Journal of Criminal Law and Criminology

Volume 74
Issue 3 *Fall*

Article 8

Fall 1983

Capital Punishment and Deterrence: Conflicting Evidence

Brian Forst

0091-4169/83/7403-927
THE JOURNAL OF CRIMINAL LAW & CRIMINOLOGY Vol. 74, No. 3
 Printed in U.S.A.

CAPITAL PUNISHMENT AND DETERRENCE: CONFLICTING EVIDENCE?

BRIAN FORST*

I. INTRODUCTION

Until 1975, belief in the deterrent value of capital punishment had no authoritative empirical support. The studies widely cited prior to that time, most notably those of Sellin[1] and Schuessler,[2] had in fact found evidence consistent with the theory that the death penalty has no deterrent effect. Then Isaac Ehrlich reported results of a complex econometric analysis of aggregate United States times-series data, results indicating that capital punishment is an effective deterrent.[3] While Ehrlich's study has received a considerable amount of criticism,[4] it has also received some support,[5] and was even cited by the United States Solicitor General in briefs to the Supreme Court supporting the death penalty.[6]

With the publication of his second major empirical study on the

* Director of Research, INSLAW, Inc., Washington, D.C.; M.B.A. University of California at Los Angeles, 1965; B.S. University of California at Los Angeles, 1964.

The author is grateful to Professors Anthony Amsterdam and Hans Zeisel for their helpful comments on drafts of this Article. The author alone is responsible for factual errors or lapses in logic that may remain.

1 T. SELLIN, THE DEATH PENALTY (1959).

2 Schuessler, *The Deterrent Influence of the Death Penalty*, 284 ANNALS 54 (1952).

3 Ehrlich, *The Deterrent Effect of Capital Punishment: A Question of Life and Death*, 65 AM. ECON. REV. 397 (1975).

4 Bowers & Pierce, *The Illusion of Deterrence in Isaac Ehrlich's Research on Capital Punishment*, 85 YALE L.J. 187 (1975); Glaser, *Capital Punishment—Deterrent or Stimulus to Murder? Our Unexamined Deaths and Penalties*, 10 U. TOL. L. REV. 317 (1978); Klein, Forst & Filatov, *The Deterrent Effect of Capital Punishment: An Assessment of the Estimates*, in DETERRENCE AND INCAPACITATION: ESTIMATING THE EFFECTS OF CRIMINAL SANCTIONS ON CRIME RATES 336 (A. Blumstein, J. Cohen & D. Nagin eds. 1978) [hereinafter Klein, *The Deterrent Effect*]; Passell & Taylor, *The Deterrent Effect of Capital Punishment: Another View*, 67 AM. SOC. REV. 445 (1977).

5 Walter Berns made the following observation on Ehrlich's time-series analysis: "His study is not garbage." W. BERNS, FOR CAPITAL PUNISHMENT: CRIME AND THE MORALITY OF THE DEATH PENALTY 98 (1979).

6 Brief for the United States as Amicus Curiae at 35-38, Fowler v. North Carolina, 428 U.S. 904 (1976); Brief for the United States as Amicus Curiae at 34-45, 9a-16a, Gregg v. Georgia, 428 U.S. 153 (1976).

deterrent effect of capital punishment in 1977, Ehrlich produced further results in support of the theory that executions deter crime in general and homicides in particular.[7] The persuasiveness of his 1977 results, based on cross-state data for 1940 and 1950, was heightened by their general consistency with his earlier results and with his theoretical predictions.[8] The credibility of these estimates was further enhanced by the fact that they were shown to hold up under various subsets of observations, several alternative measures of the key independent variable (the execution rate), alternative systems of weighting, the generalized and ordinary least-squares regression methods, the incorporation of various binary control variables, and the use of both FBI and vital statistics homicide data. In Ehrlich's words, "the consistency and stability in the results obtained upon application of efficient estimation procedures, given the sample limitations, and their consistency with specific theorectical predictions and previous findings seem remarkable."[9]

No less remarkable is the fact that Ehrlich's more recent findings are not consistent with regression estimates of the deterrent effect of the death penalty reported elsewhere, estimates also based on cross-state data. It is surely in order to reconsider the cross-sectional results in light of these apparent conflicts. Section II focuses on those conflicts.

Consideration of earlier studies in Section II provides support for the theory that deterrent effects can be more efficiently estimated from observations of temporal differences in the relevant variables in cross-sectional data. In Section III, this procedure is applied to data from the cross-section of states for 1960 and 1970, and the results are discussed and compared with results of earlier analyses. The Article concludes in Section IV with a discussion of the implications of the available evidence for the general theory of deterrence and for the applicability of that theory to the death penalty.

II. THE CONFLICTING EVIDENCE: WHAT LIES BENEATH IT?

Six months after the publication of Ehrlich's analysis of aggregate time-series data for the United States, an article by Peter Passell appeared describing an analysis of data for the cross-section of states for 1950 and 1960.[10] The strengths of the cross-sectional approach to esti-

---

[7] Ehrlich, *Capital Punishment and Deterrence: Some Further Thoughts and Additional Evidence*, 85 J. POL. ECON. 741 (1977).

