

Comments

The Deterrent Effect of Capital Punishment: Ehrlich and His Critics

Jon K. Peck†

Editors' Introduction

In the December issue, the *Journal* published a statistical debate on the deterrent effect of capital punishment between Professor Isaac Ehrlich and two sets of critics, Messrs. Baldus and Cole and Bowers and Pierce. Professor Ehrlich's original study of capital punishment, published last spring in the *American Economic Review*, used sophisticated statistical techniques to arrive at conclusions very different from earlier research in the field.^a Using regression analysis and economic theory, Ehrlich formulated and tested a model of the determinants of the murder rate, and found a significant deterrent effect associated with the use of the death penalty in the United States over the period from 1935 to 1969. The earlier research of Thorsten Sellin and others had consistently found no evidence of a deterrent effect. The issue of deterrence—in particular the technical merits of the Ehrlich study—has been raised in cases pending before the Supreme Court challenging the constitutionality of capital punishment.^b

In their critiques in the previous issue of the *Journal*, Baldus and Cole argue that Ehrlich's approach is less appropriate for testing the deterrent effect of capital punishment than are the less complex techniques used in

† Assistant Professor of Economics, Yale University.

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

b. See Editors' Introduction, *Statistical Evidence on the Deterrent Effect of Capital Punishment*, 85 YALE L.J. 164-69 (1975). *Fowler v. North Carolina*, cert. granted, 419 U.S. 963 (1974), argued, 43 U.S.L.W. 3582 (U.S. Apr. 21, 1975), restored for reargument, 422 U.S. 1039 (1975), which involves Eighth and Fourteenth Amendment challenges to capital punishment imposed under a mandatory sentencing procedure, was argued last spring and held over for reargument this Term. The Court has not yet scheduled the reargument in *Fowler*, but it recently granted certiorari in five other death penalty cases and scheduled them for argument on March 30. 44 U.S.L.W. 3439 (U.S. Jan. 22, 1976): *Proffitt v. State*, 315 So. 2d 461 (Fla. 1975), cert. granted sub nom. *Proffitt v. Florida*; *Gregg v. State*, 233 Ga. 117, 210 S.E.2d 659 (1974), cert. granted sub nom. *Gregg v. Georgia*; *State v. Roberts*, 319 So. 2d 317 (La. 1975), cert. granted sub nom. *Roberts v. Louisiana*; *State v. Woodson*, 215 S.E.2d 607 (N.C. 1975), cert. granted sub nom. *Woodson v. North Carolina*; *Jurek v. State*, 522 S.W.2d 934 (Tex. Crim. App. 1975), cert. granted sub nom. *Jurek v. Texas*. For a discussion of the facts of each of these cases, see N.Y. Times, Jan. 23, 1976, at 20, col. 7.

earlier research;^c Bowers and Pierce adopt Ehrlich's approach but argue that no evidence of a deterrent effect is found when his method is correctly applied.^d Ehrlich's reply defends the findings of his original study and attacks the analysis of his critics.^e In this issue, Professor Peck comments on the debate between Ehrlich and his critics, and Professor Ehrlich adds a brief Rejoinder.

In his study of the deterrent effect of capital punishment, Professor Ehrlich estimated a sophisticated econometric model of a type common in more traditional areas of economic analysis.¹ As in any empirical analysis, his econometric specification and choice of data require assumptions beyond those which may be derived from his theoretical analysis. These assumptions concern the functional form of the relationship between the murder rate and its determinants, the form of the remaining but unestimated equations in the overall model of the supply and (negative) demand for murder, the nature of the random disturbance term, the stability of the coefficients over time, the accuracy and appropriateness of the data, and the consequences of aggregating the behavior of individuals to the national level. The correctness of many of these assumptions depends on the correctness of others in this list. Ehrlich's assumptions have been challenged by Bowers and Pierce and by Baldus and Cole in articles in the previous issue of this *Journal*. In this Comment, I will discuss some of their criticisms of Ehrlich's analysis, suggest ways of resolving these disagreements, and point out some weaknesses of the paired-state or matching method—the approach which Baldus and Cole prefer to Ehrlich's regression technique. In addition, I will address briefly Ehrlich's reply to the two critiques of his work.²

I. Ehrlich's Critics

1. Both critiques address the issue of the proper functional form of the relationship between the murder rate and its determinants. It is claimed that evidence of a deterrent effect is found with the log-

c. Baldus & Cole, *A Comparison of the Work of Thorsten Sellin and Isaac Ehrlich on the Deterrent Effect of Capital Punishment*, 85 YALE L.J. 170 (1975).

d. Bowers & Pierce, *The Illusion of Deterrence in Isaac Ehrlich's Research on Capital Punishment*, 85 YALE L.J. 187 (1975).

e. Ehrlich, *Deterrence: Evidence and Inference*, 85 YALE L.J. 209 (1975) [hereinafter cited as *Ehrlich Reply*].

