SENSITIVITY ANALYSES OF THE DETERRENCE HYPOTHESIS:

LET'S KEEP THE ECON IN ECONOMETRICS*

by

Isaac Ehrlich

State University of New York at Buffalo

and

Zhiqiang Liu

Hong Kong University of Science and Technology

November 1998

ABSTRACT

Leamer and McManus applied Extreme Bound Analysis (EBA) in an empirical study of the deterrent effects of capital punishment and other penalties. Their analysis has questioned the validity of the deterrence hypothesis. The thrust of our paper is two-fold: First, by applying EBA to well known econometric models of demand, production, and human-capital investment, our analysis exposes and illustrates the inherent flaws of EBA as a method of deriving valid inferences about model specification. Second, since the analysis shows LMC's inferences about deterrence to be based on a flawed methodology, we offer an alternative, theory-based sensitivity analysis of estimated deterrent effects using similar data. Our analysis supports the deterrence hypothesis. More generally, it emphasizes the indispensable role of theory in guiding sensitivity analyses of model specification.

INTRODUCTION

Leamer[1] presents a scathing critique of classical sensitivity analysis concerning regression specification. The argument is that the classical approach lacks a systematic way to determine what explanatory variables truly belong in the regression when there is uncertainty about proper model specification. The remedy offered is Extreme Bound Analysis (EBA). Leamer illustrates the merits of EBA by applying it to a study of the deterrent effect of capital punishment.

EBA aims to determine the sensitivity of estimated effects of one or more "focus" variables of primary interest in a regression equation to the inclusion of additional regressors. It purports to solve the problem by allowing a researcher to partition a set of explanatory variables into "important" and "doubtful" subsets. Only the former is included in the regression with no restrictions. Linear combinations of "doubtful" variables' coefficients, in contrast, are constrained to match the researcher's prior beliefs -- generally zero. EBA produces the largest and smallest estimated coefficients of the focus variables (which themselves could be dubbed either doubtful or important), based on these restrictions. If the computed extreme bounds are too wide, or span zero, the effects of the focus variables are declared too "fragile" to warrant credibility.

Leamer's application of EBA to capital punishment is based on a study of murder rates across US states in 1950 conducted by McManus, and later published by him as a separate paper[2]. Both authors argue that EBA casts serious doubt on the validity of the deterrence hypothesis, at least as it pertains to capital punishment.[3] They conclude that the evidence on deterrence is essentially a product of the researcher's prior beliefs. "Bleeding hearts" could legitimately infer the absence of deterrence, while believers in "eye-for-an-eye" would reject as effective any penalty other than death. "Rational economists" are no more correct in inferring that offenders respond to incentives than are some criminologists who view crime merely as a product of social and environmental ills.

If true, this amounts to an impressive testimony about both the power of EBA and the validity of the economic approach to crime. We suppose that partly for this reason Leamer has titled his paper (supra note 1) grandiloquently "Let's take the con out of econometrics".[4] In this paper we suggest that the best way to keep the con out of econometrics is to keep economics in it.

In section 1 we examine the merits of EBA as a replacement for the classical approach. We find the promised remedy worse than the problem it is alleged to cure. The main methodological shortcomings of EBA have already been pointed out by McAleer, Pagan, and Volker [MPV][5]. Apparently, however, the latter's critique has escaped many economists, perhaps because it has not been backed up by convincing empirical evidence.[6]

In section 2 we use three data sets concerning consumption, production, and human capital investment to show how EBA, as applied by Leamer and McManus [LMC], would find basic economic theory to be fragile. It could dismiss laws of demand, production, and investment as a "con job", however, not because of any faults in the theory, but because of the inherent flaws of EBA itself. These flaws are also shown to be responsible for LMC's premature inferences about the validity of the deterrence hypothesis.

Since LMC’s analysis of deterrence is shown to be based on a faulty methodology, we feel an obligation not just to correct the impression their work has left about the strength of the underlying economic model, but also to offer an alternative sensitivity analysis using classical techniques. In section 3 we use similar cross-state data on the incidence of murder as well as other related crimes from both 1940 and 1950 to address a number of potential biases in estimates of supply-of-offenses functions in light of the underlying theory. We focus on 1940 and 1950 not just because LMC used 1950 data, but also because these are the most recent years in which a majority of states in the US still applied capital punishment at considerable annual frequencies.

Our analysis does not necessarily settle the issue. We do show, however, that LMC's analysis does not raise legitimate doubts about empirical results supporting deterrence, let alone prove that these results are just a product of researchers' prior beliefs. More generally, our work indicates that EBA's basic defect is that it does not put any weight on theory as the critical guideline for conducting sensitivity analysis. Indeed, when we conduct a theory-based sensitivity analysis using our cross-state data, we find additional support for the deterrence hypothesis.

1. THE BASIC MECHANICS OF EBA

EBA tests the sensitivity of the estimated effects of focus variables to alternative specifications of the linear model

(1) $Y = X\beta + u$, where $u \sim NID(0, \sigma^2 I)$,

by partitioning the regressors into alternative "important", or free, and "doubtful", or excludable, subsets, based on a researcher's prior beliefs. The independent linear restrictions corresponding to these beliefs are given by $\varphi R\beta = \varphi\gamma = r$, where $R = (0, I)$ is an exclusion block matrix, γ represents the coefficients of the designated doubtful variables in a given classification, and φ is a matrix generating all linear combinations of γ that satisfy the prior-beliefs constraint. While the latter may assume any prior values, in actual applications of EBA, r has been set as zero. The sensitivity analysis concerning focus variables' coefficients, β_0 -- a subset of β -- is then conducted by deriving the range of the restricted least squares estimates of these, $\hat{\beta}_0$, subject to all the alternative linear expressions of the restriction $\varphi\gamma = 0$.

For a simple illustration, suppose equation (1) contains only four independent variables as

(2) $Y = a + \beta_0 X_0 + \beta_1 X_1 + \gamma_1 Z_1 + \gamma_2 Z_2 + u$,

where X_0 denotes the focus variable (here allowed to be free), X_1 stands for an important variable, and Z_1 and Z_2 comprise the doubtful subset in a specific classification. The usual practice in classical analysis would be to test the effects of including or excluding the pair Z_1 and Z_2, or combinations thereof. EBA enlarges the search: in addition to accounting for these combinations, it substitutes for Z_1 and Z_2 the synthetic variable

(3) $W(\kappa) = Z_1 + \kappa Z_2$,

which encompasses all the remaining linear expressions of the prior restriction r = 0, by allowing κ to vary between negative and positive infinite values. For alternative values of κ, the restricted least squares estimates of β_0, $\hat{\beta}_0$, and their extreme bounds are estimated from the regression

(4) $Y = a + \beta_0 X_0 + \beta_1 X_1 + \eta W(\kappa) + u$.

EBA now offers two inference tests. If the distance between the extreme bounds of $\hat{\beta}_0$ is found to be either too wide or spanning zero, EBA concludes that the unrestricted OLS estimate of β_0 in equation (2), $\tilde{\beta}_0$, is too fragile to be credible.

MPV (supra note 5) term the corresponding inferences “type A” and “type B” fragility. They show that the conditions determining both are similar, except that type A involves more arbitrary assumptions because of the lack of objective criteria to determine what constitutes "too wide" a distance. We therefore focus in this paper on type B fragility, which has a straightforward interpretation: if the extreme bounds of the coefficient of "real price" in a properly identified demand function, for example, is found to span zero, this would be taken as evidence of the fragility of price theory. LMC also rely exclusively on type B fragility in their analysis of deterrence.

