sufficient accuracy that its implications on capital punishment deserve serious attention. In this paper, we amplify and extend our discussion on this subject in [1] and proceed to develop "Region I criteria" that, we argue, all models reliable enough for our purposes should meet. Then, rather than concern ourselves with finding any "best fit" model to available data, we explicitly address the questions: (1) do <u>any</u> of those we considered meet the stringent Region I criteria and (2) if so, what is the full range of models that do so? This model-evaluation procedure, we believe, provides a natural framework for full and frank description of what the analysis has achieved.

We focus on data from 33 American states and 6 years in the period 1950-65. Because changes in execution rates over this period were relatively large, it should in principle be especially useful given our objectives. We find that a noticeable subset of the models we construct meet the Region I criteria; thus we avoided the "dry hole" of a series of insensitive models whose results are of little interest.

Our results tilt slightly but not conclusively towards the view that executions exert some deterrent effect on potential killers. By this we mean that deterrence-based models seem a bit more consistent with the data examined than their nondeterrence counterparts. But the fact that state execution rates have always been very low in this country seems to render the data incapable of discriminating clearly among the various hypotheses about deterrence. Such a conclusion might be described as belaboring the obvious were it not for the fact that earlier researchers, examining the same data, have discerned strong implications on the deterrence question.

The organization of the paper is as follows: in Section I-V, we describe how



we attempted to model the different influences on state murder levels. In Sections VI and VII, we propose "standards of accountability" for such models, and discuss some inherent limitations in a data analysis of this kind. We assess in Section VIII the soundness of the models we have built, and what these models imply about the deterrent effect of capital punishment.

## I. Modelling Homicide Patterns

A <u>mathematical model of homicide levels</u> is a formula for predicting the number of killings in a locality, given the values of the variables that allegedly influence it. Because of the boundless flexibility in the choice and definition of key variables and in the functional form of the overall relationship, the number of possible homicide models is unlimited. Any given researcher, however ambitious, actually restricts his attention to an infinitesimal fraction of the set of conceivable models.

In the next four sections of the paper, we describe the set of models we have chosen to investigate. There is no reason why, on learning our many simplifying assumptions, the reader will be struck by their inherent wisdom. The models are of interest because, as we will show, some of them are rather consistent with a large amount of data; indeed, considerably more so than the models of earlier researchers.

We focus on the murder levels in various American states. Broadly speaking, there are ten factors that might plausibly influence such levels:

1) Frequency and Severity of Punishment for Homicide
2) Total State Population
3) Ethnic Composition of the Population
4) Degree of Urbanization of the Population
5) Age Distribution of the Population
6) General Economic Conditions
7) Quality of Emergency Medical Care
8) Violence in the Media
9) Availability of Guns
10) Tradition of Violence

It is clear that some of the factors above would be difficult to quantify even in principle; others, while not subject to this problem, nonetheless suffer because needed data are unreliable or nonexistent. It follows that attempting to describe the effects of particular factors on homicide rates is a hazardous undertaking in which success is uncertain. This fact underscores a point that has received insufficient attention from earlier researchers: unless a model's accuracy is empirically demonstrated, there is little reason to pay attention to the results of using it.

We concentrate here not on the murder levels in individual states in given years, but rather on the changes in those levels between two years separated by several others. As Forst [ 5 ] noted, focusing on such changes provides a natural way to include local variations in "traditions of violence" into the analysis. Furthermore, we note that the importance of particular factors in explaining homicide patterns may vary across time and place (e.g., race may be a larger component of murder patterns in Mississippi than in Minnesota, and in 1950 than in 1965). For this reason, we allow considerable time-dependence and regional dependence in the inclusion of variables into our models.

We focus in this paper on the period 1950-65, one in which executions dropped sharply in frequency. We do not consider data from earlier years because of skepticism about their completeness and accuracy (e.g., see [2 ,3 ]); we exclude years after 1965 since almost no executions have taken place anywhere in the U.S. since then.* For each of 33 states and four different pairs of years, we attempt to estimate $C$, the growth (or contraction) factor in the state's homicide level between the first year studied and the second. We will argue that $C$ can usefully be treated as a product of three factors:

$$C = C_D C_P C_T$$

where

$C_P$ reflects changes in the state's pattern of punishment over the period of interest.

$C_D$ reflects demographic changes associated with factors 2-4 on the list.

*One could argue for the inclusion of years after 1965 to help decide whether states that gave up executions suffered greater increases in murder than those that never used them. However, data from the early 1960's are quite informative on this point.



$C_T$ is a regional time-trend factor meant to reflect, at least implicitly, factors 5-9 on the list.

Our main objective is to examine how $C_P$ is actually related, if at all, to the frequency of use of the death penalty. We will consider a large number of possible definitions of $C_P$, each one of which is effectively a hypothesis about the deterrent effect of capital punishment. Through data analysis, we hope to separate those hypotheses with some empirical foundation from those that have none.

In Sections II-V, we "flesh out" our models by discussing the definition and calculation of $C_D$, $C_P$, and $C_T$, and by suggesting why it seems reasonable to treat $C$ as a product of these three quantities. In Section VI, we begin talking in earnest about statistical testing of such models.

## II. Demographic Factors and Homicide Levels

Here we discuss the influence of population size, ethnic breakdown, and degree of urbanization on a state's murder levels. The importance of the first of these factors is obvious; the last two are significant because historically, both victimization and commission rates for homicide have been substantially higher among blacks than whites nd among city dwellers than rural residents. (Since blacks tend disproportionately to live in cities, these two phenomena are somewhat related.) In the discussion below we focus on victimization patterns for murder rather than commission patterns; the two are so strongly related, however, (e.g., 7/8 of all murders are intraracial) that this is no wild simplification.

There are some subtle aspects to modeling the effects of race and urbanization on changes in a state's murder rates. Beyond the changing proportions of blacks and city-dwellers there is the issue that the "risk premium" associated with being black and/or being urban is probably not constant over time. Improvements in emergency medical care than can prevent aggravated assaults from becoming homicides tend to become available earlier in large cities than rural areas. The stresses of urban life that might be conducive to violence (noise, pollution, crowding, etc.) may change in intensity over time. And if the frustration caused by racial discrimination is conducive to violence, then the reduction of racial prejudice -- associated with court decisions, legislation, and changing attitudes -- may lessen the role of race as a determinant of murder levels. In short, a model intended to correct for such demographic influences should allow for time-dependent risk factors.

At the same time, such a model should probably treat the influence of race and urbanization on murder levels as region-dependent. Until recently, the legal status of blacks was quite different in the South than outside it; furthermore, there is a greater tradition of rural violence in the South than elsewhere. And differences between industrial cities of the East and Midwest and their more spacious counterparts in the West might be reflected in differing urban-risk increments.