1. An econometric model consists of equations derived from economic theory and fitted to a sample of data using regression analysis. See Editors' Introduction, *supra* note b, at 167 n.15.

2. Part I of this Comment was prepared before Professor Ehrlich's Reply became available; several references to the reply have been added to the footnotes.

The Deterrent Effect of Capital Punishment

arithmetic form but not with the linear form, and that therefore Ehrlich's results depend critically on his choice of functional form.³

It is quite true that the incorrect use of a logarithmic form can cause the relatively small values for execution risk in the recent years in Ehrlich's sample to appear to be statistical aberrations which strongly influence the regression line. While the critiques point out the difficulty with using a logarithmic form for values of execution risk approaching zero,⁴ there are also problems with using the linear form for these values.⁵ The theoretical analysis is not much help in choosing the correct functional form, but the data are. Statistical tests applied to the data can determine the best form over a range of possibilities which includes both the linear and logarithmic forms.⁶ Of course neither form is likely to be exactly right, but only an approximately correct shape for the function is needed.

The related question of whether the same model adequately covers both the earlier and later sample periods—whether there is a structural change in the underlying relationship over time—is also a testable proposition. Bowers and Pierce find that the sign of the estimated elasticity of the homicide rate with respect to execution risk changes when the most recent years are dropped from Ehrlich's time series. They reject the possibility that the effect of the death penalty changed in recent years and conclude that the negative association between execution risk and the homicide rate is merely a statistical artifact.⁷ But neither Ehrlich nor these critics have rigorously tested for structural change over the sample period. One such test recently reported found evidence of a significant structural shift between the periods 1938-1962 and 1963-1969.⁸ In other words, in the recent subperiod the murder rate may be determined by different factors or may be determined by the same factors in a different way. However, the evidence of what appears to be structural change may also be the result of an incorrect specification of the model for the entire time period.

3. Baldus & Cole, *supra* note c, at 181; Bowers & Pierce, *supra* note d, at 199-200.

4. See Baldus & Cole, *supra* note c, at 175 & n.23; Bowers & Pierce, *supra* note d, at 201-02. For a general discussion of this problem, see Young & Young, *Estimation of Regressions Involving Logarithmic Transformations of Zero Values in the Dependent Variable*, 29 AM. STATISTICIAN 118-20 (1975).

5. The linear form uses the natural values of all variables, including execution risk, which can never be less than zero. When execution risk approaches zero, the disturbance term cannot assume large negative values. This constraint causes the estimated regression coefficients to have a systematic error. See H. THEIL, PRINCIPLES OF ECONOMETRICS 628-32 (1971).

6. See Box & Cox, *An Analysis of Transformations*, 26 J. ROYAL STAT. SOC'Y 211-52 (1964).

7. Bowers & Pierce, *supra* note d, at 200.

8. P. Passell & J. Taylor, *The Deterrent Effect of Capital Punishment: Another View*, Feb. 1975, at 4-6 (Discussion Paper 74-7509, Columbia Univ. Dep't of Economics) (on file with *Yale Law Journal*).

2. Bowers and Pierce stress the inaccuracy in Ehrlich's crime data for the early years in his sample.⁹ The seriousness of this problem is impossible to judge. However, the FBI procedure—retroactively re-adjusting the early data based on analysis of the effects of later increases in the coverage of the reporting network—is similar to the accepted Census Bureau method of adjusting economic time series data for the effects of seasonal variation by using adjustment factors computed in part from later data. The time series generated by the FBI procedure may be amenable to regression analysis.