MPV prove that a necessary condition for type B fragility to exist in the case of any focus variable, that is, for $\hat{\beta}_0(\max) > 0$ and $\hat{\beta}_0(\min) < 0$, is given technically by

(5) $\chi^2_D > \chi^2_0$,

where χ^2_0 is the chi-square statistic for testing whether the focus variables' coefficients are significantly different from zero, and χ^2_D is the chi-square statistic for similarly testing the joint significance of the doubtful variables' coefficients in equation (1) or (2). They prove that when the focus variable is dubbed doubtful, (5) becomes also a sufficient condition for type B fragility to exist, and more important -- that it would then be mechanically guaranteed.

Breusch[7] provides more elegant proofs for these propositions. He shows that condition (5) can be expressed more tightly as

(6) $\phi\chi^2_D > \chi^2_0$,

where $\phi = [\mathrm{var}(\tilde{\beta}_0)\text{-}\mathrm{var}(\hat{\beta}_0)]/\mathrm{var}(\tilde{\beta}_0)$, and $\tilde{\beta}_0$ and $\hat{\beta}_0$ denote the unrestricted and restricted estimates of the focus variable in (2).[8] If the focus variable is free, a necessary condition for type B fragility is thus shown to depend on three parameters. The first, ϕ, is directly related to the correlation between the focus variable and the doubtful variable(s). The other two, χ^2_D and χ^2_0, indicate the statistical significance of the doubtful variables and the focus variable, respectively, in the regression model. From this analysis one can also derive a sufficient condition for type B fragility in the case of any focus variable:

(7) $[\mathrm{var}(\tilde{\beta}_0)\text{-}\mathrm{var}(\hat{\beta}_0)]\chi^2_D > (\tilde{\beta}_0 + \hat{\beta}_0)^2$.

In plain English, this analysis reveals two critical flaws in EBA:

I. When the focus variable is dubbed "doubtful", EBA's test of fragility becomes vacuous -- the focus variable automatically proves to be "fragile" simply because it is called doubtful.

II. Even when the focus variable is treated as "free", EBA could yield misleading inferences because of its mechanical nature (see section 2.D.). For by (7), it is possible to increase the likelihood that a focus variable be declared fragile if the researcher deliberately or inadvertently: a. classifies as "doubtful" those regressors which are jointly significant statistically and highly correlated with the focus variable; b. expands the number of doubtful regressors relative to the focus variable(s); c. nominates as a focus variable one that does not have a high t-ratio.

Put differently, EBA's fragility tests have the potential of rejecting a valid model as weak, or supporting a weak model as robust, simply through an arbitrary, or mechanical, classification of regressors into important (or free) and doubtful subsets, and by an arbitrary selection of regressors in general, with no guidance or backing from theory. The following examples illustrate these points graphically.

2. ILLUSTRATIONS OF EBA AND ITS INFERENTIAL IMPLICATIONS

To assess the seriousness of the potential inferential flaws of EBA, we apply it to three conventional econometric models of demand, production, and investment in a manner comparable to LMC's analyses of deterrence. Since EBA requires that all RHS variables be classified as either important or doubtful by some prior beliefs, in the following illustrations we ascribe these beliefs to groups such as "arch capitalists", "naive Marxists", "ersatz price theorists", or "radical environmentalists", with "rational economists" generally representing theory-based views. Lest the reader considers these classifications too artificial or silly, we should stress that we are simply emulating the type of classifications chosen by LMC**,** who do not rely on any coherent theory to justify the alternative classifications (see section 2.E). More important, any classification one might arbitrarily choose for conducting EBA could result, in principal, in similar inferential outcomes.

2.A. A Derived-Demand Function For Gasoline

The demand analysis illustrated is based on Douglas[9] (1992), using US time-series data from 1962-1985. Equation (8) represents a conditional derived-demand function for passenger-car gasoline:

(8) $Q^d = f(P_g, I_p, P_s, A, V)$,

where Q^d denotes the quantity of gasoline consumption, P_g the real (CPI-deflated) price of gasoline per gallon, I_p the real per-capita aggregate consumption expenditures (a measure of permanent income), P_s the real unit price (the CPI measure) of public transportation, A the stock of registered passenger cars, and V a fuel efficiency index of passenger cars (the miles per gallon traveled by the average vehicle), all in natural logarithms. The derived demand for gasoline is thus specified conditionally on the stock of cars in use, and the fuel efficiency variable is introduced to control, roughly, for the age and vintage composition of the latter.[10]

We do not attempt to address the possible simultaneity relationship between gasoline consumption and its price since EBA applies only to OLS regressions, and because the price of gasoline may have been partly regulated and largely dictated by the international market over the sample period. We examine only the inferential outcome of EBA concerning two "focus" variables: own price and income.

There are 64 (=2^6) possible prior classifications in this illustration (including the intercept as a variable) from which we select just 4 and ascribe them to the following schools of thought: 1. "Rational economists", implementing derived-demand theory, consider all RHS variables in (8) to be important. EBA does not apply in such a case, but the regression results support all the hypotheses of price theory; 2. "Ersatz price theorists" believe in the influence of gasoline's own price, but deny the relevance of income or any substitutes to passenger cars; 3. "Auto buffs" consider both income and the gasoline price as important but likewise deny the relevance of substitutes; and 4. "Radical environmentalists" discount the role of both prices and income and consider only "environmental" factors as important.

Tables 1 and 2 summarize the alternative classifications, and the extreme bounds of estimated focus variables' coefficients, based on these classifications. In Table 2, the key propositions of price theory are rejected by EBA since in 5 out of 6 of the estimated extreme bounds, the coefficients of the own-price and/or income variables have both positive and negative values.

2.B. A Production Function For Transportation Equipment

We next estimate a production function using cross-state data on the transportation equipment industry from 25 states in 1957. The data are taken from a well-known study by Zellner and Revankar [ZR][11]. They estimated a generalized, non-linear Cobb-Douglas production function, which allows the returns to scale to vary with output. Since EBA applies only to a linear regression model, however, we estimate just the conventional Cobb-Douglas production function embedded in their model.

Our focus variables are labor and capital. To allow for alternative prior classifications, as required by EBA, we augment the production function with three additional sets of variables, all in log form: (1) environmental (climate) variables, (2) randomly selected social indicators; and (3) an infrastructure indicator.[12]

Zellner and Revankar (along with most economists) would probably consider these additional variables to be of dubious relevance since there is no compelling theoretical reason for their inclusion in the production function. However, EBA does not restrict the selection of regressors or their prior classifications. We therefore select five classifications (out of 2048 possible ones) representing the following groups: 1. "Eclectics" who consider all explanatory variables to be important; (Again, EBA does not apply in this case, but the regression results indicate that the focus variables' coefficients remain positive and statistically significant, despite an expected reduction in their level.) [13] 2. "Neo-classics", viewing just capital and labor as important inputs but all other variables as doubtful; 3. "Arch capitalists", recognizing only capital and infrastructure (telephones); 4. "Naive Marxists", recognizing only labor and social indicators; and 5. "Radical environmentalists", recognizing just environmental, social, and infrastructural factors as relevant inputs.

Tables 3 and 4 summarize the results of these classifications. Except for the case of neo-classics, the extreme bounds applying to capital, or both labor and capital, span zero. EBA implies that labor and capital are too fragile to be considered relevant inputs in producing transportation equipment.

2.C. A Human-Capital-Investment-Linked Earnings Function

Our third illustration concerns an earnings function based on the Becker-Mincer human-capital investment model. We here pursue an application of EBA where the basic theoretical model is augmented by both possibly defensible and redundant additives (a hybrid case). Specifically, we treat the log of earnings as a function of 4 groups of explanatory variables: 1. the basic Mincerian set including years of schooling, job experience, experience squared, and the interaction of schooling and experience; 2. individual characteristics, including gender, marital status, race, and religion; 3. nature-of-employment indicators, including union membership and occupation and industry dummies; and 4. family background variables including father's years of schooling and father's occupation. Some (but not all) of these have been incorporated in actual studies.