A crude model corresponding to the discussion above might go as follows: suppose $\alpha$ is the murder victimization rate among rural whites in a given state and a given year. Then one might hypothesize the existence of two risk multipliers B and U such that the victimization pattern for the state is

$$\begin{aligned} \alpha &= \text{victimization rate for rural whites} \\ B\alpha &= \text{victimization rate for rural blacks} \\ U\alpha &= \text{victimization rate for urban whites} \\ BU\alpha &= \text{victimization rate for urban blacks} \end{aligned} \qquad (H_o)$$

B and U are assumed both time-dependent and region-dependent (North, South, West) but the same for all states in a given region in a given year (e.g. the same in New Jersey and Massachusetts, and in Alabama and North Carolina). Citizens are classified as "urban" or "rural" depending on whether or not they live in communities with population exceeding 100,000. This obviously crude division has a basis in homicide data: suburban and rural murder patterns are much closer to each other than either is to patterns prevailing in larger cities. The category "whites" is assumed to include Orientals; American Indians are classified as "black." The model makes no distinction between recent immigrants to a state and long-term residents.



The Vital Statistics of the United States provide data on the actual fraction of murder victims who are black for the 30 states and six years of interest to us; one can also use them to compute the fraction of urban victims. (One cannot, however, obtain further breakdowns of the data, such as the fraction of victims who are both black and urban.) From these Vital Statistics we obtain for each region in each year the maximum likelihood estimates of B and U under $H_o$ (i.e. those estimates under which the observed results would have the highest probability of arising). The results are given in Table 1 below.

TABLE 1: CALCULATED DEMOGRAPHIC RISK PREMIUMS FOR HOMICIDE IN THREE REGIONS

| | 1950 | 1953 | 1958 | 1960 | 1963 | 1965 |
|---|---|---|---|---|---|---|
| $B_N$ | 12.41 | 10.83 | 8.89 | 8.19 | 7.25 | 7.67 |
| $U_N$ | 1.62 | 1.67 | 1.77 | 1.86 | 2.02 | 2.05 |
| $B_S$ | 6.69 | 6.61 | 5.92 | 6.28 | 5.97 | 5.67 |
| $U_S$ | 1.47 | 1.69 | 1.50 | 1.34 | 1.40 | 1.54 |
| $B_W$ | 5.94 | 8.08 | 6.60 | 6.82 | 7.35 | 6.12 |
| $U_W$ | 1.08 | 1.11 | 1.22 | 1.34 | 1.34 | 1.62 |

N = North; S = South; W = West

These maximum likelihood estimates of B and U are of interest only if the model $H_o$ bears some sensible relationship to the actual distribution of murder risk among the four cited demographic groups. This matter can be investigated empirically. We examine data for the six years listed above from 30 American states which account for about 80% of all homicides in the U.S. These states are

| NORTH | | | SOUTH | | WEST |
|---|---|---|---|---|---|
| Connecticut | Massachusetts | New York | Alabama | Mississippi | California |
| Illinois | Michigan | Ohio | Arkansas | N. Carolina | Colorado |
| Indiana | Minnesota | Oklahoma | George | S. Carolina | New Mexico |
| Kansas | Missouri | Pennsylvania | Kentucky | Tennessee | Oregon |
| Maryland | New Jersey | Wisconsin | Louisiana | Texas | Washington |

The other 20 states are excluded from the analysis because they either i) have no cities and very few blacks, ii) have very few killings, or iii) have key cities whose boundaries have changed drastically, making the definition of urban status perilous.

If we neglect multiple murders, we might express the hypothesis $H_o$ probabilistically as follows: if fraction $p_1$ of a state's residents (in the year studied) are rural whites, fraction $p_2$ are rural blacks, $p_3$ are urban whites, and $p_4$ urban blacks, then, for a randomly chosen homicide that year, the probability $P_B$ that the victim was black follows:

$$P_B = \frac{Bp_2 + BUp_4}{p_1 + Bp_2 + Up_3 + BUp_4} \qquad (1)$$

Thus if N murders occurred in the state that year, then, under an independence assumption for different killings, the number $\hat{N}_B$ with black victims would follow the standard binomial distribution:

$$Pr(\hat{N}_B = K) = \binom{N}{K} P_B{}^K (1 - P_B)^{N-K} .$$

If N is reasonably large, then $\hat{P}_B$, the fraction of killings with black victims (i.e., $\hat{N}_B/N$), would be approximately normally distributed with mean $P_B$ and variance $\sigma_B{}^2 = P_B(1 - P_B/N)$ if $H_o$ is correct.

Similarly, the probability $P_R$ under $H_o$ that the victim of a randomly chosen murder lived in the city follows:

$$P_R = \frac{Up_3 + BUp_4}{p_1 + Bp_2 + Up_3 + BUp_4} . \qquad (2)$$

$\hat{P}_R$, defined analogously to $\hat{P}_B$, is likewise normally distributed; its parameters are $P_R$ and $P_R(1 - P_R)/N$. Actually, because blacks live in cities in disproportionate numbers, $\hat{P}_B$ and $\hat{P}_R$ are bivariate normal with positive correlation. These facts are useful in assessing the extent to which discrepancies between $P_B$ and $\hat{P}_B$, and $P_R$ and $\hat{P}_R$, can be attributed to chance.

For each state and each year of interest, we calculate $P_B$ from (1) with the relevant B and U values from Table 1 and estimates of $p_1 - p_4$ based on The Census of the United States. We then record the actual value of $\hat{P}_B$ from Vital Statistics of the



United States, and proceed similarly for $P_R$ and $\hat{P}_R$. Typical of the results obtained are those presented in Table 2 for Southern states in 1960.

Over the entire 6-year set of 180 racial predictions, the average absolute error $\hat{P}_B - P_B$ was 3.7 percentage points, compared to the error of 2.7 points* one would have expected from chance fluctuations alone under the binomial process described. The predicted fraction of murder victims who lived in cities erred on the average by 4.0%; the expected error were the model perfectly correct would be 2.9%*. But the differences between the predicted and actual risk breakdown, while small, were nonetheless statistically significant.

TABLE 2: PREDICTED AND ACTUAL DISTRIBUTION OF HOMICIDE RISK BY RACE, RESIDENCE AND STATE: SOUTH 1960

| State | Fraction of Victims who are black | | Fraction of Victims who are urban | |
|---|---|---|---|---|
| | $P_B$ (Predicted) | $\hat{P}_B$ (Actual) | $P_R$ (Predicted) | $\hat{P}_R$ (Actual) |
| Alabama (344) | .736 | .725 | .279 | .276 |
| Arkansas (133) | .638 | .639 | .082 | .098 |
| Georgia (430) | .720 | .714 | .276 | .253 |
| Kentucky (185) | .341 | .319 | .225 | .265 |
| Louisiana (269) | .751 | .743 | .375 | .420 |
| Mississippi (175) | .821 | .817 | .078 | .080 |
| North Carolina (420) | .683 | .686 | .159 | .167 |
| South Carolina (238) | .771 | .681 | .051 | .055 |
| Tennessee (270) | .574 | .589 | .542 | .493 |
| Texas (726) | .481 | .510 | .484 | .488 |

NOTE: The number after each state's name is its number of homicides in 1960, according to Vital Statistics of the United States. The smaller this number, the larger the expected discrepancy between $P_B$ and $\hat{P}_B$ (or $P_R$ and $\hat{P}_R$) based on chance fluctuations.