The inaccuracy of the data is compounded by the inability of Bowers and Pierce to replicate Ehrlich's estimates.¹⁰ I do not know who is correct, but the differences seem too large to attribute to small differences in the definition of the variables or to correlations among the explanatory variables.¹¹ Ehrlich is clearly right in suggesting that the large differences in the estimates of the serial correlation coefficient reported by him and by Bowers and Pierce raise the question of error in the computation of this and other more important coefficients.¹² Good standards of documentation require that authors spell out exactly their computational procedures and identify the computer programs used.¹³

3. As Baldus and Cole emphasize, the level of aggregation of the analysis has an important effect on the validity of the conclusions.¹⁴ If there are substantial state or regional variations in behavior, and if arrest and punishment for homicide primarily involve the behavior of individuals and states, then estimation of the model ought to be done at the level of the state or, better, the individual. Nothing in Ehrlich's general approach precludes an application of his model to more disaggregated data.¹⁵ However, aggregate analysis is frequently

9. Bowers & Pierce, *supra* note d, at 187-92.

10. Compare *id.* at 195 n.19 (Table III) with Ehrlich, *supra* note a, at 410 (Table 3); see Ehrlich Reply, *supra* note e, at 210-11.

11. For an explanation of why this problem—known as multicollinearity—may lead to large differences in estimated coefficients with small differences in the data sample, see R. WONNACOTT & T. WONNACOTT, *ECONOMETRICS* 59-63 (1970).

12. Ehrlich Reply, *supra* note e, at 211. The serial correlation coefficient measures the correlation between successive values of the disturbance term. If there are computational errors in the calculation of this coefficient, there will also be errors in the estimated elasticities for the explanatory variables.

13. In his reply, Ehrlich does identify the computational procedures he used, and they appear to be reliable. *Id.* at 211 n.7.

14. Baldus & Cole, *supra* note c, at 175-77.

15. Ehrlich asserts that his new unpublished research using statewide data supports the findings of his original study. Ehrlich Reply, *supra* note e, at 217. Since I have not seen this work, I cannot comment on it. However, a recently published study based on state data and using a theoretical model similar to Ehrlich's found no evidence of a deterrent effect. Passell, *The Deterrent Effect of the Death Penalty: A Statistical Test*, 28 *STAN. L. REV.* 61 (1975).

The Deterrent Effect of Capital Punishment

performed in conventional economic problems with reasonable effectiveness. Such an analysis will be free of a statistical bias if the additional variables which ought to be included at the disaggregate (*e.g.*, state) level of analysis are unrelated to the explanatory variables at the aggregate level. However, the omissions will reduce the ability of the aggregate model to predict the murder rate.

4. The most basic point raised by Ehrlich's critics is the challenge by Baldus and Cole to his use of economic theory and regression analysis rather than experimentally based statistical procedures. The above enumeration of the problems with Ehrlich's analysis might suggest that an econometric approach to this problem is hopeless and that the paired-state comparisons endorsed by Baldus and Cole should prevail.¹⁶ But there are also difficulties with the latter approach, especially in a situation in which some relevant variables cannot be controlled or even measured.

Ehrlich begins with a careful economic analysis of the determinants of murder. Many researchers, including some economists, might object that his theory applies at best only to a small subset of homicides and that he has failed to identify and take into account some important causal factors. To the extent that these variables can be specified and measured, they can be tested and controlled for; to the extent they cannot be so identified, one must assume they are not related to the other explanatory variables. If this assumption is false, Ehrlich's results are biased. For some of the variables, such as the urban migration rate, which Baldus and Cole suggested should be included,¹⁷ the assumption seems plausible; for others, such as the level of crimes against property, the assumption seems implausible.

Any analysis must proceed conditionally on the specification of the model. Ehrlich's econometric approach involves many assumptions, some of which are testable; if these assumptions are true or nearly true,¹⁸ his method is able to use them to detect effects which might be too small to detect in a less fully specified model. No one claims that the deterrent effect of capital punishment is large, and it is consequently very important to bring the data to bear on the question as efficiently as possible. Ehrlich imposes a theoretical framework on his empirical analysis and necessarily makes more assumptions than are required in the paired-comparison analysis. If these assumptions

16. The paired-state or matching method compares homicide rates in neighboring abolitionist and retentionist jurisdictions which are as alike as possible with respect to other influences on the homicide rate. See Baldus & Cole, *supra* note c, at 171-72, 177.

17. *Id.* at 180.