[8] Foremost among these predictions is that the elasticity of the murder rate with respect to the murder conviction rate will exceed, in absolute value, the elasticity of the murder rate with respect to the conditional execution rate. *Id.* at 744, 771.

[9] Ehrlich, *supra* note 7, at 778.

[10] Passell, *The Deterrence Effect of the Death Penalty: A Statistical Test*, 28 STAN. L. REV. 61 (1975).

mating deterrent effects have been widely recognized and will not be repeated here.[11] Passell's results for 1950 were based on an analysis of data for the forty-one states for which data were available; his findings for 1960 were based on data for forty-four such states. Passell reported results under both the ordinary and the two-stage least-squares estimating techniques, both with and without a binary Southern variable, using both the linear and loglinear models, and using various constructions of the execution rate variable. He concluded with the statement: "We know of no reasonable way of interpreting the cross-section data that would lend support to the deterrence hypothesis."[12]

While there are important differences in the approaches employed by Passell and Ehrlich,[13] the similarities are extraordinary: both Passell and Ehrlich examined data for cross-sections of states for 1950; both used an approximation of the loglinear regression model; both used unweighted regressions; both used FBI homicide data and National Prisoner Statistics data on executions and murder convictions; and both formed the execution rate variable alternatively as the coincident ratio of executions to convictions, as a lagged ratio, and as a ratio of the four-year mean number of executions to the number of convictions. In addition, both incorporated the following as control variables: median time served in prison by convicted murderers; percent of the resident population between the ages of fifteen and twenty-four; percent of families in poverty; and a race variable. Given these similarities, it can only be regarded as astounding that Passell finds no evidence that the death penalty deters homicide while Ehrlich's results suggest that it does deter.[14]

---

[11] *See id.* at 62; EHRLICH, *supra* note 7, at 742; Forst, *The Deterrent Effect of Capital Punishment: A Cross-State Analysis of the 1960's*, 61 MINN. L. REV. 745-47 (1977); Klein, *supra* note 4, at 341-42.

[12] Passell, *supra* note 10, at 80.

[13] Two fundamental differences are discussed in detail below. *See infra* notes 15-29 and accompanying text. These two differences relate to Passell's and Ehrlich's assumptions about the specification of the homicide function and their willingness to accept results for the 1960 cross-section of states. Other differences also may have contributed to the divergence in Ehrlich's findings from those of Passell. Unlike Passell, Ehrlich incorporated variables reflecting median family income, proportion of residents living in urban places, and whether the state was an execution state (the use of which was not found to alter materially Ehrlich's deterrence finding). Further, whereas Passell used only unweighted regressions, Ehrlich used both weighted and unweighted regressions, again with similar results in each instance. On the other hand, Passell's results were reported both with and without a Southern dummy variable (with similar results in each case), whereas Ehrlich did not incorporate this variable.

The results of both Ehrlich's and Passell's robustness tests, together with results reported below, *see infra* Tables 1 and 2, suggest that the sharp contrast in their respective conclusions has less to do with these latter differences than with assumptions about the specification of the homicide function and the use of the 1960 cross-section.

[14] It is noteworthy that using unweighted regression with the 1950 cross-section (which comes as close to Passell's analytic structure as any employed by Ehrlich), Ehrlich's result for

A. USE OF THE 1960 CROSS SECTION

Equally astounding is the fact that Passell reports results for the 1960 cross-section of states, while Ehrlich does not. Ehrlich gives two reasons for not analyzing the 1960 cross-section of states. One reason is that the "number of executions per state in a given year" was too small in 1960—the "large percentage changes from year to year . . . amount to severe 'sampling errors.'"[15] His other reason for not analyzing the 1960 data is that the cross-state variation in the level of enforcement of capital punishment was insufficient to provide a meaningful test of the deterrence hypothesis.[16] Neither reason holds up under scrutiny.

First, it is irrelevant that the percentage change in the number of executions from year to year was larger in many states in 1960 than in 1940 or 1950, the years for which Ehrlich did report results. If a case were to be made against the use of the 1960 cross-section based on "sampling errors," it should be based on year-to-year *differences* in the execution *rate*, not percentage changes in the number of executions. Percentage changes in the number of executions are surely large, but neither the percentage changes nor the number of executions are very relevant to sampling error in the key independent variable—the execution rate. Moreover, the random part of the fluctuations in the execution rate that does exist, small as it is, has been virtually eliminated through the statistical smoothing procedures employed by Ehrlich and others.[17]

Ehrlich's second justification for not analyzing the 1960 cross-section data is that the level of enforcement of capital punishment did not vary enough from one state to the next in that year to allow a meaningful test of the deterrence hypothesis. It is, of course, not obvious what level of variation in an independent variable is enough to provide a meaningful test of a hypothesis. If the level of variation in the execution

---

the execution rate variable was among the least significant reported for this variable in his 1977 article (t = −1.38). Ehrlich, *supra* note 7, at 753. Under weighted regression, or using the 1940 cross-section of states, or both, he obtained t-statistics in the neighborhood of −3.5 to −5.0. *Id.* at 752-53.