While not a perfect model, $H_o$ does emerge as fairly accurate, perhaps surprisingly so since it wholly ignores the local conditions in individual states. Its errors are only about one percentage point over the "random-noise" level which suggests that a typical estimate of $P_B$ from (1), or $P_R$ from (2), errs by .02 or less.** All things

*These numbers are already adjusted downward because the estimates of B and U are data-based. In all, one parameter of the model was estimated per ten data points.

**This estimate is based on the relationships (i) MEAN SQUARE PREDICTION ERROR = MEAN SQUARE FLUCTUATION + MEAN SQUARE MODEL ERROR, and (ii) MEAN ABSOLUTE MODEL ERROR ≤ $\sqrt{\text{MEAN SQUARE MODEL ERROR}}$.

considered, a demographic correction based on $H_o$ and the calculated constants in Table 1 seems an adequate first approximation to the changing effects of race and urbanization. We will use this approximation in the models we are constructing; the precise way we do so is described below.

Among the dimensions along which states within a region differ considerably are several of great interest to us, including the frequency of use of capital sanctions. The strength of $H_o$, which ignores such differences, does not imply the irrelevance of these factors towards murder levels; it does suggest, however, that proportionally speaking, they exert roughly the same influence on blacks and whites, and on city dwellers and rural residents (i.e., their effect is absorbed in $\alpha$ in $H_o$).

We therefore assume that, to first order, the homicide rate h (i.e., murders per 100,000 residents) in a given state can be approximated by the expression

$$h = \alpha(p_1 + Bp_2 + Up_3 + UBp_4)$$

where B and U are as defined earlier for the appropriate region and time, $p_1 - p_4$ are obtained from census data, and $\alpha$ is the state's murder rate (per 100,000 people) among rural whites.

We will consider changes in h between two different years, using throughout the subscripts 1 and 2 to represent the first and second year respectively. Were $\alpha$ to remain constant between the two years, then in our approximation, the murder rate would change by a factor $D_{12}$ given by:

$$D_{12} = \frac{p_{12} + B_2p_{22} + U_2p_{32} + B_2U_2p_{42}}{p_{11} + B_1p_{21} + U_1p_{31} + B_1U_1p_{41}}$$

The total number of felony killings would change by the factor $I_{12}D_{12}$ where $I_{12}$ is the ratio of total population in the later year to that in the earlier. Were $\alpha$ not to remain constant, the number of murders would change by the factor

$$\frac{\alpha_2 I_{12} D_{12}}{\alpha_1} .$$

Let $C_D = I_{12}D_{12}$. The arguments above imply that, except for chance fluctuations,



the ratio C of the second year's murder level to that of the first should be nearly proportional to $C_D$, with $\alpha_2/\alpha_1$ the proportionality factor. Put another way, data patterns imply not only the utility of a particular definition of $C_D$, but also the appropriateness of treating $C_D$ as a factor of C. We will accept this second implication as well as the first in building overall models of homicide patterns.

## III. Punishment for Homicide

We use three variables to measure the frequency and severity of punishment for homicide:

$P_G$ = the fraction of killings that ultimately result in a conviction for homicide,

L = the median prison sentence for those convicted of homicide and not executed,

$P_e$ = the fraction of killings whose convicted perpetrators are eventually executed.

Our key assumption is that, at any given time, the number of murders in a locality is proportional to a factor $f(P_G,P_e,L)$ that reflects the impact of punishment patterns on homicide levels. $f(P_G,P_e,L)$ is assumed to have the following properties:

(1) $f(P_G,P_e,L)$ is time independent.

The alternative assumption, after all, makes studies of past patterns of dubious relevance.

(2) $f(P_G,P_e,L)$ is the same for all demographic groups in a given state.

This assumption is buttressed by the results reported in the last section. The hypothesis that, for example, blacks and whites respond very differently to sanction policies is inconsistent with the fact that, across the states of a given region, rather large variations in punishment patterns had no observed effect on the black-white murder risk differential.

(3) $f(P_G,P_e,L)$ is the same in all states.

Without homogeneity assumptions such as this one, the data base would be subject to endless fragmentation.

Taken together, these assumptions imply that, for two separate years, the ratio $\alpha_2/\alpha_1$ of the murder rates within a given population subgroup (e.g., rural white) is proportional to $C_p$, where $C_p = \frac{f(P_G(2),P_e(2),L(2))}{f(P_G(1),P_e(1),L(1))}$ .

Our aim is to estimate which "influence functions" (f's) seem most consistent with actual data. The study of such functions does not imply that all potential killers are aware of current $P_G$, $P_e$ and L values, that all those cognizant are influenced by them, or that those who are influenced believe their personal risk levels exactly those given by aggregate statistics. The f's are simply intended to reflect the net macroscopic effect of changes in homicide punishment levels.

How does one estimate the current values of $P_G$, $P_e$, and L in a given state? To approximate $P_G$, we follow a procedure of Forst [5] which, roughly speaking, estimates the conviction rate for killing committed two years earlier. To estimate L, we use state data from the National Bureau of Prisons on the mean (or median) time served by those released from prison in given years; these statistics, it has been suggested, are better guides to incarceration patterns than sentences pronounced in courtrooms. Estimating $P_e$ is a bit trickier. Since executions have been rather rare in most states, it is not uncommon for calculated execution rates to vary from year-to-year by factors of two or three. Thus how $P_e$ levels are perceived by potential killers depends both on (i) how far back their memories extend and (ii) how much greater weight they place on recent patterns than on those further back.

Lacking any clear idea how to estimate $P_e$, we proceed, like Ehrlich [4] and Forst [5], to allow for several possibilities. Our particular approach is to focus on execution patterns over the five most recent calendar years. We divide that period into two equal parts (actually a 3-2 split because of data limitations), and assume $P_e$ is given by an expression of the form:

$$P_e = Wr_1 + (1 - W)r_2$$

where

$r_1$ = ratio of executions for murder in the last three years to non-negligent homicides in the three-year period preceding last year (i.e., we allow a one-year lag between murders and executions).

$r_2$ is defined analogously for the two calendar years before the last three.

W is a parameter of the model allowed to vary from .6 to 1; it is a damping factor



allowing some crude approximation of the diminshing influence of receding years. When W = 1 only the last three years are considered; when W = .6, $P_e$ is essentially the average execution rate for the past five years. This definition of $P_e$, with its one-year lag between homicides and executions, avoids two pitfalls noted by earlier scholars: (1) a spurious negative correlation between execution and murder levels based solely on chance fluctuations and (2) a misleading positive correlation because higher murder levels, by stimulating public fear, induce more executions. While people rarely know actual $P_e$ values, we are hoping that their perceptions of personal risk are monotonically related to them.*

Like earlier researchers, we are prevented by data limitations from calculating $P_e$ for capital homicides (i.e., those subject to the death penalty) rather than all homicides. But we suspect that if researchers with access to large data bases find the distinction intractable, it was no less so for potential murders. A more serious limitation on the analysis is that, under any reasonable definition, $P_e$ almost never exceeded .04 in the U.S. experience. Thus at most we can assess whether rare executions have deterrent power compared to none; we cannot provide useful information on what might happen if, in the future, high $P_e$ values arose under mandatory-penalty laws.