18. See, *e.g.*, Fisher, *On The Cost of Approximate Specification in Simultaneous Equation Estimation*, 29 *ECONOMETRICA* 139 (1961).

are wrong, his conclusions are wholly or partially invalid. But many of these assumptions are testable.¹⁹

The matching approach has its own set of difficulties. It imposes relatively little explicit structure on the problem and is perhaps less likely than the econometric approach to find effects which are weak. The fundamental problem, however, is that the data are not generated in a controlled experiment. In making matched pairwise comparisons, the choice of pairs is inevitably subjective. To ensure that states are matched on all relevant variables requires a theory just as detailed as in an econometric analysis. In a classical experimental procedure, the control states and treatment levels are assigned randomly to observations to ensure that the effects of omitted variables are not systematic. This is obviously impossible in a "social experiment" where the death penalty is the treatment variable. Even if states are correctly matched in terms of the averages of all relevant variables, other differences may be important. For example, of two states with the same average permanent income, one may have a much greater proportion of low income families than the other. If low income families were disproportionately responsible for homicides, the pairing of the two would be inappropriate.

Another difficulty is the problem of spillovers between states. If state *A* has and uses the death penalty and its paired state *B* does not, it is possible that potential murderers would migrate to state *B* to avoid the death penalty and would thereby cause state *B* to have a higher homicide rate. One could not extrapolate from this comparison to a situation where all states or no states used the death penalty, since in that situation no one could avoid the penalty by migrating to an abolitionist state. Put another way, a murderer might be determined to commit the crime, regardless of the penalty, but might choose to commit it in the state where the consequences are less costly to him. The results of this behavior would give the appearance of a deterrent effect when none is present.

A final difficulty is the possible response of punishment policies to homicide rates. For example, if high or rising homicide rates led states to institute the death penalty and low or declining rates led states to abolish the penalty, retentionist states would tend to have higher homicide rates. This relationship could cancel out a possible negative correlation which would be produced if the penalty were in fact a deterrent. The paired-comparison approach cannot adequately

19. The argument for assuming the validity of certain basic (and statistically unverifiable) economic propositions in order to test more specific hypotheses is summarized in Koopmans, *Measurement Without Theory*, 29 *REV. OF ECON. & STAT.* 161-72 (1947).

The Deterrent Effect of Capital Punishment

separate these effects, and consequently could fail to yield evidence of an underlying deterrent relationship. On the other hand, the particular regression procedure used by Ehrlich specifically takes account of the possible response of punishment variables to the murder rate.²⁰

Any properly executed statistical analysis must include careful diagnostic checking of the model and consideration of alternative models which might also be consistent with the data. When using aggregate time series data, a number of different specifications are usually in general accord with the evidence. Thus the choice among competing models must be based on underlying theoretical analysis and on such empirical clues as particular data points which do not fit the model, correlations among explanatory variables, and stability of the important results under alterations in untestable and weakly maintained assumptions. For the reader to understand and evaluate the author's conclusions, some summary statement of the author's examination of this evidence is vital.²¹ Any sound regression analysis must provide such a statement. Diagnostic procedures may be less necessary for a paired-comparison analysis, which requires fewer assumptions, but the diagnostics should include, for example, an investigation of the sensitivity of the findings to the choice of pairings.

II. Ehrlich's Reply

1. In his original article, Ehrlich emphasized as support for his analysis that the ranking in order of magnitude of the estimated effects of his three deterrence variables conforms to his theoretical predictions.²² In his reply, he emphasizes that Bowers and Pierce, in attempting to replicate his results, obtained the same ranking for these estimates.²³ However, he does not show that this ranking would be inconsistent with other theories of the relationship between capital punishment and the murder rate. In the absence of such a demonstration, this evidence does not provide the strong support he claims.

20. In a regression equation in which the dependent variable (here the murder rate) exerts a causal influence on an explanatory variable (here the fraction of those convicted of murder who are executed), the equation must be estimated as if embedded in a larger model which simultaneously determines both variables. Ehrlich used a procedure appropriate for simultaneous equation estimation. See Ehrlich, *supra* note a, at 406; *Ehrlich Reply*, *supra* note e, at 219 (use of three-round regression procedure related to two-stage least squares).

21. In his reply, Ehrlich states that the results of diagnostic tests performed in the course of his research now in progress show that the logarithmic form is optimal for estimating his equation. *Ehrlich Reply*, *supra* note e, at 218 (likelihood ratio tests for optimal functional form within the class of single-parameter power transformations).

22. Ehrlich, *supra* note a, at 411, 416.

23. *Ehrlich Reply*, *supra* note e, at 211, 213, 219 n.29.