Our sample, drawn from the 1988 PSID, consists of 838 observations. It is selected based on the availability of the variables discussed above and is restricted to household heads holding full-time jobs and reporting hourly wages. When we implement Mincer's basic equation, we obtain estimates of rates of return to schooling and job experience consistent with those reported in Mincer (1974)[14]. The expanded hybrid model leaves both of these focus variables statistically significant, although the rate of return estimates become lower (not a surprising result; some of the added variables may be proxies for the focus variables). EBA radically changes these inferences.

The prior-belief classifications we select out of roughly 2^{49} possible ones approximate the real or contrived positions of: 1. "Traditional human capitalists" who consider only the core Mincerian variables as important; 2. "Teachers' unions", who consider only schooling as important; 3. "Labor unions", who consider only years of job experience and employment characteristics as important; and 4. "Genealogists", who recognize only family background and individual characteristics as important.

Tables 5 and 6 summarize the results of these classifications and the extreme-bounds coefficient estimates of schooling and job experience. In all cases, the extreme bounds include negative and positive values. EBA implies that both schooling and job training are too fragile to warrant acceptance as relevant determinants of earnings.

2.D. Assessing the Inferences

In each of the preceding illustrations the inferential implications of EBA are attributable to one of the following mechanical reasons:

(1) In the trivial set of tests, the fragility inferences are automatically guaranteed by EBA simply because the focus variables are called doubtful (see section 1). This is true for cases 2 and 4 in Table 2; cases 3, 4 and 5 in Table 4; and cases 3, 4 and 5 in Table 6. All these tests are vacuous.

(2) In the non-trivial tests, where focus variables are left free, they are judged to be "fragile" as a result of the arbitrary and selective classification of variables into important and doubtful subsets (out of a very large number of such possible combinations), as well as the arbitrary incorporation of doubtful explanatory variables.

It is hardly surprising that the price of gasoline is judged to be fragile by "ersatz price theorists" who classify income as doubtful; that both own-price and income are similarly judged by "auto buffs" who consider prices of substitute goods to be doubtful; that capital is rejected as a credible factor of production when labor is classified as doubtful; or that schooling (job experience) is judged fragile by "labor unions" ("teachers' unions") who consider schooling (experience) to be doubtful. In each of these cases, some arbitrarily designated "doubtful" variables are, in fact, important on theoretical grounds to assure unbiased estimates of the focus variables' effects. This is analogous to the "omitted variable problem" in classical inference tests.

What is more surprising is that in our third illustration, both schooling and job experience are found to be fragile by "traditional human capitalists" who, based on theory, view other RHS variables in the "eclectics" specification to be of marginal worth. Recall that when the latter are either excluded from the regression (see footnote 14), or included as important (in the first row of Table 6), the focus variables remain statistically significant. Yet EBA declares the focus variable "fragile" when these other variables are classified as "doubtful".[15] Why?

As is well known, any ordinary (unrestricted) regression analysis, where all variables are by designation "important", would provide imprecise estimates of the model being implemented if the regression equation is stacked up with many additional, but irrelevant, regressors. This exposes the OLS estimates of the coefficients of the model's key ("focus") variables to an increased risk of multicollinearity, or otherwise reduces their efficiency by allowing for an unrestricted estimation of the coefficients of the additional irrelevant variables. EBA augments this potential mischief by relying on arbitrary linear combinations of the doubtful variables. These might turn out to be proxies for or significantly correlated with the focus variables. The "fragility" verdict imposed on the latter by EBA is then, again, just an arbitrary artifact of its mechanics.

2.E. LMC's Application of EBA to Deterrence

It is fitting to consider at this point the application of EBA to deterrence in Leamer and McManus (supra notes 1 and 2, respectively). Since the latter contains the former, we focus on McManus's presentation.

LMC adopt a variant of the supply-of-murder function specified in Ehrlich[16] (1977), but estimate it in a linear form:

(9) $q_m = f(Pc, Px, T, W, X, LF, UE, NW, AGE, URB, SOUTH)$.

The dependent variable is the murder rate across 44 states in 1950, and the focus variables are the probability of conviction (Pc), the conditional probability of execution given conviction (Px), and the length of imprisonment (T), representing deterrence variables. LMC group the other regressors as: a. “economic variables”, including median income (W), a measure of income inequality (X), and the rates of labor force participation (LF) and unemployment (UE); and b. “social variables”, including the fractions of the population that are non-white (NW), in the age group 15-24 (AGE), or in urban areas (URB), plus a dummy variable for Southern states (SOUTH).[17]

Tables 7 and 8 reproduce McManus's Tables 1 and 2. The 5 prior classifications of regressors used by LMC are quite in line with (and have actually inspired) the classifications we pursued in the preceding sections. Note that McManus has chosen to report the more conservative maximum and minimum estimates over the feasible ellipse and within the 90 percent data confidence region of the estimated coefficients of focus variables, rather than the extreme bounds themselves.

In 8 out of the 15 cases reported by McManus, the maximum and minimum estimates of deterrent effects span zero. He concludes, as did Leamer, that the deterrence hypothesis is thus too fragile to be believed. As should be apparent to the reader at this point, however, in 7 out of these 8 cases the estimates are computed when the focus variables are dubbed doubtful, which automatically guarantees a fragility verdict. Since the test is vacuous, the inference is erroneous and misleading.

The sole non-trivial case where the deterrence hypothesis seems to be challenged by EBA in Table 8 is the "eye-for-an-eye" case where the maximum and minimum coefficients of the probability of conviction (Pc) span zero when the imprisonment term (T) is dubbed doubtful. In terms of deterrence theory, this makes as much sense as dubbing a close substitute's price doubtful when estimating a commodity's own-price effect. By the deterrence hypothesis being tested, an unbiased estimate of the effect of imprisonment risk cannot be achieved without controlling for the prospective length of imprisonment when the two are correlated empirically.[18]

In Tables 7 and 8, the dummy variable allowing for different intercepts for executing and non-executing states, EPOS, is considered to be an important variable. McManus also shows (see Table 3 of his paper) that when EPOS is considered doubtful, there are 3 additional non-trivial cases in which Px now appears fragile. But this sensitivity of the regression results to the introduction of EPOS has been fully documented and discussed in Ehrlich (supra note 16) using classical sensitivity analysis. EBA reveals no new information concerning this variable. Moreover, it does not consider the possible relevance of EPOS (see section 3 below). All other 7 inferences in McManus's Table 3 are, again, vacuous since they involve the classification of focus variables as doubtful.

3. THE DETERRENCE HYPOTHESIS: AN ALTERNATIVE SENSITIVITY ANALYSIS

LMC's work may have created some doubt in the minds of students of the economics of crime concerning the validity of the general deterrence hypothesis. And although our preceding analysis has exposed the technical flaws of the EBA methodology LMC have implemented in reaching their critical inferences, it is incumbent on us to present an alternative, more comprehensive, sensitivity analysis by which to judge the theoretical model.

Note that apart from its inherent limitations, EBA addresses only one aspect of sensitivity analysis -- regression specification. No valid inferences can be reached about this issue, however, without addressing other targets of sensitivity analysis. In the context of our cross-sectional study of crime, these include improper functional form, heteroscedasticity, heterogeneity, measurement errors, simultaneity, and contemporaneous correlation. Some of these issues were tackled in Ehrlich (supra note 16). In what follows we extend the analysis using classical techniques.[19]

We begin with model selection. The theoretical "market model" of crime[20] contains a system of structural equations: supply-of-offenses, private demand for protection, public demand for law enforcement, and production functions controlling the probability and severity of the relevant criminal sanctions. The model also anticipates possible interactions among closely related offense categories.