We will explore various $f(P_G,P_e,L)$'s of the general form:

$$f(P_G,P_e,L) = 1 - \alpha P_e{}^a - \beta(P_G{}^\dagger)^b - \gamma L^d \qquad (3)$$

where $P_G{}^\dagger$ is chosen as either $P_G$ or $1 - P_G$. We consider only nonnegative values of a and b, and $f(P_G,P_e,L)$. The additive form of (3) arises from the premise that there is probably a deterrent effect associated with conviction per se; the impact may grow as incarceration, the likely consequence of conviction, tends to get longer, and may grow further as the ultimate consequence -- execution -- gets more common.

Through varying α, β, γ, a, b, and d, we can examine a wide range of possibilities. A negative value of α implies that executions stimulate murders. (Several reasons this may happen have been suggested, including the greater urgency of eliminating witnesses and an allegedly brutalizing effect of state-sponsored murders.) When α = 0, executions are assumed to achieve no deterrence in themselves. Perhaps an occasional execution just to show that $P_e > 0$ has almost as much deterrence value as far more frequent executions; choosing "a" barely above 0 and α > 0 allows consideration of this

---

*Because murder levels were fairly stable in the period studied, $P_e$ values were nearly proportional to the number of annual executions, a quantity that might well be accurately perceived.

hypothesis. Perhaps, instead, for $P_e$'s below a certain threshold, criminals perceive their risks as so minimal as to be effectively zero; this possibility can be examined by choosing $\alpha > 0$ and large values of $a$. The choice $a = 1$ allows the standard linear relationship. Since $P_e$ values are small, functions of the form $\alpha P_e{}^a$ can fairly well approximate many more elaborate functions.

It is not a priori clear whether it is the probability of punishment or the probability of not being punished that enters the thinking of a potential murderer; that seems to depend on whether he is an optimist or a pessimist. Thus we allow both $P_G$ and $1 - P_G$ to appear as conviction-risk indicators. We do not, however, use $1 - P_e$ as well as $P_e$, for $1 - P_e$ varies very little from 1. Technically it is inappropriate to consider $L$ separately from $P_G$ (if $P_G = 0$, who would care about increases in L?) but we doubt this is a practical difficulty.

While we have several parameters to vary, this should not obscure the fact that we are incorporating punishment into our models in a very particular way. If this general approach is wrong, the specific results of using it could be highly misleading. While this problem afflicts all mathematical models of this kind, it creates a genuine need for external "standards of accountability" to indicate whether a model is accurate enough for its stated purposes. Such standards will be described and applied later in the paper.

## IV. Time-Trends

So far, we have not considered how economic conditions, the availability of guns, the prevailing age distribution and the quality of emergency medical care affect homicide levels. Instinctively, one thinks that such factors vary far more between different regions than within a given region. One hears, for example, of the economic stagnation of the North or the ubiquity of guns in the South, but one rarely gets the impression that conditions differ greatly in, say, Pennsylvania and New York. Because of both this and the lack of reliable data on several of these factors, we assume that, between two separate years, the cumulative effect on murder levels of changes on these dimensions can be approximated by a regional time-trend factor, $C_T$. The overall murder-growth ratio $C$ is assumed proportional to $C_T$ as well as $C_P$.

In calculating regional time-trend factors, we certainly do not wish inadvertantly to include, and thus to "weed out," the effects of changing punishment patterns. We describe in the next section a procedure to estimate the $C_T$'s aimed specifically at avoiding this problem.

## V. Summary and Example

To summarize the last four sections, we aim to examine the change in a given state's murder levels between two different years. Data analysis led us to approximate $C$, the ratio of murders in the second year to those in the first, by the relationship $C = C_D(\alpha_2/\alpha_1)$, where $C_D$, $\alpha_1$, and $\alpha_2$ were defined in Section II. We subsequently assumed that $\alpha_2/\alpha_1$, the demographically adjusted ratio of murder rates in the two years, is given by: $\alpha_2/\alpha_1 = C_P C_T$, where $C_P$, defined in Section III, reflects the effect of changing punishment patterns and $C_T$ estimates the effect of trends unrelated to punishment. Both $C_D$ and $C_T$ are treated as regionally-dependent, $C_P$ is not. We will consider many different hypotheses about the influence of punishment on the number of killings; they in turn generate differing $C_P$ values. We will first attempt through data analysis to assess the soundness of the framework we have developed. If the results are encouraging, we will then look seriously at which hypotheses about punishment seem most consistent with the data examined.

This approach, and in particular the procedure to estimate $C_T$, are perhaps made clearer by an illustrative example, based on data from Louisiana from the years 1950 and 1960. The data relevant to the example are summarized below:

| Louisiana | 1950 | 1960 |
|---|---|---|
| $P_1$ | .46 | .52 |
| $P_2$ | .233 | .207 |
| $P_3$ | .21 | .169 |
| $P_4$ | .097 | .104 |

| Louisiana | 1950 | 1960 |
|---|---|---|
| Population (1000's) | 2684 | 3257 |
| No. of Homicides | 236 | 259 |
| $P_G$ | .38 | .46 |

Suppose the hypothesis under investigation is that $f(P_e,P_G,L) = 1 - .2P_G$, under which executions are irrelevant to murder levels while a 5% rise in the conviction rate causes roughly a 1% drop in the number of killings. The population growth factor over the period is 3257/2684 = 1.213 ; using this figure, the given $P_1 - P_4$ values, and the appropriate $B_S$ and $U_S$ values from Table 1 (e.g., $U_S = 6.28$ for 1960), we obtain $C_D = .909$ . $C_P$ for the given punishment function is .983 , a reflection of the salubrious impact expected from the 8% increase in conviction rates over the period. Multiplying these quantities yields a growth factor C of $1.084C_T$ for Louisiana's number of murders. But how do we estimate $C_T$ , the time-trend factor from 1950 to 1960 for the Southern region?

Since Louisiana had 236 homicides in 1950, the "expected" number in 1960 would be $236(1.084C_T) = 255.8C_T$ under the model assumed. One can perform similar calculations for the 11 other Southern states; they lead to the estimate that, neglecting fluctuations both in 1950 and 1960, the <u>total</u> number of killings in the South in the latter year would be $(4325.8)C_T$ . The actual 1960 total for the South was 3814. Choosing $C_T = .882$ would achieve exact agreement for the whole region between predicted and expected murder levels, though not, we stress, for the individual states. It is under this regional criterion that we will estimate $C_T$'s . The factors thus obtained are essentially the maximum-likelihood estimates for $C_T$ in each region under a random fluctuations model we will soon introduce.