2. Ehrlich argues that the presence of errors of measurement in a variable can only weaken the variable's estimated effect and hence that improvements in the data could only strengthen his conclusions.²⁴ Economists are nearly always forced to use imperfect data and yet have drawn many strong empirical conclusions from them. But only the most harmless sorts of measurement error unambiguously reduce the estimated effects of a variable. If measurement errors in a variable are correlated with each other over time or are systematically related to other variables in the analysis, reducing the measurement errors will not necessarily strengthen the estimated effect. For example, when crime rates are high, the police may feel increased pressure simply to arrest someone for a crime even if that person is subsequently released. In Ehrlich's model the probability of arrest of the murderer, which is measured by the percent of murders cleared by an arrest, would then contain a measurement error correlated with the homicide rate. Because of the possibility of nonrandom measurement errors, Ehrlich should not assume that the true effects of execution risk and the other deterrence variables are systematically larger than the estimated effects.

3. Ehrlich argues that Bowers and Pierce have selectively deleted observations in order to distort his empirical findings.²⁵ However, examining subsets of the data to see if they are mutually consistent is an important part of validating a regression analysis. It is particularly important to examine the recent data, which in Ehrlich's analysis are most affected by the logarithmic transformation and which are in any event most relevant for policy analysis. Ehrlich correctly points out that estimation from a subset will be inefficient because of the reduced variability in the sample,²⁶ but he overlooks the purpose for estimating over subperiods—checking the specification of the model. When he analogizes the subperiod estimations performed by Bowers and Pierce to the selective deletion of observations in a regression analysis of the relationship between corn prices and quantities of corn demanded, he misses the point. In the terms of his example, Bowers and Pierce are asking whether the nature of the demand for corn has changed since the explosion in food prices of a few years ago; they are not trying to reverse the slope of the regression line by eliminating particular data points that give the line a negative slope.

4. Ehrlich defends his use of the logarithmic form for estimating his equation, but he does not address the problem of how to treat very

24. *Id.* at 213, 227.

25. *Id.* at 209, 214-17.

26. *Id.* at 215-16. It should also be noted that there are alternative estimation procedures—"full information" methods—which are more efficient than Ehrlich's.

The Deterrent Effect of Capital Punishment

small or zero values in the variables, a problem which is central to the criticisms made by Bowers and Pierce.²⁷

5. Ehrlich attacks the conclusion of Bowers and Pierce that the regression results are better for samples which omit the data from the 1960's.²⁸ While he correctly identifies statistical errors in their discussion, their basic finding—that whether the estimated effect of execution risk on the murder rate is positive or negative depends on the ending point of the sample period—casts doubt on the stability of Ehrlich's results.

None of the studies considered here can be said to have resolved the question whether the death penalty deters murder. The regression approach can be improved by better specification, better diagnostic testing and disaggregation. The analysis should consider time series data on states in ways that to some extent bring it closer to the paired-comparison method. Thus a resolution and synthesis of these approaches may be possible. Ehrlich concludes in his reply that no statistically meaningful evidence has been presented against his analysis and that his analytical framework withstands the criticisms raised against it. Without denying the usefulness of his approach, which appears to me potentially fruitful, I believe that his particular finding of a deterrent effect rests on as yet inadequately tested assumptions and on an incompletely validated model. I shall await with considerable interest his further contributions on the questions raised in this debate without, for the moment, concluding that he has established a statistically significant deterrent effect.

Finally, I would like to note that according to Ehrlich's equations, a one percent change in per capita income or labor force participation has a much greater effect on the homicide rate than does a one percent increase in the use of capital punishment.²⁹ Even if the deterrent effect of capital punishment were of statistical significance, it may be so small relative to other influences on the murder rate that it is of little practical significance.

27. See p. 361 *supra*. Ehrlich also asserts that using the natural rather than logarithmic values of the variables yields evidence of a deterrent effect. As independent support for this conclusion, he cites an unpublished paper by J. Yunker, *The Deterrent Effect of Capital Punishment: Comment*, Oct. 1975 (unpublished manuscript, on file with *Yale Law Journal*). I believe Ehrlich would reject the analysis in that paper on grounds similar to his criticism of Sellin's work. Further, Yunker's conclusion that "the real [deterrent] effect is at least five times the size of the effect estimated by Professor Ehrlich" is unmitigated conjecture. *Id.* at 17. The paper also appears to contain technical errors.

28. *Ehrlich Reply*, *supra* note e, at 219-20.

29. See Ehrlich, *supra* note a, at 409-10 (Tables 2 & 3) (elasticities for labor force participation and per capita income substantially larger than for conditional probability of execution).