Data exigencies preclude a full implementation of the complete system. We focus, instead, on estimating a supply-of-murder equation separately, or in conjunction with aggravated assault and robbery. A detailed justification of the regression specification is given in Ehrlich (supra note 16). We here go beyond that specification by treating key deterrence variables as endogenous.

3.A. Functional Form:

Ehrlich (supra note 16) provided a theoretical justification for using a log-linear format for the supply-of-offenses equation, which also accounts, in part, for potential measurement errors in crime counts. This was backed by a Box-Cox analysis of transformations using 1950 data. LMC, in contrast, relied exclusively on a linear format. To settle this issue, we conduct additional tests using a larger, pooled 1940 and 1950 sample.[21]

Since the Box-Cox analysis requires all variables to have positive values, whereas the conditional execution risk measure (Px) is zero in non-executing states, we base the analysis on the subset of executing states. The optimal transformation coefficient (λ) is found to be 0.0158 with a standard deviation of 0.127. $\lambda=0$ is approximately at the center of the 99 percent confidence interval about this coefficient, while $\lambda=1$ is far outside it. By this analysis the linear format ($\lambda=1$) must be rejected, whereas the log-linear format ($\lambda=0$) cannot be rejected, as optimal.[22]

We also conduct a modified Box-Cox test treating the probability of conviction Pc and the conditional probability of execution Px as endogenous variables through a non-linear, two-stage-least-squares procedure. The estimated value of λ, 0.028, is not significantly different from zero (with a standard error of 0.482). This further indicates the adequacy of the log format and justifies its use in the analyses of sections 3.C and 3.D.

3.B. Heterogeneity and the Role of EPOS:

Because the conditional execution risk measure, Px, equals 0 (percent) in non-executing states, the log transformation cannot be applied in regressions using the full set of states. Ehrlich (supra note 16) applied an approximate-logarithmic specification by transforming all variables by λ=0.001 and introducing a dummy variable, EPOS, which equals 1 for executing states and 0 for non-executing states.

EPOS was employed to account for a number of potentially serious estimation biases, addressed below. Indeed, the estimated regression coefficient of Px in particular was found in Ehrlich (supra note 16) to be sensitive to the introduction or exclusion of EPOS in the regression -- LMC's EBA analysis just reiterates this finding. Models 1 and 2 in Table 9 further underscore this sensitivity.[23]

The apparent heterogeneity of different states, as implied by the statistically significant effect of EPOS, deserves further scrutiny. In what follows we consider alternative explanations of EPOS's role, and pursue more general remedies for the heterogeneity problem it exposes.

3.B.a. Correcting for a bias introduced by the approximate-log transformation of Px

Besides capturing the effect of possible missing factors, EPOS corrects for a potential bias caused by the approximate-log transformation of the execution-risk measure, Px. While the optimal transformation is found to be logarithmic even for Px (in executing states)[24], transforming it by λ = 0.001 can grossly understate its level in non-executing states.[25] EPOS counters this "transformation error", which would bias the estimated Px effect toward zero while also distorting those of other regressors.

3.B.b. Correcting for observed vs. perceived risk of execution

An alternative remedy is to apply the superior log transformation even to Px, but only after assigning a positive value to Px=0. This remedy has an independent justification: a zero Px value does not necessarily imply that the perceived (conditional) risk of execution is also zero. Indeed, nine states have changed the status of executions between 1940 and 1950. The true value of Px in the non-executing states may thus be positive-- call it k>0.

Since k is unknown, we search for it via a grid search procedure, subject to the constraint that it is bounded between zero and the minimum value of Px in executing states. The log-likelihood value of the regression increases monotonically with k over our search interval. The optimal k is thus set as 1.4 -- just below the minimum value of Px in this sample's executing states.[26]

Models 3 and 4 of Table 9 present the estimated regressions with k=1.4. When EPOS is excluded in model 4, the estimated effect of Px, although smaller than in model 1, has the expected negative sign and is statistically significant at the 1 percent level, while all other estimated coefficients are comparable to those of model 1. When EPOS is included in model 3, all estimated coefficients, except that of EPOS and the constant term, are predictably identical to those in model 1.[27]

The estimated effects of income (W), income inequality (X), and the percent of non-whites in the population (NW) are positive and significant, but the effects of the labor force participation rate (LF) and the unemployment rate (UE) are insignificant. These variables are proxies for relative illegitimate and legitimate earning opportunities[28]. The effect of AGE is generally insignificant while those of URB and D40 are negative and positive, respectively, and statistically significant.[29] We have also iterated on the regression results (in all tables) by excluding the generally insignificant regressors AGE, UE, and LF. In all such iterations, the estimated effects of our deterrence measures are hardly affected, or slightly improved.

In models 5 and 6 we iterate on 3 and 4 by assuming that the conditional risk of execution in non-executing states is k=1 rather than 1.4 (that is, log k=0).[30] When EPOS is excluded in Model 6, the estimated effect of Px, albeit smaller, remains statistically significant. When EPOS is included in model 5, all estimated coefficients, except that of EPOS and the constant term, are again identical to those in model 1 (see footnote 27).

3.B.c. Using a linear transformation just for Px

Another way to avoid the bias created by the approximate-log transformation of Px=0 is to use a linear transformation just for this variable, that is, introducing Px in its natural form while all other variables are entered in log form. Although this hybrid model may not represent an optimal functional form, it avoids the systematic measurement bias addressed above.

In models 7 and 8 we re-estimate models 3 and 4 using this hybrid transformation. To enable easy comparisons, the regression coefficients of the "linear Px" variable are reported as estimated elasticities at the sample means. The most remarkable change is that the results are no longer sensitive to the presence of EPOS in the regression: the sign and statistical significance of the estimated Px coefficients hardly vary across models 7 and 8.

These results indicate that it is largely the bias resulting from the approximate-log transformation of Px=0 which is responsible for the sensitivity of its coefficient to the exclusion of EPOS.

3.B.d. Accounting for missing variables via a fixed-effects regression model

A more general way to resolve the apparent heterogeneity of our full sample is by estimating a fixed-effects (FE) regression model using the 1940-50 data. In addition to capturing the effects of state-specific missing variables, the FE model eliminates the bias caused by systematic measurement errors in Px across states in any one year, as the estimated effects of all regressors are based strictly on within-state variations in these regressors. However, if Px changes from a positive value to zero (or vice versa) between 1940 and 1950, the transformation bias discussed in 3.B.b would still apply here in the year in which Px becomes zero.

To account for this problem we introduce an interaction term involving the state dummies and a status-switch dummy variable (EXCHANGE), denoting a change in the execution status within a given state. In switching states, this variable assumes the value 1 in the year where Px>0, and 0 in the year where Px=0, while in non-switching states it is always 0.

Table 10 contains the results. To simplify the table, the estimated effects of all dummy variables are not reported, but the interaction-term effects are discussed below. Since we need to have at least two full observations per state, the sample is reduced to 82 states having both 1940 and 1950 data. The imprisonment variable is dropped from the regressions because T50 is used as a proxy for T40 as well (see footnote 21).[31]

The estimated coefficients in the FE regressions depend only on within-state changes in regressors, and are thus not strictly comparable to those reported in Table 9. Nevertheless, the results are statistically significant for all the deterrence variables, including Px, even when EXCHANGE is excluded. Moreover, when Px is specified in linear form in models 5 and 6, the inclusion of EXCHANGE has virtually no effect on the estimated regressions. The coefficient estimates for NW, X, AGE, and URB are insignificant here because these hardly vary within states between 1940 and 1950.