Observe that we estimated the time-trend factor only <u>after</u> correcting through $C_P$ for the hypothesized effect of changing punishment levels. In this way we isolate trends in punishment of vital interest to us from more general trends.

For the state of Louisiana, the model predicts a 4.4% drop in the number of murders from 1950 to 1960 (1.064 x .882 = .956); in fact, the number of killings rose 9.7%.



Is this prediction error reflective of large systematic errors in this particular model? We consider this question and its more general counterpart in Section VII; obviously, the answer should affect our assessment of the model's utility.

## VI. <u>Evaluating Homicide Models: Some General Considerations</u>

Recorded numbers of killings are not perfect reflections of prevailing levels of "murderousness," both because of (1) police errors in identifying and recording homicides and (2) differences between the <u>intended</u> number of killings during a period and the actual number. (This latter quantity includes both unsuccessful attempts at murder and assults not intended to be lethal that nonetheless leave their victims dead.) These two factors are presumably unrelated to the systematic influences on murder discussed earlier; we refer to their combined effect as the <u>random component</u> of the recorded homicide level. In [1] we argued that, because of such randomness, a recorded annual murder total should be regarded as one sample from a normal distribution with some mean $\lambda$ and variance $1.04\lambda$. $\lambda$, we suggested, is in some sense the "true" murder level for the period, devoid of the mischievous effects of sheer chance.

Our interest in such fluctuations stems from the realization that, because of their existence, no mathematical model of the systematic effects on murder levels could be expected to have perfect predictive power. In general, however, only a fraction of the discrepancy between predicted and actual levels of killing can be accounted for by such fluctuations; the remainder is caused by imperfections in the model itself. In [1], we considered how to estimate how much of a model's prediction error cannot reasonably be attributed to random effects. The procedure, summarized below, provides useful information about the accuracy of the model's key assumptions.

Consider a particular model for estimating a locality's murder level in terms of its punishment patterns and social and economic conditions. Suppose that one uses the model to predict the number of killings in N different situations, and then obtains the corresponding actual numbers. Associated with the $i^{th}$ prediction is the normalized residual $r_i$ that follows:

$$r_i = \frac{w_i - x_i}{1.02\sqrt{x_i}}$$

where

$w_i$ = predicted murder level in $i^{th}$ situation

$x_i$ = actual murder level in $i^{th}$ situation.

With $r_i$'s in hand, one then calculates the quantity $\bar{K}$ given by:

$$\bar{K} = 0 \qquad \text{if } S \leq \sqrt{2/\pi}$$

$$2\bar{K}\,(.5 - R(\bar{K})) + \sqrt{2/\pi}(e^{-\bar{K}^2/2}) = w \qquad \text{if } S \geq \sqrt{2/\pi}$$

where

$$S = A\sum_i^N |r_i|/N\,, \qquad R(\bar{K}) = \frac{1}{\sqrt{2\pi}}\int_{\bar{K}}^{\infty} e^{-x^2/2}\,dx$$

and $A$ is a correction factor for those parameters in the model estimated from the data now being used to assess it. This quantity $\bar{K}$, we tried to show, is a useful measure of the model's level of systematic error. The average absolute nonrandom error $\bar{E}$ in the N predictions is roughly proportional to $\bar{K}$, the proportionality factor being $1.02(\sum_{i=1}^{N} \sqrt{x_i}\,)/N$. In Section VIII, we will present $\bar{K}$ values for various models arising from the framework described earlier, and interpret these values in light of our primary interest in the deterrent effect of capital punishment.

Proceeding further, it is important to discuss some of the hazards inherent in capital punishment research. Suppose that each hypothesis about the systematic influences on murder levels could be characterized uniquely by a point (x,y) in two-dimensional space, where x reflects the assumed effect of capital punishment and y the effect of other factors. (Clearly this supposition is not true, but it simplifies the discussion greatly without changing the general point.) Let the hypothesized effect of executions be monotonic increasing in x, with $x = 0$ corresponding to no effect at all. It is too much to hope that we will ever determine the hypothesis $(x_o, y_o)$ that is perfectly correct in a given setting. A more reasonable goal is to attempt through data analysis to identify the range of hypotheses that have sufficiently low systematic errors that their y-values warrant serious attention (i.e., those whose systematic errors are not clearly larger than the effect being investigated). This "interesting" group of hypotheses, identified perhaps by their low $\bar{K}$-values, might in a hypothetical situation cover Region I in Figure 1.



Any given model builder, who of necessity imposes restrictions on the variables included and functional forms considered, examines only a small region of the two-dimensional "hypothesis space." Broadly speaking, there are three possible relationships between the region he studies and the "interesting" Region I; they are suggested by Regions A, B, and C in Figure 1.

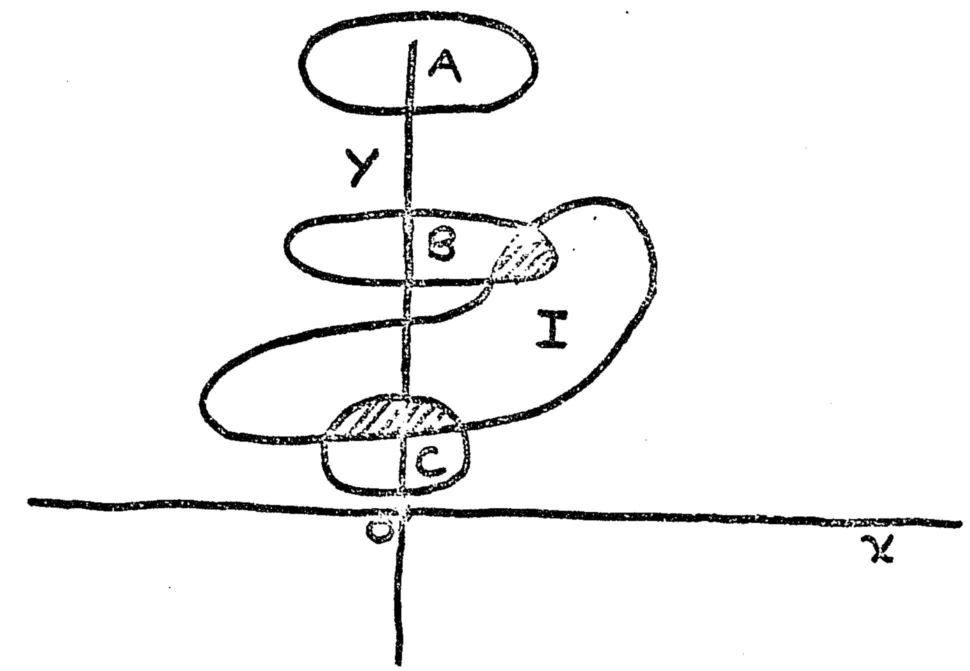


Figure 1

It is quite possible that all the models considered by a given researcher (e.g., linear in certain variables) have very large systematic errors (Region A). In that case, one should no more use the models to determine the effect of executions than use a watch with only an hour hand to estimate the time to the nearest second. Even the "best-fit" model among those considered is of little relevance. While it is probably not his fault, the researcher's contribution to the study of homicide patterns is primarily a negative one.