15 Ehrlich, *supra* note 7, at 786.

16 *Id.* at 742, 748-49, 786.

17 The year-to-year fluctuations in the execution rate that do exist have been damped through the application of two-stage least-squares estimates, Forst, *supra* note 11, at 757-58; Passell, *supra* note 10, at 73-77, and with the use of the average number of executions over three- to five-year periods, Ehrlich, *supra* note 7, at 750-54; Forst, *supra* note 11, at 756-57; Passell, *supra* note 10, at 68. Even in the absence of these data smoothing procedures, however, one would expect the estimates of the ratio of executions to homicide convictions for states to be considerably less accurate for 1940 and 1950 than for more recent periods. The estimates of the number of state homicide convictions used by Ehrlich for 1940 and 1950 are based on murder convictions only, Ehrlich, *supra* note 7, at 774-75; the general superiority of the more recent data is discussed below. *See infra* note 20.

rate variable for which Ehrlich estimated deterrent effects in his time-series analysis can be accepted as sufficient, however, then the level for the 1960 cross-section of states should also be considered sufficient. The coefficient of variation in the execution rate is in fact *larger* for the 1960 cross-section of states than for Ehrlich's aggregate time series.[18] Under the circumstances, Professor Ehrlich's decision not to report results for the 1960 cross-section appears to represent another instance of a tendency, which has been noted earlier,[19] for Ehrlich to select analytic options that produce findings of deterrence when the case for options that produce different results is at least as strong.

Actually, cross-state data for 1960 and the decade that followed are especially worthy of analysis. First, the more recent data have been widely acknowledged to be more accurate than the earlier data.[20] Second, and perhaps most important, much of the criticism of Ehrlich's 1975 results centers on lack of robustness of those findings with respect to the inclusion of more recent data;[21] hence, it surely seems in order to test the deterrence hypothesis using cross-sectional data for this period in controversy. Finally, to the extent that the magnitude of the deterrent effect is itself a function of other factors,[22] the current deterrent effect of the death penalty may be estimated more reliably with data that reflect more recent conditions.

---

[18] I calculate the coefficient of variation equal to 1.54 for the execution rate, based on data for 33 states for which information about executions and convictions are available for 1960. It is 1.34 based on a three-year average of executions centered about 1960. For the aggregate time-series data, on the other hand, the coefficient of variation for the execution rate for the period 1933 to 1969 is 0.946, based on an independent calculation using Ehrlich's data sources. The source of the 1960 state data is the FEDERAL BUREAU OF PRISONS, NATIONAL PRISONER STATISTICS (1960). Ehrlich's data sources for his 1975 article are given in an unpublished memorandum prepared by Ehrlich in May 1975.

[19] *See* Klein, *supra* note 4, at 343-57.

[20] Ehrlich has indicated that, in general, the more recent data are superior. Ehrlich, *The Deterrent Effect of Criminal Law Enforcement*, 1 J. LEGAL STUD. 259, 272 (1972). He has also reported results for the 1960 cross-section in a study that found deterrent effects for noncapital sanctions. Ehrlich, *Participation in Illegitimate Activities: A Theoretical and Empirical Investigation*, 81 J. POL. ECON. 521 (1973) [hereinafter Ehrlich, *Participation in Illegitimate Activities*]. He has further pointed out that state data on the numbers of homicides and homicide convictions are, in particular, inferior for the 1940 and 1950 cross-section of states. Ehrlich, *supra* note 7, at 775, 779, 785. Whereas the number of homicides in each state for 1960 is based on total homicides reported by the police, the numbers for 1940 and 1950 are based on samples of police departments drawn only from urban areas. Klein has shown that such errors do not, in general, bias the coefficients of the sanction variables toward zero. *See* Klein, *supra* note 4, at 347-49.

[21] *See* Bowers & Pierce, *supra* note 4, at 197-204; Klein, *supra* note 4, at 353-55; Passell & Taylor, *supra* note 4, at 446-48.

[22] *See infra* text accompanying note 32.

B. ASSUMPTIONS ABOUT THE FORM OF THE HOMICIDE FUNCTION

A second major problem with estimates of the deterrent effect of capital punishment is their lack of robustness with respect to the mathematical specification of the homicide function. In particular, deterrence has been found to appear stronger when the homicide function is assumed to be loglinear than when it is assumed to be linear. This phenomenon has been found in analyses of both time-series[23] and cross-sectional data.[24] Ehrlich rejected the simple linear model based upon an application of the Box and Cox tests for optimal transformations to a class of power transformations,[25] concluding that the loglinear model "cannot be rejected as optimal."[26]

