December 25, 2009

Estimating the Impact of the Death Penalty on Murder

John Donohue, *Yale University*
Justin Wolfers, *University of Pennsylvania*

Estimating the Impact of the Death Penalty on Murder

John J. Donohue, III, *Yale Law School,* and Justin Wolfers, *The Wharton School, University of Pennsylvania*

This paper reviews the econometric issues in efforts to estimate the impact of the death penalty on murder, focusing on six recent studies published since 2003. We highlight the large number of choices that must be made when specifying the various panel data models that have been used to address this question. There is little clarity about the knowledge potential murderers have concerning the risk of execution: are they influenced by the passage of a death penalty statute, the number of executions in a state, the proportion of murders in a state that leads to an execution, and details about the limited types of murders that are potentially susceptible to a sentence of death? If an execution rate is a viable proxy, should it be calculated using the ratio of last year's executions to last year's murders, last year's executions to the murders a number of years earlier, or some other values? We illustrate how sensitive various estimates are to these choices. Importantly, the most up-to-date OLS panel data studies generate no evidence of a deterrent effect, while three 2SLS studies purport to find such evidence. The 2SLS studies, none of which shows results that are robust to clustering their standard errors, are unconvincing because they all use a problematic structure based on poorly measured and theoretically inappropriate pseudo-probabilities that are

designed to capture the key deterrence elements of a state's death penalty regime, and because their instruments are of dubious validity. We also discuss the appropriateness of the implicit assumption of the 2SLS studies that OLS estimates of the impact of the death penalty would be biased against a finding of deterrence.

Does the death penalty provide greater deterrence of murders beyond that afforded by a sentence of life imprisonment? This question has been actively debated for centuries, with those arguing that capital punishment is a more severe punishment that will provide greater deterrence opposed by those who argue that state-sanctioned executions provide an environment conducive to unsanctioned homicides. Alternatively, some have argued that the death penalty is not as dreadful to potential murderers as the thought of life imprisonment or that the death penalty is less cost-effective than its alternatives. In the past half-century, this debate has turned from social theorists to empiricists.

Given the availability of relatively high-quality data on American murder rates, executions, criminal justice statistics, and other relevant control variables as well as the number of researchers conducting sophisticated econometric studies over the last thirty years, one would think that a consensus would have emerged about the answer to this ostensibly simple question. Indeed, some believe that it has.

Radelet and Akers (1996) surveyed seventy past presidents of the academic criminology associations asking them "on the basis of their knowledge of the literature and research in criminology" (Rubin, 2006b) whether the death penalty lowered the murder rate. Only eight of these eminent criminologists responded affirmatively to the statement that "the death penalty acts as a deterrent to the commitment of murder—that it lowers the murder rate," while fifty-six (or 84%) argued against deterrence. (Three past presidents had no opinion, while a further three failed to respond to the survey.) Radelet and Lacock (2009) administered this same survey in 2008 to an updated list of top criminologists (not including those included in the 1996 survey), and generated similar results: 88% of the seventy-six respondents thought there was no deterrent effect of the death penalty. Dieter (1995) surveyed a nationally representative sample of U.S. police chiefs and county sheriffs, finding only 26% found the statement that the "death penalty significantly reduces [the] number of homicides" to be accurate, while 67% believed it to be inaccurate (7% were unsure).

Yet Becker (2006) argued that "the preponderance of the evidence does indicate that capital punishment deters." Joanna Shepherd's 2004 congressional testimony concurred: "In the economics literature in the past decade . . . there is a very strong consensus . . . all of the modern economic studies in the past decades have found a deterrent effect." Paul Rubin (2006) echoed this assessment before the Senate Judiciary Committee claiming that "The literature is easy to summarize: almost all modern studies *and all the refereed studies* find a significant deterrent effect of capital punishment. Only one study questions these results."

We provide the emphasis in the Rubin quote because the reference to "one study" is to Donohue and Wolfers (2005), hereinafter "DW"—our own rather critical response to recent death penalty research published in the *Stanford Law Review* in December 2005. In that paper we evaluated many of the death penalty studies that Rubin deemed to establish the deterrent effect of the death penalty. In each case the foundation for these claims proved to be quite shaky, albeit for varying reasons, including coding errors, inappropriate study designs, improper calculation of standard errors, and reliance on invalid instrumental variables.

Our aim in this paper is not to provide a single "best" estimate of the impact of the death penalty on murder, but rather to provide a systematic review of the issues confronting researchers working on this question, as well as to review the state of the recent and growing literature on the deterrence question. Section 1 provides an overview of the history of the econometric debate on the deterrent effect of the death penalty, briefly summarizing the methodologies and conclusions of some of the major studies evaluating the impact of the death penalty. Section 2 looks at New York State's experiment with capital punishment, which lasted from 1995 to 2004, as a way to illustrate some of the modeling complexities that must be addressed in trying to estimate the impact of capital punishment on murder. Perhaps surprisingly, the two counties in New York City (Manhattan and the Bronx) with District Attorneys (DAs) who strongly and vociferously opposed the death penalty experienced the largest drops in murders.