In the situation depicted by Region B, the researcher's explorations did uncover a subset of Region I (shaded area). But for the purposes of the study, this subset is not a representative one, for its x-values considerably exceed the Region I average. Those relying exclusively on this study might well get an exaggerated impression of the deterrent effect of capital punishment. In the Region C situation, the researcher has unearthed a representative subset.

Through such statistical measures as $\bar{K}$, one can get some idea whether a particular researcher's models fall into the unfortunate Region A category. (In practice, few capital punishment studies even make reference to this possibility.) But the "successful" researcher has no way of knowing whether he is in the Region B or Region C situation. The hope must be that ultimately, many different researchers, generating models from a variety of perspectives, will collectively define the contour of Region I. These last remarks do not imply that no individual death-penalty scholar can make substantive statements on the basis of his results. They do suggest, however, that there exists genuine limits to what he can say responsibly.

## VII. Region I Criteria

To assess the predictive power of various homicide models, we use data from 33 U.S. states that produce 92% of all national homicides (the 30 listed in Section II plus Arizona, Florida, and Virginia).

As indicated earlier, we consider various punishment functions (f's) and, as in the Louisiana example, proceed for each of the 33 states to calculate predicted changes in murder levels over the two periods 1950-1960 and 1960-1965. We then compare these predicted changes with those actually recorded, and calculate $\bar{K}$-values to reflect the accuracy of the individual models used. We also examine data for the period 1953-1958 and 1958-1963. However, we were unable to get reliable information on $P_G$ and L values for these years, and thus restricted our attention to those models whose punishment functions were based solely on use of the death penalty (i.e., $f(P_e,P_G,L) = 1 - \alpha P_e^{\ a}$).



Before reviewing charts of $\bar{K}$ values in the next section, we should make clear how to interpret them. As already noted, the $\bar{K}$-value for a model is roughly proportional to $\bar{E}$ , the average absolute systematic error of its predictions; for the data used, $\bar{E}$ is approximately 20.5K . (The 20.5 arises because over the period studied, a typical state in the group of 33 had about 200 homicides per year; see [1].) Thus, for example, if $\bar{K} = .8$ for a given model, its prediction errors caused by its imperfections averaged about 16 homicides apiece. This systematic error, we stress again, is distinct from the random prediction errors that no homicide model can avoid.

We discussed in the last section Region I hypotheses, the objects of search of those doing death-penalty data analysis. With its $\bar{K}$-value in hand, how does one determine whether a particular hypothesis has Region I status? Consider the following tentative criterion:

> A homicide model should be included in Region I if its observed systematic error is smaller than the effect it attributes to capital punishment.

The rationale for this criterion is suggested by an example. Consider a model $M_o$ to predict changes in state murder levels between 1950 and 1960; suppose that, under $M_o$ , reduced deterrence caused by lower state $P_e$ values should lead to 600 more U.S. killings in 1960. Suppose further that the estimated total systematic error in $M_o$'s various state predictions (given by $N\bar{E}$ , where N predictions are made) is 300. In this case, even if one believed that ALL the systematic error resulted from overstating the impact of executions, and that 300 should therefore be subtracted from the hypothesized effect, a net deterrent effect of 300 would remain. In other words, the size of $M_o$'s observed systematic error is not large enough to raise doubts <u>in itself</u> about the <u>sign</u> of the effect attributed to executions.* $M_o$ would seem under these circumstances to warrant entry into Region I.

This proposed criterion for Region I, while neutral on its faces, suffers two

*This is not to say that doubts could not arise on other grounds; this issue is discussed at the end of Section VIII.

serious problems. It is rather harsh on models that attribute only small effects to the death penalty; indeed, those that assume executions do not affect homicide levels need zero systematic error to meet the standard. In other situations, however, the criterion seems too lenient, as the example below might suggest.

Consider a model $M_1$, under which each execution has twice the deterrent effect implied by $M_o$ discussed above (e.g., if $f(P_e,P_G,L) = 1 - \alpha P_e^{\ a}$ under $M_o$, then $f(P_e,P_G,L) = 1 - 2\alpha P_e^{\ a}$ under $M_1$). Suppose that $M_1$, which would predict roughly 1200 extra homicides in 1960 because of lower $P_e$ values, sustains a systematic error about 900. The entire hypothesized rise in deterrence in $M_1$ over $M_o$ (1200-600) has shown up as additional systematic error (900-300), which suggests that $M_1$ has strongly exaggerated the effect of capital sanctions. One would probably not want $M_1$ in Region I even though it meets the stated criterion.

In the Appendix we describe modified criteria for Region I to deal with these two difficulties. The admissions standard finally becomes: a model is included in Region I if and only if its $\bar{K}$-value falls below some threshold Q; the precise value of Q is obtained from considerations discussed in the Appendix.

An important qualifying remark should be made now: a model's inclusion into Region I does NOT guarantee that its assessment of the sign of the effect of executions must be correct. It is possible that even a model with $\bar{K} = 0$ is wrong in this regard, because it attributes to changing execution rates an effect that properly belongs to another variable that, over time, has moved collinearly with $P_e$. Region I models, unlike their counterparts outside the region, warrant serious attention because of their high consistency with actual data patterns; serious attention, however, is a far cry from unqualified acceptance.

## VIII. Results

K-values for a subset of the models we studied are presented in Table 3. These values have been adjusted upward to correct for data-based parameter estimates that



would tend artificially to reduce $\bar{K}$. A "degree of freedom" was also subtracted for each parameter in a model's punishment function (e.g., 2 for $f(P_e,P_G,L) = 1 - \beta P_G^{\ b}$). This latter subtraction, though technically required only for "best fit" parameter values, facilitates direct comparisons of the power of different models.

A review of Table 3 suggests that the variables $P_G$ and L add virtually no explanatory power to models of homicide patterns. If this inference is correct, it might well indicate not that potential killers are indifferent to conviction rates or prison sentences, but rather that they are insensitive to the limited changes in state $P_G$ and L values over the period studied. The values of $P_e$ do seem to have some influence on $\bar{K}$-values for the years 1950-60-65: models that assume some deterrent effect of executions achieve a modest reduction in $\bar{K}$ values compared to those that assume no such effect. For the years 1953-58-63, when changes in state $P_e$ values were lower, no really suggestive pattern emerges. It is perhaps of interest, however, that the "best fit" model for these years does assume some deterrent effect ($\alpha = 1.2$, $a = .7$).