The execution rate variable, however, takes on values of zero in both of Ehrlich's empirical analyses of the deterrent effect of capital punishment. Since the logarithmic function is undefined at zero, one must reject the loglinear specification as a feasible model. Thus, it is safe to assert that the homicide function is, in fact, neither linear nor loglinear. Further, for that matter, we cannot be sure that it is even within the class of power transformations considered by Ehrlich, since he finds the logarithmic form (an inadmissible specification) to be "unambiguously . . . super[ior]" to all other power transformations considered.[27]

Assumptions about the precise form of the homicide function need not, however, stand in the way of accurate estimation of the average deterrent effect of the death penalty. Indeed, the deterrent effect is typically estimated as a constant trade-off number, as though the underlying function were linear, even when the homicide function is assumed to be nonlinear.[28] Given the lack of robustness of the previous findings with respect to the specification of the homicide function, an estimation procedure that is relatively insensitive to the precise form of this function would appear to be in order.

Instead of assuming that the homicide function is of a particular form, one need only assume that it is continuous and estimate the relevant parameters of the differential of the function. One advantage of

---

23 *See* Bowers & Pierce, *supra* note 4, at 199-203; Klein, *supra* note 4, at 356-57; Passell & Taylor, *supra* note 4, at 448-49.

24 *See* Ehrlich, *supra* note 7, at 758-59.

25 *Id.* at 781-83.

26 *Id.* at 760.

27 *Id.*

28 For example, Ehrlich has estimated an average trade-off of eight lives saved for each execution for the period 1933 to 1967 in the United States based on loglinear regression estimates. Ehrlich, *supra* note 3, at 398. Subsequently, he estimated an average trade-off in the neighborhood of 20 to 24 lives saved, based on the application of a loglinear model to state data for 1940 and 1950. Ehrlich, *supra* note 7, at 779.

this approach is that it exploits the fact that the differential of any continuous homicide rate function is always of the additive form, regardless of the particular form of the homicide function. This property is described more fully in the Appendix.

A second advantage of estimating the differential of the homicide function, rather than the homicide function itself, is that it focuses on *changes* in the variables of interest rather than on the levels of the variables. The question of deterrence deals primarily with the effect of change in a sanction variable on a crime rate measure. An estimation procedure based upon actual changes in the relevant variables in a cross-section of jurisdictions provides a more sensitive and direct test of the deterrence hypothesis than a procedure based on the levels of the variables in the cross-section. In addition, Klein has acknowledged that forming parameter estimates from the changes in variables in two successive cross-sections can eliminate biases due to spatial heterogeneity, biases that are common to estimates based on conventional analyses of cross-sectional data.[29]

III. Analysis of State Data for 1960 to 1970

The decade from 1960 to 1970 provides a particularly important opportunity to test the hypothesis that the death penalty deters homicide. For many states, this period represented an ending of capital punishment; for the other states, capital punishment had already been abolished, either de jure or de facto, prior to 1960. Under the theory that the death penalty deters homicide, one would predict that the homicide rate would have increased more in those states in which use of capital punishment ended during the 1960's than in the states in which the death penalty was not being used in the first place. One would further predict that the homicide rate would have increased most in the states with the greatest decline in the use of capital punishment—that is, those states with the highest execution rates in and around the beginning of the decade.

Since 1960 and 1970 are census years, with richer data available than in off-census years, one can examine the accuracy of these predictions by accounting for temporal changes in a host of other factors associated with the explosion in the homicide rate in the United States during that decade.[30] In addition, these two census years complement the census years selected in Ehrlich's 1977 cross-sectional study—1940 and 1950.

29 L. Klein, A Textbook of Econometrics 350, 358 (1974).

30 Homicides increased by more than 50% in the 1960's—from 5.1 homicides per 100,000 residents in 1960 to 7.8 in 1970. Forst, *supra* note 11, at 748.

Data for the 1960's are especially important, however, because they permit scrutiny of that crucial period on which previous time-series findings of a deterrent effect of capital punishment, derived from aggregate United States data, have been found to rely so heavily.[31] After decades of steady decline, the homicide rate rose sharply during the 1960's—the same decade in which use of the death penalty was stopped. It is surely in order to examine the 1960's more closely to establish whether the homicide rate increase was in fact related to the discontinuation of the death penalty.

To test the deterrent effect of the death penalty during the 1960's, one must analytically distinguish the deterrent effect of the death penalty from the other factors that significantly affected the homicide rate during that decade. The full set of variables considered plausible as homicide rate determinants (and for which data were available) include the following: the probability of execution given conviction,[32] Pe|c; the rate at which homicides result in conviction, C/Q; the average prison term served by persons convicted of homicide and released in the base year; proportion of families in poverty, Pov; proportion of nonwhites, NW; median family income, Y; the non-homicide crime rate, Cr; proportion of residents of the ages twenty-one to twenty-four; proportion of males; proportion of urban residents; school enrollment rate; resident population; divorce rate; proportion of adults employed; and a binary Southern variable.[33] The failure of changes in most of these variables to emerge as significant correlates of the increase in the homicide rate from 1960 to 1970 suggests the model.[34]

$$h = f\,(Pe|c,\ C/Q,\ Pov,\ NW,\ Y,\ Cr) \qquad (1)$$

The theory behind the inclusion of sanction variables in a crime equation (in this case Pe|c and C/Q) has been well-developed by Becker,[35] Ehrlich,[36] and others, and needs no further elaboration here. The income and race variables are included as statistical controls: while theories about their links with crime are both important and abundant, those theories lie outside the context of this paper. The rate at which

---

[31] *See* sources cited *supra* note 21 and accompanying text.