Section 3 illustrates that OLS estimates of the impact of executions on murder during the post-moratorium period (post-1976) consistently show no statistically significant evidence of deterrence, while Section 4 notes that a number of studies find greater evidence of deterrence using instrumental variables techniques. Unfortunately, the 2SLS studies have some

major flaws. First, they all use a problematic structure based on poorly measured and theoretically inappropriate pseudo-probabilities that are designed to capture the key deterrence elements of a state's death penalty regime. Second, their estimated deterrence effects are statistically insignificant if clustering the standard errors is necessary. Third, since researchers have isolated few credible sources of exogenous variation in execution policy, there is little reason to credit 2SLS estimates that rest on such flawed instruments. Section 5 goes on to analyze the issue of possible endogeneity bias in the OLS estimates, and argues that in the post-moratorium period that bias may well operate in *favor* of deterrence. If so, then 2SLS estimates that find a stronger deterrent effect than that of the OLS estimates—which is the case in the studies by Dezhbakhsh et al. (2003) (hereinafter "DRS"), Mocan and Gittings (2003) (hereinafter "MG") and Zimmerman (2004)— are presumptively invalid.

Section 6 offers concluding remarks and notes that if the 2SLS studies are unreliable and the OLS studies provide an upper-bound estimate of the impact of the death penalty, then the absence of any statistically significant effect in the OLS estimates presents a major challenge to those arguing for a "strong deterrent" of capital punishment. Of course, there is a fundamental difficulty in teasing out the impact of the death penalty in post-moratorium period in the United States in that there may not have been sufficient variation across states in execution policy to yield precise estimates of the relationship between capital punishment and homicide.

But recent evidence from the massive increase, and then subsequent decline, in executions in Singapore suggests that potential murderers tend not to be responsive to levels of execution that are dramatically higher than those in modern day Texas. Indeed, the time path of homicide in Hong Kong looks strikingly similar to that of Singapore even though the former never used capital punishment and formally abolished it shortly before Singapore began its experiment in extremely heavy reliance on the death penalty.

1. Some History of the Econometric Debate

1.1. The Pioneering, but Now Superseded, Early Work

In 1975, Isaac Ehrlich developed a sophisticated econometric model using national time-series data and claimed to show that each execution

between 1933 and 1969 saved eight lives. Although Ehrlich merits credit as an original and innovative contributor to an important conceptual literature on the economics of deterrence, his paper precedes the major advances in micro-econometric evaluation of panel data. Specifically, a national time-series analysis is incapable of providing robust empirical estimates of the impact of the death penalty because it cannot identify whether any changes in murder rates are occurring in the states that invoke capital punishment. Indeed, the national time-series approach can only yield a valid estimate in the unlikely event that two conditions hold: (i) the rate of executions is orthogonal to the large, unexplained movements in the murder rate over time; and (ii) an execution anywhere in the United States is equally likely to deter a murder throughout the United States (even in jurisdictions that have no death penalty) or the unexplained differences in murder rates across different jurisdictions are constant over time.[1]

In response to criticisms of his time-series approach, Ehrlich produced a second cross-sectional study in 1977 that looked at murder rates and executions across states in two years—1940 and 1950. But it is now recognized that cross-sectional studies are even less suited for estimating the causal impact of executions because they cannot easily account for the large and persistent, unexplained differences in crime rates across states. Specifically, the cross-sectional analysis cannot address the unobserved heterogeneity that underlies the fact that in the United States murder rates tend to be substantially *lower* in nonexecuting states than in high execution states, as the regional breakdown in Table 1 suggests. Clearly, there are reasons why, say, Maine with a 2004 murder rate of 1.4 per 100,000 is safer than Mississippi, where the murder rate of 7.8 is more than five times as high, but fully

---

1.   A particularly telling problem with the Ehrlich time-series analysis was that his finding of a deterrent effect emerged only because Ehrlich used a log specification that gave disproportionate weight to the fact that the small reduction in the execution rate from a very low level to virtually zero in the late 1960s was accompanied by a very large jump in murders. Stopping the analysis in 1962 rather than 1969 or using a nonlogarithmic model would generate no deterrent effect. This nonfinding seems more intuitively plausible in that the 80% decline in the execution rate over the period of 1933–1962 occurred during a period of falling murder rates (from 8.8 per 100,000 to 4.6, a decline of 47.7%) and the large post-1962 increase in murder rates occurred identically in states that never had the death penalty as well as those that did. See Figure 3 in DW. Numerous other conceptual and data problems with Ehrlich's work are discussed in Section 4.1 and in the Appendix.