The $\bar{K}$ values in Table 3 are up to 60% lower than the lowest obtained in [1] for the homicide models presented in several other recent death penalty studies. (Their $\bar{K}$ values ranged from 1.58 to 3.07.) But this fact, as well as the general comments in the last paragraph, are of little interest unless some of our models are accurate enough to deserve Region I status (i.e., we are not in the Region A situation of Section VI). In deciding whether to include a model in Region I, we examine all 55 state predictions for 1950-1960 and 1960-1965 in which nonzero changes in $P_e$ values made relevant the values assigned to "execution parameters" $\alpha$ and a. Applying in this context the criteria of the Appendix, we obtain the standard: A model should be included in Region I if and only if its $\bar{K} \leq .88$*. This means that we uncovered models with $\bar{K}$'s up to .88 whose systematic errors fell below the effects they attributed to capital punishment, and which

*We use only 1950-60-65 results in setting Region I criteria because, as noted earlier, we have no data on $P_G$ and L for 1953-58-63, and want all models to be subject to the same standard. The 1953-58-63 results should be viewed as supplementary.

Table 3: $\bar{K}$-VALUES FOR A SERIES OF HOMICIDE MODELS

The punishment functions used generally are of the form

$$f(P_e, P_G, L) = 1 - P_e^{\,a} - P_G^{\,b} - L^c$$

and are summarized by the vector (α, a, β, b, γ, c). Note that positive (negative) values of α, β, γ imply a deterrent (counterdeterrent) effect.

MODELS THAT ASSUME THE IRRELEVANCE OF CAPITAL PUNISHMENT

| Punishment Function | $\bar{K}$ |
|---|---|
| (0,0,0,0,0,0) | .81(.62)* |
| (0,0,.22,.28,0,0)** | .82 |
| (0,0,.08,.19,0,0)**† | .83 |
| (0,0,0,0,360,.012)** | .82 |
| (0,0,.34,.28,-.4,.02)** | .82 |

MODELS THAT ASSUME AN IMPACT OF CAPITAL PUNISHMENT

| Punishment Function | $\bar{K}$ when "memory factor" w in definition of $P_e$ is: .7 | .8 | .9 |
|---|---|---|---|
| (4,1,0,0,0,0) | .77(.61)* | .76(.63) | .76(.67) |
| (-4,1,0,0,0,0) | .90(.69) | .91(.69) | .92(.70) |
| (8,1,0,0,0,0) | .77(.72) | .75(.77) | .77(.82) |
| (-8,1,0,0,0,0) | .97(.80) | .99(.81) | 1.01(.82) |
| (20,1,0,0,0,0) | 1.39(1.27) | 1.40(1.40) | 1.47(1.57) |
| (-20,1,0,0,0,0) | 1.26(1.11) | 1.29(1.13) | 1.35(1.15) |
| (400,2,0,0,0,0) | .87(.77) | .85(.83) | .84(.91) |
| (.4,.5,0,0,0,0) | .81(.62) | .80(.62) | .80(.63) |
| (.04,.01,0,0,0,0) | .86(.63) | .86(.63) | .86(.63) |

Linear Model with Minimum $\bar{K}$:
1950-60-65: (7.2,1,0,0,0,0) with w = .78 ($\bar{K}$ = .75)
1953-58-63: (3.7,1,0,0,0,0) with w = .6 ($\bar{K}$ = .61)
Overall Model with Lowest $\bar{K}$:***
1950-60-65: (35,1.4,0,0,0,0) with w = .78 ($\bar{K}$ = .74)
1953-58-63: (1.2,.7,0,0,0,0) with w = .6 ($\bar{K}$ = .61)

*$\bar{K}$ values in parenthesis are for 1953-58-63; all others are for 1950-60-65. Recall that $P_G$ and L values were not available for 1953-58-63.

**Lowest $\bar{K}$ for punishment function of form considered, in which all 0's shown were constrained to take those values. Thus, for example, "best" model of form $f(P_e,P_G,L) - 1 - \beta P_e^{\,b}$ used β = .22, b = .28. This model has a higher $\bar{K}$ than (0,0,0,0,0,0); this means that having the two parameters β and b to vary yielded so little reduction in prediction errors that, when "degrees of freedom" were subtracted for them, they hadn't improved predictions at all.

†$1 - P_G$ rather than $P_G$ used as conviction-risk variable.

***We did consider models with α, β, γ simultaneously nonzero; use of nonzero β and γ never caused serious improvement.



had not clearly overstated the effects of executions.

The fact that we have identified some Region I hypotheses means that, in one crucial respect, our effort at model building has been successful. The next task is to characterize the Region I models we have obtained. Those models in Table 3 with $\bar{K} \leq .88$ are representative of the complete Region I subset we found. The subset is composed predominantly of models that attribute a noticeable but not gigantic deterrent effect to capital punishment. But -- and it is a crucial but -- certain models assuming no deterrence at all (e.g., $f(P_e,P_G,L) = 1$ for all $P_e$, $P_G$, L experienced) are also present in Region I.

In their assessments of the effects of executions, the Region I models differ strikingly. Their curious coexistence within Region I probably has something to do with the historical rarity of U.S. executions compared to U.S. homicides. Just as it is hard to tell through four tosses whether a coin is fair, it is hard with very low $P_e$ values to make strong statements about α and a . It is unsurprising that, when executions are rare, there is considerable ambiguity about the effect of any one. What is surprising is that, despite this ambiguity, many scholars have taken strong positions on the existence or nonexistence of a deterrent effect.

Opponents of the death penalty often state that there is "not one shred of evidence" that executions can deter murders. This viewpoint is not supported by the results here in which, in two completely separate periods, the models most consistent with the data assumed not only a deterrent effect but a fairly substantial one. On the other hand, given the full range of observed Region I models, it would distort these results to say that we have proved that "the death penalty works." Perhaps the best synthesis of the results is the statement that they lean slightly but not decisively towards the presence of some deterrent effect. Even this remark is subject to a caveat expressed earlier: because we considered only a limited set of models, it is possible that the Region I models we uncovered are not representative of the whole region.

There are two reasons why an attempt at data analysis can fail to yield decisive information on a question of interest: (1) its underlying model is not accurate enough to provide useful information and/or (2) not enough data are available to discriminate clearly among different hypotheses. These two circumstances have quite different implications on how to proceed next; in the first case, one should work at improving the model; in the second, one should search for more data. In our own situation, we have tried to suggest, it is the second rather than the first problem that prevents unequivocal statements on the effects of executions.

It is not clear how one could get additional data to reduce the difficulties we have encountered, for low $P_e$ values are far from limited to the years we studied, or even to the United States. Indeed, the changes in state $P_e$ levels over the period we explored were the largest recorded in the last 50 years. Having chosen certain data that seemed particularly likely to be revealing, we cannot be especially hopeful about the discriminatory power of other data.

## IX. Conclusion

It seems fair to say that those who, on other grounds, have strong views on the appropriateness of capital punishment need not abandon them because of conclusive evidence on the deterrence question. In a certain sense, our results actually might be viewed as decisive, for they suggest strongly that those who hope through analyzing past data to resolve this controversial question once and for all are, quite probably, harboring an illusion.