[32] Hereinafter this variable will be referred to as the execution rate.

[33] The data sources are given in Forst, *supra* note 11, at 765-67. Symbols are given in text only for those variables that were found to be significant.

[34] Of the six independent variables shown in Equation (1), all but Pe|c had t-statistics that exceeded 1.64 in absolute value (i.e., significant at .10) in the full regression equation. Pe/c is included here because it is the variable of primary interest. The full model was estimated with an ordinary least-squares regression analysis of the 33 states for which data were complete for both 1960 and 1970.

[35] Becker, *Crime and Punishment: An Economic Approach*, 76 J. POL. ECON. 169-217 (1968).

[36] *See* Ehrlich, *supra* note 3; Ehrlich, *Participation in Illegitimate Activities*, *supra* note 20.

TABLE 1

REGRESSION COEFFICIENTS (AND STANDARD ERRORS) FOR VARIABLES USED TO EXPLAIN STATE HOMICIDE RATE INCREASES, 1960 TO 1970

	R^2	N[a]	Intercept	Δ(E/C)	$\Delta(E_{+1}/C)$	$\Delta(\bar{E}/C)$	Δ°(E/C)[b]	Δ(E/Q)	Δ(C/Q)	Δ Pov	ΔNW	ΔY	ΔCr
a.	.566	33	−3.638	18.16					−5.384	−23.99	37.04	.0002556	.001736
			(2.00)	(8.51)					(1.56)	(7.44)	(13.0)	(.000474)	(.000448)
b.	.493	33	−4.746		5.006				−4.968	−26.46	38.45	.0005108	.001509
			(2.09)		(11.3)				(1.68)	(7.96)	(14.8)	(.000501)	(.000495)
c.	.514	33	−4.502			17.33			−5.014	−25.92	37.83	.0004534	.001636
			(2.05)			(15.3)			(1.64)	(7.79)	13.8)	(.000487)	(.000486)
d.	n.a.	32	−3.841				23.05		−6.003	−23.69	38.39	.0004168	.001527
			(2.24)				(12.8)		(1.71)	(7.82)	(13.7)	(.000551)	(.000525)
e.	.321	50	−2.161					13.17		−13.38	35.00	.0001917	.001441
			(1.89)					(30.1)		(6.82)	13.4)	(.000445)	(.000460)
f.	.315	33	−4.066	7.497					−4.234	−14.70	36.13	.001176	
			(2.46)	(9.92)					(1.88)	(8.67)	(16.0)	(.000505)	

[a] Conviction data for 1970 were missing for the following 17 states: Alabama, Alaska, Arkansas, Florida, Indiana, Iowa, Louisiana, Michigan, Nebraska, New Jersey, North Carolina, Oregon, Pennsylvania, Rhode Island, Texas, Virginia, and Wisconsin. The average term of incarceration, used to construct Δ°(E/C), was not available for Vermont for 1970.

[b] Equation (d) was constructed using the 2SLS method, with the difference variable Δ°(E/C) produced in the first stage from the differences in the following variables: proportion of families in poverty, proportion of nonwhites, median family income, proportion of residents of the ages 21 to 24, proportion of males, proportion of urban residents, school enrollment rate, resident population, divorce rate, proportion of adults employed, and a binary Southern variable.

crimes other than homicide are committed is included to ensure that the deterrence estimates are not contaminated by the omission of other factors that generally caused crime to increase from 1960 to 1970.[37]

Following the discussion of the previous section and the technical Appendix, we incorporate these variables in an empirical counterpart to Equation (1):

$$\Delta h = a + b \bullet \Delta Pe|c + c_1 \bullet \Delta(C/A) + c_2 \bullet \Delta Pov + c_3 \bullet \Delta NW + c_4 \bullet \Delta Y + c_5 \bullet \Delta Cr, \qquad (2)$$

where Δ represents the 1970 value of a variable minus the 1960 value. This approximation to the differential of the homicide rate serves as a basis for the estimates shown in Table 1.[38] Equations (a) through (d) form the variable Pe/c in four different ways, to ensure that the results are robust with respect to alternative ways of measuring the independent variable of primary interest. The first construction, E/C, is based on executions and murder convictions in the same year. Equation (b) uses executions in the year after the convictions, so that Pe/c here represents an objective forecast of the probability of execution given conviction.[39] Equation (c) constructs the numerator of the execution rate as the average annual number of executions for the three-year period centered about the base year, to reduce random error in the execution rate variable.[40] Equation (d) is also designed to reduce this error and to eliminate any bias resulting from the reverse effect that changes in the homicide rate may have on the execution rate. Equation (e) tests the deterrence hypothesis based on data for all fifty states by replacing the variables E/C and C/Q with the single variable E/Q. Equation (f) is the same as (a) except that the variable Cr is excluded to eliminate any biases resulting from the influence of other included variables on the variable Cr.[41]

---

37 A rationale for the inclusion of the non-homicide rate variable is given in Klein, *supra* note 4, at 355-56. This variable, however, is likely to be determined by other variables in the right-hand side of (1), so that its inclusion as a regressor could distort our inferences. Hence, results are reported below both with and without Cr.