There are nine states which experienced a change in execution status between 1940 and 1950.[32] Only in three of these EXCHANGE has a positive and statistically significant effect -- Delaware, Kansas, and Montana. The statistical significance of EXCHANGE appears independent of whether the status changed from positive to zero, or vice versa, or whether it changed from legal abolition to moratorium. It appears linked to the distortion created when we use a log transformation for Px in the year where executions stop.[33]

3.B.e. Threshold effects and EPOS

So far we have stressed the role of EPOS as a correction for biases resulting from the approximate-log transformation of Px, or missing factors that distinguish executing from non-executing states. It may also be tempting to view EPOS as accounting for a threshold effect associated with the enforcement status of capital punishment (that is, whether Px is positive or zero), as distinct from the extent of its imposition when enforced. Such interpretation would be premature, however, because of a simultaneity problem associated with EPOS. A moratorium or abolition status (Px=0) can be expected to be endogenously determined, much like the frequency of executions (Px>0), along with other deterrence variables. In particular, both Px>0 and Px=0 are expected to respond in a similar way to the perceived risk of murder victimization, as well as the economic, legal, and political environments in the state.[34] The apparent sensitivity of Px to the introduction or exclusion of EPOS in our full sample can thus be explained partly by reverse causality: an exceptionally low murder rate reduces the incentive to resort to the death penalty as a sanction. It is for this reason that in the next section we continue to treat Px as a single, albeit endogenous, variable in our full sample.[35]

3.C. Simultaneity:

We have so far treated deterrence measures as exogenous variables. The market model of crime suggests, however, that the crime rate, public expenditures on enforcement, and thereby the instruments of law enforcement, including capital punishment, are simultaneously determined. Higher crime rates may lower the probability of apprehending and convicting offenders because of a "crowding effect" on the efficacy of law enforcement activity if enforcement budgets are constant. But optimal law enforcement also requires that enforcement budgets, and the willingness to resort to harsher penalties, increase in response to a higher risk of victimization, as is the case for private-protection efforts (see Ehrlich supra note 20 and references therein).

We use the conventional Hausman test to determine the endogeneity of all the deterrence variables: Pc, Px, and T. We first test the endogeneity of Pc and Px while treating the imprisonment length (T) as predetermined. This is because our imprisonment measure -- the median time spent in State prisons by offenders first released of specific crimes -- is indeed predetermined. We then conduct an independent test, however, to validate this assumption.

We also apply the conventional Basmann test to validate the over-identification restrictions we impose when excluding some of our reduced-form variables from the supply-of-offenses equation.

3.C.a. The underlying structure

In addition to the supply-of-offenses function, our underlying econometric structure contains production and demand functions for private protection and public enforcement. The production functions governing the probabilities of apprehension and punishment are conventionally specified to include real expenditures on law enforcement as well as technical variables affecting the productivity of these expenditures, such as population density and the volume of crime itself (because of the adverse "crowding effects" these may generate). The demand for enforcement activity, in turn, and thus the production of apprehension and punishment probabilities, depend on the perceived risk of, and losses from, victimization, which are produced by the crime rate in the population and by population wealth (see Ehrlich supra notes 28, 20).

The conditional probability of execution (Px) is controlled, in addition, by legislative and judicial decisions, including a legal prohibition or moratorium on executions, which may be affected by the political makeup of the legislature and the judiciary, or the ideological and moral sentiments of the electorate they represent.

3.C.b. The instrumental variables used

In addition to the exogenous variables in the murder equation, we have used in the reduced form the total number of police personnel per 100,000 population, POL, the state population, N, the accumulated rates of all crimes and violent crimes (excluding murder) in the 5 years preceding 1940 and 1950, church membership, defined as the portion of a state population belonging to a church, a dummy variable distinguishing Southern states (SOUTH), and another identifying the political party of the candidate winning the last Presidential elections in the state (PARTY).[36]

POL is a proxy, albeit imperfect, for enforcement efforts directed against all crimes. Since murder constitutes only a small fraction of the total, POL can be considered exogenous in the regression. The same holds for the accumulated flows of total offenses and violent crimes over the preceding 5-year period, which induce a demand for both law enforcement and self-protection by potential victims. Church membership may capture the impact of population homogeneity and solidarity on the productivity of law enforcement, and the political makeup of the population may affect the emphasis placed by the authorities on probability and severity of punishment as means of combating crime.[37]

3.C.c. The results

Table 11 summarizes the results. Note, first, the test statistics reported at the bottom of the table. The row titled HN shows the results of Hausman's tests concerning the endogeneity of the pair Pc and Px. As the χ^2 statistics indicate, Hausman's tests reject (at the 1 percent level) the null hypothesis that Pc and Px are jointly exogenous. The row titled HN-T, in turn, contains independent Hausman tests of the hypothesis that T is an exogenous variable, conditional on Pc and Px being endogenous. These fail to reject the null hypothesis that T is indeed exogenous. The row titled BN contains the results of Basmann's tests of our overidentification restrictions. These fail to reject at the 5 percent level the null hypothesis that our overidentification restrictions are valid.

The results presented in Table 11 are highly consistent with our theoretical expectations, as well as with those reported in Table 9, and they provide new insights:

i) The estimated deterrent effects (elasticities) associated with the probability of conviction and the conditional probability of execution given conviction are statistically significant and larger than their GLS counterparts in Table 9. The simultaneity bias may affect the estimated deterrent effects in either a negative direction (due to the crowding effect) or a positive direction (due to the optimal demand for enforcement). The 2SLS results indicate that the latter bias is the more important empirically.

ii) The estimated elasticities of the murder rate with respect to the conviction risk are significantly larger than the corresponding elasticities with respect to the conditional execution risk in all specifications. This confirms one of the strong predictions of the deterrence hypothesis [38].

iii) The estimated elasticities of measures of relative opportunities in legitimate and illegitimate activity (median income and income inequality, labor force participation, and unemployment) have generally the right signs but most are not statistically significant. The AGE, URB and D40 effects remain the same as in Table 9.

iv) The estimated key deterrent effects are generally robust to alternative transformations of Px (linear v. log) and alternative values of k. Indeed, since the estimated effect of Px is now based on its predicted values from the reduced-form equation, $\hat{P}x$, the results should be less sensitive to the k-value assigned to Px=0.

v) There may be missing exogenous variables that explain not just variations in the observed values of Px in executing states, but also the decision to abolish the death penalty legally or de-facto. Treating $P\hat{x}$ as a continuous variable across executing and non-executing states may thus be open to question. To test for the importance of this potential distortion, we estimate the supply-of-murder equation using only the sub-sample of executing states. That the estimated elasticities in models 4 and 5 are in the same ball park as those in corresponding models based on the full set of states, adds credibility to the 2SLS analysis.[39]

3.D. 3SLS Estimates of Interrelated Crimes:

Finally, we extend our 2SLS analysis by allowing for interdependencies among a set of violent crimes -- murder, robbery, and assault -- that may interact as either complements or substitutes. Also, the error terms in the corresponding equations may be significantly correlated. Complete data concerning the deterrence variables for robbery and assault, however, are available only for 1950. We thus restrict our sample accordingly.

To account fully for the cross effects of a deterrence variable associated with one crime on the incidence of the other two we would need to treat each as an additional regressor in the other equations. This would be impractical, however, given the small sample size in 1950 and the high correlation among the predicted values of the probabilities of conviction for each crime. We do introduce, however, the conditional probability of execution for murder Px5 (here measured as the average number of executions for murder in the preceding 5 years over the number of convictions in the sample year) as a regressor in all three equations. In each equation we treat the crime-specific probabilities of conviction Pc_i and the conditional risk of execution for murder (Px5) as endogenous variables. The equation set is then estimated via Zellner and Theil's three-stage-least squares procedure (3SLS).[40]

The pattern of results for all three crime categories, as reported in Table 12, is consistent with those reported in Table 11 for murder alone. The main new insight concerns the effect of execution risk on robbery and assault. Since the imposition of the death penalty was restricted only to murder in 1950, we expect little direct deterrent effect on the incidence of assault or robbery coming from Px5, while its indirect effect (resulting from the impact on murder) could be even positive if assault, for example, were a substitute for murder. The small and insignificant effect of Px5 on that crime may be consistent with this interpretation. The negative and more pronounced, albeit insignificant, effect of Px5 on robbery indicates a degree of complementarity between the two crimes, as may indeed be the case, since some murders are committed as a by-product of robbery.[41]

CONCLUSION

The basic objective of EBA is admirable: to help determine which regressors belong in a regression analysis where either the researchers, or their critics, are in doubt about the relevant regression specification. But this important objective cannot be achieved mechanically through a statistical package.