Our particular approach to constructing models, geared to the idiosyncracies of homicide patterns, is probably not directly useful in many other settings. But a central theme of this paper -- that mathematical models of a process should be subject to clear standards of accuracy tied to their specific purposes -- does seem relevant in a wider context, as are, perhaps, some of the specific procedures we developed pursuant to that theme. This author hopes very much that other researchers will agree on the importance of further work in this area. Mathematical modeling in the social sciences will not achieve full credibility until its practitioners are willing to assess and admit the limitations of their models.



## References


1. A. Barnett, "The Deterrent Effect of Capital Punishment: A Test of Some Recent Studies,"

2. W. Bowers and G. Pierce, "The Illusion of Deterrence in Issac Ehrlich's Research on Capital Punishment," Yale Law Journal, 85, p. 187 (1975).

3. I. Ehrlich, "Deterrence: Evidence and Inference," Yale Law Journal, 85, p. 209 (1975).

4. I. Ehrlich, "Capital Punishment and Deterrence: Some Further Thoughts and Additional Evidence," Journal of Political Economy, 85, p. 741 (1977).

5. B. Forst, "The Detrrent Effect of Capital Punishment: A Cross-State Analysis of the 1960's," Minnesota Law Review, 61, p. 743 (1977).

## APPENDIX

### Region I Criteria: Some Details

Region I models are those whose systematic errors do not "undercut" their assessments of the sign and magnitude of the effect of executions. This does not mean a Region I model must be correct, but it does mean that it might be. We describe below when we believe a model deserves Region I status.

For simplicity we restrict our attention here to punishment functions of the form $f(P_e, P_G, L) = 1 - \alpha P_e^{a}$; the general case is treated similarly. Suppose that, within this paper's framework, such a punishment function is used in predicting changes in murder levels for various states and periods. Let $E(\alpha, a)$ be the $\bar{K}$-based estimate of the total systematic error over all predictions made, and let $I(\alpha, a)$ be the total absolute extent to which changes in state $P_e$-values affected the predictions (i.e., the total effect attributed to capital punishment). To estimate $I(\alpha, a)$, we proceed for each state and period to calculate in two ways the predicted change in murder levels: (1) using actual $P_e$ values and f( ), and (2) setting $C_P = 1$, as if $P_e$ were unchanged over the period. The absolute difference between these two predictions is roughly the hypothesized net effect of the change in execution rates on the number of killings. $I(\alpha, a)$ is simply the sum of these absolute differences.

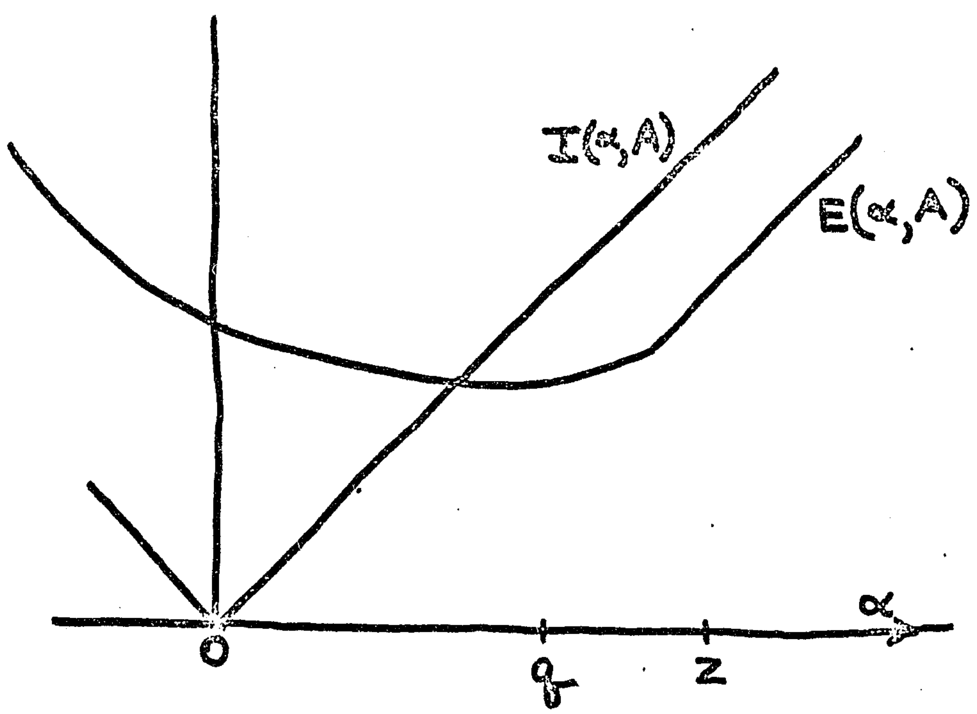


Figure 2



Suppose that the particular punishment function $f(P_e, P_G L) = 1 - qP_e^{A}$ is perfectly correct. Then we would expect the graph of $E(\alpha, A)$ as $\alpha$ is varied to take a shape like that depicted in Figure 2. While $E(\alpha, A)$ takes its minimum at $\alpha = q$, this minimum is probably not zero because of systematic errors in other parts of the model (e.g., time trend corrections). As $\alpha$ is increased above q, $E(\alpha, A)$ also increases, although initially not as fast as $I(\alpha, A)$. The reason for this is that although the increases in $\alpha$ directly introduce systematic errors into the predictions, these errors are probably opposite in sign in some states to other nonrandom errors already present. Only when $\alpha$ is so large that such other errors have been completely cancelled (i.e., past Z in Figure 2) does $\partial E/\partial\alpha$ approach $\partial I/\partial\alpha$.

In practice, a data-based graph of $E(\alpha, A)$ vs. $\alpha$ would probably not take its minimum at exactly $\alpha = q$, much as a maximum likelihood estimate of a parameter is rarely precisely right. But it still seems reasonable that once $(\partial E/\partial\alpha)/(\partial I/\partial\alpha)$ has exceeded a certain value, $\alpha$ is probably somewhat beyond its optimal value if the exponent A is accurate.

These comments and the discussion in Section VII lead us to the following formulation: a model with punishment function $f(P_e, P_G, L) = 1 - P_e^{a}$ predicts well enough to warrant inclusion in Region I if and only if either

(1) $I(\alpha, a) > E(\alpha, a)$ and $(\partial E/\partial\alpha)/(\partial I/\partial\alpha)$ at $(\alpha, a) <$ some specified $\lambda$,

OR (2) The model's $\bar{K}$ value does not exceed that of another model that entered Region I under (1).

The second criterion, pursuant to Section VII, is meant to avoid discrimination against hypotheses that attribute only small effects to executions. Once $\lambda$ is specified, the first criterion implies a maximum permissible $\bar{K}$-value for entry into Region I. In our calculations we use the threshold $\lambda = \frac{1}{2}$, largely because of our empirical observation that $(\partial E/\partial\alpha)/(\partial I/\partial\alpha)$ grows fairly rapidly beyond the value of $\alpha$ that minimizes $E(\alpha, a)$ for a given 'a'. While this choice of $\lambda$ is a bit arbitrary, the general results of this paper are unaffected by fairly large changes either way.

END