38 The intercept coefficient, a, can be interpreted as a balancing term that reflects the net effect of omitted variables and stochastic measurement errors in the included variables. The other coefficients can be interpreted as average values of the partial differential coefficients for each of the independent variables. The coefficients of all equations in Table 1 other than (d) are estimated using ordinary least-squares regression analysis; equation (d) is estimated using a two-stage least-squares (2SLS) regression analysis, with the execution rate variable endogenized as noted in the table.

39 Ehrlich estimated the average lag from conviction to execution to be about 12 to 16 months. Ehrlich, *supra* note 3, at 407. I have estimated it in the same vicinity—14 months. Forst, *supra* note 11, at 756 n.50.

40 *See supra* note 17 and accompanying text.

41 *See supra* note 36 and accompanying text.

TABLE 2

REGRESSION COEFFICIENTS (AND STANDARD ERRORS) FOR VARIABLES USED TO EXPLAIN STATE HOMICIDE RATE INCREASES, 1960 TO 1970, EACH OBSERVATION WEIGHTED BY THE SQUARE ROOT OF THE 1960 STATE POPULATION

	R^2	N[a]	Intercept	Δ(E/C)	$\Delta(E_{+1}/C)$	$\Delta(\bar{E}/C)$	Δ°(E/C)[b]	Δ(E/Q)	Δ(C/Q)	Δ Pov	ΔNW	ΔY	ΔCr
a.	.583	33	−4.531	19.20					−5.881	−25.61	48.26	.0006002	.001383
			(2.50)	(10.6)					(1.83)	(8.86)	(15.6)	(.000548)	(.000454)
b.	.534	33	−6.360		−6.260				−6.109	−29.14	56.63	.001006	.001043
			(2.49)		(14.1)				(1.94)	(9.14)	(17.7)	(.000555)	(.000503)
c.	.539	33	−5.777			13.06			−5.824	−28.22	50.34	.0008635	.001271
			(2.52)			(18.6)			(1.94)	(9.18)	(16.8)	(.000553)	(.000494)
d.	n.a.	32	−4.294				26.48		−6.440	−24.96	48.48	.0006215	.001287
			(2.83)				(17.6)		(1.92)	(9.42)	(16.4)	(.000618)	(.000501)
e.	.379	50	−2.349					9.400		−14.31	46.87	.0003158	.001238
			(2.28)					(36.0)		(8.02)	(16.0)	(.000510)	(.000447)
f.	.434	33	−4.961	9.528					−5.252	−19.07	52.16	.001357	
			(2.86)	(11.5)					(2.07)	(9.82)	(17.8)	(.000558)	

[a] *See* Table 1 n.a.
[b] *See* Table 1 n.b.

None of these six formulations shows the execution rate to be a deterrent against homicides. In two instances, (a) and (d), the result for this variable is consistent with the hypothesis that executions, on balance, *provoke* homicides. More will be said about this hypothesis in the concluding section.

The results for the conviction rate variable, C/Q, on the other hand, are consistent with the deterrence hypothesis. While this finding is likely to be somewhat exaggerated by measurement errors in the number of homicides (a data element that appears both as the numerator of the dependent variable and as the denominator of the homicide conviction rate), the finding nonetheless does provide support for the theory that homicides are averted by the conviction and incarceration of homicide offenders.[42]

Findings similar to those of Table 1 are obtained when each observation is weighted by the square root of the 1960 state population, to reduce potential bias in the regression estimates associated with heteroscedasticity.[43] These results are shown in Table 2. It should be noted that while the standard errors tend to be slightly larger in Table 2 than in Table 1, the results are substantively quite similar.

IV. CONCLUSION

These findings, as well as earlier ones by Passell, Sellin, Schuessler, Zeisel, and others, do not invalidate the theory that capital punishment occasionally may deter individuals from committing homicides, nor do they invalidate the general theory of deterrence. They do suggest, however, that on balance the death penalty does not have a perceptible influence on the homicide rate. Even if some people are in fact deterred from committing murder by the prospect of being executed rather than subjected to a lengthy term of imprisonment, the available evidence suggests that as many people, and perhaps more, are provoked to commit murder by the existence of the death penalty.

Several different mechanisms may underlie the tendency of capital

42 Note that measurement errors in the number of homicides are not sufficient to produce an artificial appearance of deterrence in equation (e) of Table 1, where the potential for such distortion is also present. A more thorough treatment of this bias is presented by Klein, *supra* note 4, at 347-49.