As our analysis and illustrations show, when EBA calls a focus (specific deterrent) variable "doubtful" it will surely show it to be doubtful. And by calling some additional regressors which are an integral part of the same theory "doubtful", it can again show the focus variable to be fragile. Much of this is common to classical sensitivity analysis, but EBA exacerbates the problem because of arbitrary designations of regressors, or combinations thereof, as important and doubtful. Moreover, a valid sensitivity analysis of regression specification must address other potential "sensitivities" that are implied by the tested hypotheses. Lincoln's adage applies here: "if you call a tail a leg, how many legs has a dog got? five? no, four, because calling a tail a leg does not make it a leg."

Some researchers may have the notion that EBA is a useful device for summarizing how wide the range of estimated coefficients can be in cases where the literature produces conflicting results by relying on different regression models. Indeed, this appears to have been a key motivation for many researchers’ reliance on EBA. In light of our empirical illustrations, however, this notion is questionable. First, by designating “not-so-significant” regressors as doubtful variables, one can show that one’s focus variables’ estimates are robust even though they are not. The apparent robustness of estimates when such designations are made is illusory. Second, much of the claim of EBA to be a superior method of investigating the sensitivity of focus variables’ estimates to changes in model specification stems from subjecting the sensitivity tests to all possible linear combinations of doubtful variables. In terms of economic theory, what do these linear combinations represent? Do they make sense? And why should a researcher be concerned with any of them? As our illustration and analysis in parts 2.C. and 2.D. indicate, these linear combinations may introduce their own distortions. The alleged advantage of EBA, in view of these imprecise and crude tests, is at best presumptuous, and at worst misguided.

In conducting sensitivity analyses of a regression model, it is important not to lose sight of the basic objective of the model: testing hypotheses. The validity of a regression specification must be judged by the extent to which it implements the theory underlying these hypotheses. A regression model (a theory) can be considered useful if it offers a number of discriminating hypotheses that are testable and rejectable by data, if the hypotheses are consistent with a wider range of observed behavior, and if the tests are corroborated by evidence from independent samples and independent studies. Valid criticisms of, or challenges to, such a regression model should be based on alternative, systematic models, the robustness of which should be judged on similar criteria.

Whether the deterrence hypothesis is valid has clearly not been proven by our analysis because we are able to implement only partially the market model underlying this theory. We are not able to estimate the complete structure of the model, some of the data are noisy, and we lack direct measures of some important theoretical constructs. But our results are consistent with key behavioral propositions: the deterrent, or preventive, effects of both probability and severity of punishment are established after taking into account a number of restrictions posed by the theory. The results are consistent with a key proposition of deterrence theory: that the probability of conviction has a larger deterrent effect on the incidence of murder (in elasticity terms) than the conditional probability of execution (or any other specific penalty) given conviction. Also, the probability of imprisonment has a greater deterrent effect than its severity. Although challenges abound, these and other implications of the economic model of crime have been corroborated through independent studies using independent samples and investigating alternative crime categories.

References

William A. Bartley, and Mark. A. Cohen, The Effect of Concealed Weapons Laws: An Extreme Bound Analysis, 36 Economic Inquiry 258-65 (1988)

S. T. Breusch, Simplified Extreme Bounds, in Modelling Economic Series: Readings in Econometric Methodology 72-81 (C. W. Granger eds., Oxford University Press 1990)

Russell Davidson, and James G. MacKinnon, Several Tests for Model Specification in the Presence of Multiple Alternatives, 49 Econometrica 781-93 (1981)

K.M Deravi, C.E Hegji, and D.H Moberly, Government Debt and the Demand for Money: An Extreme Bound Analysis, 28 Economic Inquiry April 390-401 (1990).

Douglas, Evan J., Managerial Economics, (4th ed., New York Prentice Hall 1992).

Isaac Ehrlich, Participation in Illegitimate Activities: An Economic Analysis. In Essays in the Economics of Crime and Punishment (G. S. Becker and W. M. Landes eds., New York Columbia University Press 1974).

Isaac Ehrlich, The Deterrent Effect of Capital Punishment: A Question of Life and Death, 65 American Economic Review 397-417 (1975).

Isaac Ehrlich, Capital Punishment and Deterrence: Some Further Thoughts and Additional Evidence, 85 Journal of Political Economy August 741-88 (1977).

Isaac Ehrlich, Crime, Punishment, and the Market for Offenses, 10 Journal of Economic Perspectives Winter 43-67 (1996).

Isaac Ehrlich & George D. Brower, On the Issue of Causality in the Economic Model of Crime and Law Enforcement: Some Theoretical Considerations and Experimental Evidence, 77 American Economic Review May 99-110 (1987).

Isaac Ehrlich & Joel C. Gibbons, On the Measurement of the Deterrent Effect of Capital Punishment and the Theory of Deterrence, 6(1) J. Legal Stud. January 35-50 (1977).

Richard Fowles & Peter D. Loeb, Effects of Policy-Related Variables on Traffic Fatalities: An Extreme Bounds Analysis Using Time-Series Data, Southern Economic Journal October 359-66 (1995).

Kishore Gawande, Are U.S. Nontariff Barriers Retaliatory? An Application of Extreme Bounds Analysis in the Tobit Model, Review of Economics and Statistics November 677-88 (1995).

William H. Greene, Economic Analysis (2nd ed. New York Macmillan 1993).

Steven E. Landsburg, The Armchair Economist (New York The Free Press 1995).

Edward E. Leamer, Let's Take the Con out of Econometrics, 73 American Economic Review March 31-43 (1983).

Ross Levine & David Renelt, A Sensitivity Analysis of Cross-County Growth Regressions, 82 American Economic Review September 942-63 (1992).

G. S. Maddala & Forrest Nelson, Switching Regression Models with Exogenous and Endogenous Switching, in Proceedings of the American Statistical Association (Business and Economics Section) 423-26 (1975).

Michael McAleer, Adrian R. Pagan & P. A. Volker, What Will Take the Con out of Econometrics? 75 American Economic Review June 293-307 (1985).

Michael McAleer & Michael R. Veall, How Fragile Are Fragile Inferences? A Re-evaluation of the Deterrent Effect of Capital Punishment, 71 Review of Economics and Statistics 99-106 (1989).

Walter S. McManus, Estimates of the Deterrent Effect of Capital Punishment: The Importance of the Researcher's Prior Beliefs, 93 Journal of Political Economy February 417-25 (1985).

Jacob Mincer, Schooling, Experience, and Earnings (New York Columbia University Press 1974).

Johan Torstensson, Determinants of Intra-industry Trade: A Sensitivity Analysis, 58 Oxford Bulletin of Economics and Statistics August 507-24 (1996).

Arnold Zellner & Nagesh Revankar, Generalized Production Functions, 37 Review of Economic Studies 241-50 (1970).

Arnold Zellner & H. Theil, Three Stage Least Squares: Simultaneous Estimation of Simultaneous Equations, 30 Econometrica 63-68 (1962).

FOOTNOTES

* We are indebted to Arnold Zellner, Adrian Pagan, Nagesh Revankar, Yong Yin, and an anonymous referee for valuable comments and suggestions. We alone are responsible for errors.