43 "Heteroscedasticity" refers to the condition in which the variance of the dependent variable is not constant—specifically, the case in which the variance is related to the value of an independent variable. Heteroscedasticity is a problem in regression analysis because more weight tends to be given to observations with larger error variances. It is especially common in regression analyses involving cross-sectional data in which the observations consist of geographic units of varying population sizes. The standard correction in such a case is to weight each observation by the square root of the population for that observation.

punishment actually to provoke murder. First, as Ehrlich suggested,[44] the deterrent effect of capital punishment may be offset by the incentive it may create for murderers to kill policemen, witnesses, and others who could help to bring about their arrest, conviction and execution. Second, judges and juries may be less inclined to convict murderers when the risk of execution is high, in which case both the incapacitation and deterrent (both general and individual) benefits of conviction would be lessened.[45] Third, capital punishment provides a unique opportunity to commit suicide by having the state willfully and legally perform the killing.[46] Fourth, use of the death penalty by the state, despite an intention to convey the message that killing is unacceptable, may convey the opposite message to the general public. "Do as we do" may thus overpower "do as we say." By raising the sanction stakes from life imprisonment to death, the state may unwittingly induce criminals to "up the crime ante" and respond in kind with more homicides.

What, then, are we to make of the apparent conflict in evidence between the findings of Isaac Ehrlich, on the one hand, and those reported here (and by Passell, Zeisel, Bowers and Pierce, Schuessler, and Sellin), on the other? The evidence to date suggests that Ehrlich's findings are a product of his strong a priori belief in the theory of deterrence;[47] the availability of a variety of small, volatile data sets; and a larger variety of options for analyzing these data sets. A sufficiently strong a priori belief in anything can induce even the most well-intentioned of scientists to gravitate toward findings that support the belief.

Failure of the evidence to support the deterrent effect of the death penalty may be discomfitting to all who, through a personal experience, believe that sanctions influence behavior. We know that we are deterred by fines and other sanctions that apply to violations of traffic rules, income tax laws, and other matters of direct personal experience. Moreover, we all have some vague notion of how terrible it must be to be

---

44 *See* Ehrlich, *supra* note 3, at 398.

45 *Id.* at 405. "Individual deterrence" refers to the tendency for a sanction to control the subsequent behavior of the offender sanctioned; "general deterrence" refers to the tendency for a sanction to control the behavior of others.

46 Evidence of such behavior has been described by several commentators. *See* E. ABDEL FATTAH, A STUDY OF THE DETERRENT EFFECT OF CAPITAL PUNISHMENT WITH SPECIAL REFERENCE TO THE CANADIAN SITUATION, 39-40 (1972); Diamond, *Murder and the Death Penalty: A Case Report*, in CAPITAL PUNISHMENT IN THE UNITED STATES, 445 (H. Bedau & C. Pierce eds. 1976); Solomon, *Capital Punishment as Suicide and as Murder*, in CAPITAL PUNISHMENT IN THE UNITED STATES 432 (H. Bedau & C. Pierce eds. 1976); West, *Psychiatric Reflections on the Death Penalty*, in CAPITAL PUNISHMENT IN THE UNITED STATES 426-27 (H. Bedau & C. Pierce eds. 1976); Von Weber, *Selbstmord als Mordmotiv*, MONATSSCHRIFT FUR KRIMINOLOGIE UND STRAFRECHTSREFORM 61 (1937).

47 *See* Ehrlich, *supra* note 3; Ehrlich, *supra* note 7; Ehrlich, *Participation in Illegal Activities*, *supra* note 20.

executed. Hence, it might seem difficult to imagine that others are not deterred by the death penalty. Widespread public support for capital punishment, fed by frustration about crime and by a retributive zeal, is probably rooted in such personal beliefs about the efficacy of a strategy of deterrence.

As Anthony Amsterdam has observed, however, the personal experiences and beliefs of social scientists and the vast majority of the general public in fact may bear little resemblance to the experiences and beliefs of those predisposed to commit murder.[48] The behavior of the violence-prone minority simply may not correspond to our tidy general theories of human behavior, such as those embodied in the hypothesis that the death penalty deters murder. The evidence suggests that people prone to criminal violence may be as inclined to respond to an increased use of the death penalty by committing more homicides as they are to respond by committing fewer. The vast majority of society, on the other hand, is likely to be sufficiently deterred by social and legal sanctions other than the death penalty to render capital punishment a superfluous sanction for them. In short, the deterrence theory may apply to most sanctions, but the evidence indicates that the death penalty is an important exception, one for which other theories of behavior appear to be more applicable.

---

48 Amsterdam, *Capital Punishment*, 5 THE STANFORD MAGAZINE at 42 (1977). According to Amsterdam, the "real mainstay" of the general theory of deterrence is intuition, not evidence:

> You and I ask ourselves: we are not afraid to die? Of course! Would the threat of death, then, not intimidate us to forbear from a criminal act? Certainly! Therefore, capital punishment must be a deterrent. The trouble with this intuition is that the people who are doing the reasoning and the people who are doing the murdering are not the same people. You and I do not commit murder for a lot of reasons other than the death penalty. The death penalty might perhaps also deter us from murdering—but altogether needlessly, since we would not murder with it or without it. Those who are sufficiently dissocialized to murder are not responding to the world in the way that we are, and we simply cannot "intuit" their thinking processes from ours.

Id. at 47. Amsterdam has suggested to the author that persons predisposed to commit murder may tend to be moved by a variety of specific psychologies different from those that characterize the behavior of others. Letter from Anthony Amsterdam to Brian Forst (Nov. 22, 1982). One such difference could be a tendency for murderers to be more risk-prone than others—risk-prone, that is, not only in the conventional von Neumann-Morgenstern expected utility sense (i.e., characterized by preference for a gamble involving an expected return that is smaller than the amount of the bet), but in the unconventional sense that a gamble involving enormous negative stakes may be more attractive than an otherwise identical gamble with smaller negative stakes. Institution of the death penalty thus may tend to cause the minority who are predisposed to murder to be even more inclined to commit the crime than before.

APPENDIX

Rigorous empirical analysis of deterrence requires a statistical procedure that can distinguish the effects of criminal sanctions from other factors that affect crime rates (e.g., demographic and economic variables). The conventional approach for making such a distinction is to analyze the data using a multiple regression program on a computer. The estimates reported in Tables 1 and 2, like conventional estimates of deterrent effects, are derived from regression analysis.

The estimates reported in Tables 1 and 2 differ from conventional regression estimates, however, in that they are based on a less restrictive assumption about the mathematical form of the regression equation. As noted in the main text of this Article, *see supra* text accompanying notes 23-27, conventional regression estimates of deterrent effects are highly sensitive to assumptions about the form of the regression equation (e.g., whether it is assumed to be purely linear or linear in the logarithms of variables). A less restrictive approach is to assume only that the crime rate equation is a continuous function; one may then estimate the differential of the function rather than the crime rate equation itself. The differential of any continuous crime rate equation is additive and approximately linear, regardless of the precise specification of the crime rate equation.

Expressed symbolically, the differential equation corresponding to any continuous crime rate function,

$$h = f(x_1, x_2, \ldots, x_n),$$

is always of the additive form

$$dh = \frac{\delta h}{\delta x_1} dx_1 + \frac{\delta h}{\delta x_2} dx_2 + \ldots + \frac{\delta h}{\delta x_n} dx_n , \qquad \text{(A2)}$$

regardless of the particular form of Equation (A1). Denoting h as the homicide rate and x_1 as the execution rate, the average trade-off between executions and homicides can be obtained directly by estimating the partial differential coefficient,

$$\frac{\delta h}{\delta x_1}$$

via regression analysis. While this coefficient will, in fact, be a constant only when Equation (A1) is linear, the coefficients obtained by estimating the additive expression, (A2), with a linear regression model will approximate the average trade-offs, which are of primary interest, even if Equation (A1) is nonlinear.

Suppose, for example, that the homicide function were of the log-linear form,

$$h = ax_1^{b}, x_2^{b_2} \ldots x_n^{b_n}.$$

The differential of h would then be

$$dh = \frac{\delta}{\delta_1} dx_1 + \frac{\delta h}{\delta x_2} dx_2 + \ldots + \frac{\delta h}{\delta x_n} dx_n$$

$$= ab_1 x_1^{b_1 - 1} x_2^{b_2} x_3^{b_3} \ldots x_n^{b_n} dx_1 +$$

$$ab_2 x_1^{b_1} x_2^{b_2 - 1} x_3^{b_3} \ldots x_n^{b_n} dx_2 + \ldots +$$

$$ab_n x_1^{b_1} x_2^{b_2} \ldots x_n^{b_n - 1} dx_n \text{ ,}$$

which is approximately

$$dh = ax_1^{b_1} x_2^{b_2} \ldots x_n^{b_n} (b_1 dx_1 + b_2 dx_2 + \ldots + b_n dx_n) \text{ ,}$$

and which, at a particular margin, would be approximately

$$dh = h_0(b_1 dx_1 + b_2 dx_2 + \ldots + b_n dx_n) \text{ .}$$

The coefficients of primary interest are customarily estimated either by assuming that Equation (A1) is linear and estimating the coefficients of interest directly, or by assuming that the equation is nonlinear and estimating the parameters of the function and then calculating the coefficients of interest at the *average* values of the relevant variables. Thus, even when nonlinearity is assumed, the effects of primary interest are customarily expressed as constants.

Under the alternative method used here, those constants are estimated directly as the partial differential coefficients,

$$\frac{\delta h}{\delta x_i}$$

of expression (A2). Accordingly, the observations consist of measures of *changes* in the variables, corresponding to dh and the dx_i; rather than measures of the levels of the variables, as in conventional regression applications.