1 Edward E. Leamer, Let's Take the Con out of Econometrics, 73 American Economic Review March 31-43 (1983).

2 Walter S. McManus, Estimates of the Deterrent Effect of Capital Punishment: The Importance of the Researcher's Prior Beliefs, 93 Journal of Political Economy February 417-25 (1985).

3. Leamer (supra note 1, at p.42) concludes that "any inference from these data about the deterrent effect of capital punishment is too fragile to be believed". McManus (supra note 2, at p.425) reaches a similar inference.

4. In a private conversation, Bob Clower, who accepted Leamer's paper for publication, told the first author that he became convinced by Leamer's work that we do not know anything about deterrence, since Leamer's piece proves that it is all in the eye of the beholder! Steven Landsburg, who is quite sympathetic to the deterrence hypothesis, also mentions Leamer (supra note 1) in Steven E. Landsburg, The Armchair Economist (New York The Free Press 1995), as a challenge to the hypothesis.

5 Michael McAleer, Adrian R. Pagan & P. A. Volker, What Will Take the Con out of Econometrics? 75 American Economic Review June 293-307 (1985).

6. Recent studies which have relied on EBA include: K.M Deravi, C.E Hegji, and D.H Moberly, Government Debt and the Demand for Money: An Extreme Bound Analysis, 28 Economic Inquiry April 390-401 (1990); Ross Levine & David Renelt, A Sensitivity Analysis of Cross-County Growth Regressions, 82 American Economic Review September 942-63 (1992); Richard Fowles & Peter D. Loeb, Effects of Policy-Related Variables on Traffic Fatalities: An Extreme Bounds Analysis Using Time-Series Data, Southern Economic Journal October 359-66 (1995); Kishore Gawande, Are U.S. Nontariff Barriers Retaliatory? An Application of Extreme Bounds Analysis in the Tobit Model, Review of Economics and Statistics November 677-88 (1995); Johan Torstensson, Determinants of Intra-industry Trade: A Sensitivity Analysis, 58 Oxford Bulletin of Economics and Statistics August 507-24 (1996); William A. Bartley, and Mark. A. Cohen, The Effect of Concealed Weapons Laws: An Extreme Bound Analysis, 36 Economic Inquiry 258-65 (1988).

7 S. T. Breusch, Simplified Extreme Bounds, in Modelling Economic Series: Readings in Econometric Methodology 72-81 (C. W. Granger eds., Oxford University Press 1990).

8. Since ϕ gives the fraction of variance reduction due to the restrictions imposed on doubtful variables, γ, its value must fall between 0 and 1. Therefore, this condition is tighter than the one without the term ϕ.

9 Douglas, Evan J., Managerial Economics, (4th ed., New York Prentice Hall 1992).

10. The data are taken from Douglas (ibid, p. 185-87), except for the aggregate consumption proxy for permanent income, which we have taken from the Economic Report of the President, 1988 Washington: US Government Printing Office.

11 Arnold Zellner & Nagesh Revankar, Generalized Production Functions, 37 Review of Economic Studies 241-50 (1970).

12. Climate variables include: annual '(heating) degree days', average percentage of possible sunshine hours per day, and average humidity at 7 am. Social indicators include: median school years completed, total crime rate (offenses known to the police) per 100,000 population, number of prisoners received in state and federal prisons, and murder rate. The infrastructural indicator is the number of registered telephones. All data are taken from the Statistical Abstract of the US in 1957.

13. The coefficients of labor and capital in the conventional Cobb-Douglas specification embedded in ZR's model (all other regressors omitted), are of similar magnitudes and significance as the estimates reported in ZR (supra note 11). Variables with significant coefficients in the "eclectic" specification are annual degree days, crime rate, median school years, and number of prisoners received in state and federal prisons.

14. Jacob Mincer, Schooling, Experience, and Earnings (New York Columbia University Press 1974). The estimated basic regression is: log hourly wage = 0.649 (2.92) + 0.09*S (5.62) + 0.04*EXP (4.07) - 0.00065*EXP2 (-6.84) - 0.00006*S*EXP (-0.09). The t-ratios are given in parentheses.

15. A somewhat similar pattern emerges in our second illustration where the introduction of environmental, social, and infrasctructural variables in the eclectic specification does not affect the significance of labor and capital in production, but when these same variables along with labor are dubbed doubtful, capital is judged to be fragile.

16 Isaac Ehrlich, Capital Punishment and Deterrence: Some Further Thoughts and Additional Evidence, 85 Journal of Political Economy August 741-88 (1977).

17. The measures used for non-standard variables are: qm: urban murder rate per 100,000 population based on UCR samples in 1950; Pc: ratio of murder convicts entering state prisons to the estimated number of total state murders in 1950; Px: ratio of the average number of executions in 1946-50 to the estimated number of convictions in 1950; T: median length of time served in prison by offenders first released in 1951; W: median family income in 1949; X: fraction of families with less than half of the median income in 1949.

18. There is a little more to the story. LMC's regressions have been weighted erroneously by the fourth root of the population size, instead of the intended square root (see Michael McAleer & Michael R. Veall, How Fragile Are Fragile Inferences? A Re-evaluation of the Deterrent Effect of Capital Punishment, 71 Review of Economics and Statistics 99-106 (1989)). Second, we have discovered a few large errors in their reported Non-Whites variable. Third, they have relied exclusively on a linear regression specification, which we show in section C to be an inadequate specification. For

comparison, we have corrected LMC's data for these three problems, and report the corresponding maximum and minimum estimates over the feasible ellipse and within 95 percent data confidence region of the focus variables' estimates in the square brackets below McManus's results in Table 8. The sole non-trivial case in which the extreme bound estimates of the Pc coefficient spanned zero in McManus (Supra note 2) now disappears.

19. In all of the following regressions, we use a correction for heteroscedasticity, based on the analysis in Ehrlich (supra note 16), which LMC have also adopted, albeit with an error (see footnote 18). All variables are multiplied by the square root of the state's urban population in the sample year.

20 See Ehrlich Isaac Ehrlich, Crime, Punishment, and the Market for Offenses, 10 Journal of Economic Perspectives Winter 43-67 (1996).

21. The pooling of the two sample years, after allowing for different intercepts to account for some sampling differences, is supported by Chow's test. Specifically, the relevant F statistics for models 4, 6, and 8 of Table 9 below -- 0.114, 0.115, and 0.102, respectively -- are all smaller than the critical value, F95%[9,67]=2.023. Data limitations for 1940 dictate that in the pooled 40-50 regressions Px be defined as the ratio of the average number of executions over the 4-year period preceding and including the sample year, relative to the estimated number of convictions in that year. Also, T50 serves as the imprisonment-length variable for both 1950 and 1940, and W and X in 1940 are computed from wage and salary income, while in 1950 they are computed from family income.

22. Russell Davidson, and James G. MacKinnon [Several Tests for Model Specification in the Presence of Multiple Alternatives, 49 Econometrica 781-93 (1981)] provide a test for comparing linear and log-linear specifications. This test also rejects the linear form, but fails to reject the log-linear one at the 5 percent level of significance.

23. The results of models 1 and 2 differ slightly from those in Ehrlich (ibid), Table 8, because here only Px is transformed by λ=0.001, while all other variables are in log form. Also, we here include the variables W, X, LF, and UE (see equation (9)), except that the latter two are related to the male (not the total) population. The dummy variable D40 (1940=1) is incorporated to account for some data differences between 1940 and 1950 (see footnote 21). For this reason we have also considered interaction terms between W, X and D40, but these were found to have insignificant effects and virtually no effect on other regressors' coefficients. The same holds for the SOUTH dummy. The latter variables have therefore been excluded from the regressions.

24. We have also applied a Box-Cox procedure to search for the optimal transformation for Px alone while keeping all other variables in logarithmic form. The test, based on the set of executing states, unequivocally rejects the linear, but not the logarithmic, specification for Px.

25. For example, a 10 percent difference in Px [transformed via (Px0.001-1)/0.001] in executing states where Px varies from 20 percent to 10 percent is 0.7 (or 3-2.3), whereas the same 10 percent difference in the transformed Px between states where Px=10 percent and 0 is 1002.3. The actual values of the transformed Px variable vary between 0.385 and 4.616 in executing states, but are uniformly -1000 in non-executing states.

26. Even better results are obtained if we ignore the upper bound. We have also implemented an alternative approach for estimating k involving a two-step procedure. In the first, we run a reduced-form regression for Px (using the reduced-form model of section 3.C), based on the sub-sample of executing states. In the second, we use the estimated coefficients of the latter regression to project the value of Px in non-executing states, k', based on the mean values of the reduced-form variables in the latter states. We then use the projected k' to estimate the relevant supply-of-murder equations. The results are similar to those of model 3 of Table 9, with estimates of the coefficient of Px higher in magnitude, and those of Pc somewhat lower in magnitude than those of model 3. Moreover, the results are insensitive to the introduction or exclusion of EPOS. A similar pattern emerges when we use this alternative method of estimating k in tables 10-12 (see sections 3.C and 3.D below).

27. This is because when EPOS is used as a regressor, the estimated regression coefficients are influenced only by variations in the regressors within each subset of states, distinguished by EPOS. And since there is no variation in Px (k) in non-executing states, all the estimated coefficients, except that of EPOS and the constant term, would be invariant to the assigned value of k (including $k \approx 0$). Only the coefficients of the latter two depend on k: a high enough value of k would eliminate the significance of EPOS.

28 See Isaac Ehrlich, Participation in Illegitimate Activities: An Economic Analysis. In Essays in the Economics of Crime and Punishment, G. S. Becker and W. M. Landes eds., New York Columbia University Press, 1974.

29. Ehrlich (supra note 16) provides an explanation for why variations in UE and LF across states, rather than over time, may not be efficient measures of legitimate labor market opportunities. URB, measuring the percentage of a state's population in urban areas, accounts for the fact that the FBI's crime data from 1940 and 1950 are based on sample of the state's urban population, whereas other regressors relate to the state population as a whole. URB thus serves in part as a "correction" factor for systematic reporting differences across urban and rural areas, but it may also account for the fact that in urban areas better access to medical facilities improved the odds of survival of assault victims, especially in 1940. Ehrlich (ibid) reports some supporting findings, based on Vital Statistics data.

30. This is consistent with William H. Greene's (Economic Analysis, 2nd ed. New York Macmillan 1993 p.329) recommendation that only positive values of a variable such as Px be relied upon in the search for an optimal transformation.

31. Note that the fixed-effects regression model does not require using a positive value (k) for Px=0. Indeed, the regression results are invariant to k because any biases resulting from measuring Px=0 are fully corrected for by the state dummies and the EXCHANGE interaction term. As in models 1 and 3 of Table 10, the assigned value of k can affect only the coefficients of the estimated interaction terms or the state-specific dummy variables. A high enough value of k eliminates the significance of EXCHANGE.

32. Among these, 5 switched from an executing to a non-executing status (Delaware, Montana, New Hampshire, Utah, and Wyoming), while 4 switched from a non-executing to an executing status (Kansas, Nebraska, New Mexico, South Dakota).

33. We have also implemented a potentially superior random-effects (RE) regression model, which yielded similar estimates for Pc and Px. We do not report these, however, because Hausman's specification test rejects the RE model as a valid specification.

34 See Isaac Ehrlich & Joel C. Gibbons, On the Measurement of the Deterrent Effect of Capital Punishment and the Theory of Deterrence, 6(1) J. Legal Stud. January 35050 (1977).

35. Indeed, we find that treating Px>0 and Px=0 (that is, EPOS) as separate endogenous variables creates a serious multicollinearity problem. We have also examined the possibility of a self-selection bias in connection with the status of execution in executing and non-executing states (as distinguished by EPOS), by applying a model of switching regressions with endogenous switching suggested by G. S. Maddala & Forrest Nelson, Switching Regression Models with Exogenous and Endogenous Switching, in Proceedings of the American Statistical Association (Business and Economics Section) 423-26 (1975). By this method execution status is treated as endogenous, and the estimated structural coefficients are allowed to be different in the subsets of executing and non-executing states. In both classes of states, the selection-bias-corrected estimates of the effects of Px, Pc, and T on murder are found to be consistent with the deterrence hypothesis generally, and estimates of deterrent effects reported in Table 9 specifically. However, in this analysis we treat deterrence variables themselves as exogenous -- our non-executing sub-sample is too small to allow their treatment as endogenous.

36. Church membership covers all major religious groups. No data were published for our sample years. We therefore use the reported data from 1936 and 1952 as proxies for these variables in 1940 and 1950, respectively. (The sources are "Religious Bodies", 1936, and "Churches and Church Membership in the US", 1952.) Since there is no compelling reason for including SOUTH, CHURCH, and PARTY (Democrats=1) in the structural equation, we have excluded them in the reported regressions. Our estimates, however, are robust to the treatment of these three variables. The exclusion of other instruments (along with SOUTH, CHURCH, and PARTY) from the structural equation is supported by Basmann's overidentification tests reported below. Similar tests and results were reported in Isaac Ehrlich & George D. Brower, On the Issue of Causality in the Economic Model of Crime and Law Enforcement: Some Theoretical Considerations and Experimental Evidence, 77 American Economic Review May 99-106 (1987).

37. Several instruments exhibit significant predictive value in the reduced form regressions. The probability of conviction (Pc) is found to be positively related to POL (significant at the 5 percent level) and the accumulated violent crime rate (its t-value is 1.656). It is negatively and significantly related to T and LF. CHURCH and SOUTH are found to increase Pc, while the Democrat (PARTY) dummy, URB, and the accumulated total crime rate lower it. In the regression predicting Px, the variables T, NW, UE, the population size, and the accumulated flow of total and violent offenses are statistically significant, and all but the accumulated violent crime rate have positive signs. The Church membership and South are negatively related to Px, whereas PARTY is positively related to Px. While it would be tempting to explain some of these results on theoretical grounds, this would not be warranted, since our set of instruments represents only part of the theoretically relevant reduced form corresponding to the complete system.

38 See Isaac Ehrlich, The Deterrent Effect of Capital Punishment: A Question of Life and Death, 65 American Economic Review 397-417 (1975), and Ehrlich, Supra note 16.

39. We could not implement a 2SLS/fixed-effects procedure. This is because the predicted values of Pc and Px, obtained from fixed-effects reduced forms, show little variation between 1940 and 1950 for each state, and are thus highly collinear with the state dummies.

40. Arnold Zellner & H. Theil, Three Stage Least Squares: Simultaneous Estimation of Simultaneous Equations, 30 Econometrica 63-68 (1962). The three-round procedure involves the following steps: (1) we obtain the predicted values for all four probability measures (Px5, Pc, P-ass, P-rob) from reduced-form regressions similar to the one used in the previous subsection; (2) we compute the 2SLS estimates for each equation, and obtain an estimate of the covariance matrix of the error terms; (3) we then compute the GLS estimator. It has been shown that this estimator is both consistent and asymptotically efficient. We do not iterate on the computation of this three-round estimator because such iteration does not lead to a gain in asymptotic efficiency (see Greene, Supra note 30).

41. The 3SLS estimates of the effects of deterrence variables in the case of murder are uniformly larger than their counterparts in a simple SUR estimation procedure (not reported here), consistent with our results in the preceding section. The contemporaneous disturbances in different equations are found to be significantly correlated, which contributes to the efficiency gain in our 3SLS estimates over the 2SLS estimates for each crime (not reported here).