Table 1. Homicide and Execution Rates by Region: 2002

Region	Homicide rate (per 100,000)	Execution rate (per 100,000)
Northeast	4.1	0.0
Midwest	5.1	0.014
West	5.7	0.002
South	6.8	0.059

explaining these enduring differences without the benefit of a state fixed-effect dummy has proven to be a daunting challenge.

By the time various scholars—backed up by a 1978 National Academy of Sciences report—were done pointing out the infirmities in Ehrlich's analysis, few outside of the University of Chicago believed that either his national time-series analysis or his cross-sectional analysis afforded substantial support for the view that each additional execution saves many lives.[2]

1.2. The Move to Panel Data

It is now widely recognized that panel data models with state and year fixed effects, while hardly foolproof, are far more likely to identify the causal impact of a legal or policy change, such as the death penalty, than time-series or cross-section models (Nerlove, 2002). As we will see, the difficulties in trying to reliably estimate the impact of the death penalty using the best tools are daunting enough; there is really no hope that we can do so using less reliable statistical methodologies, such as time-series and cross-section studies. Unfortunately, not everyone has gotten this message. While such papers continue to be published, they likely should be ignored.[3] For this reason, we limit our attention to the issues involved in the specification

---

2. Cameron (1994) provides a detailed review of the pre-"panel data" literature by Ehrlich, his critics, and other scholars using time-series or cross-section approach. He concludes that "the presence of capital punishment on the statute book acts as some kind of deterrent but variations in its use do not." Interestingly, Ekelund et al. (2006) finds that the presence of a death penalty statute *increases* murder but higher use dampens this increase.

3. Ehrlich and Liu (1999) and Liu (2004) (both using Ehrlich's original state-level, cross-section data from 1940 and 1950) and Narayan and Smyth (2006) (using national-level time-series data for 1965–2001) may offer insights about innovations in these more primitive statistical tools, but, given the now extensive panel data analyses, these less-discerning approaches cannot be expected to advance our understanding of the causal impact of the death penalty on murder.

and estimation of panel data models, and to one new matching study that compares high-execution Singapore with abolitionist Hong Kong.

John Lott and David Mustard created a panel dataset beginning with 1977 data (for use in their evaluation of state right-to-carry concealed handgun laws) that greatly influenced a new round of estimates of the impact of the death penalty on crime.[4] Lott and Mustard essentially followed the Ehrlich model and then shared their state and county data with a number of researchers, who used it to analyze the death penalty: MG and Zimmerman (2004) conducted their analyses on state data using both OLS and 2SLS methods, and DRS relied on 2SLS estimation on county data to offer support for the view that the death penalty deters murder.[5] Unfortunately, this piggy-back approach has meant that some of the conceptual errors of Ehrlich have persisted over time, and some of the data and specification problems that were introduced in the original Lott and Mustard dataset have infected subsequent papers that have used their data.[6]

At the same time, another major state panel data study by Katz et al. (2003) using OLS estimation concluded that there was little empirical

---

4. The 1977 date was mandated by Lott and Mustard's desire to use arrest rate data, which became available by county in that year. While Lott and Mustard were following Ehrlich's lead in using arrest rates as an explanatory variable in their crime model, the problems with using this variable are discussed in Section 4.1.

5. Section 4.1 describes some general inadequacies in Ehrlich's theoretical and econometric specification that have carried over to the models employed by DRS, MG, and Zimmerman. MG relied primarily on an OLS panel data regression, but also presented one table of 2SLS estimates, which we critique in Section 4.2.2. The OLS estimates in both Zimmerman (2004) and Zimmerman (this issue) do not indicate that capital punishment deters (that is, the explanatory variable "executions/death sentences" measured contemporaneously or lagged one year was negative but not statistically significant).

6. The DRS dataset, which came from Lott and Mustard, contains demographic variables that provide the percentages of county population of the following age groups: ages 0–9, 10–19, 20–29, 30–39, 40–49, 50–64, and 65 and over; race groups: black, white, and other; and sex groups: male and female. After summing these variables within groups, we find that the minimum values for total age, total race, and total sex are 99.97%, 99.97%, and 70.56%, respectively, which suggests something has gone quite wrong in the sex breakdown. The maximum values for total age, total race, and total sex are 132.67%, 105.76%, and 156.66%, respectively. Clearly, the age groups, race groups, and sex groups do not sum up to 100% for all counties over all years. Most notably, the sum of female and male population percentages falls below 90% for nearly 80% of the observations (this percentage excludes observations that would be dropped from the regressions because one or more of the demographic variables are missing data, which constitute about 10% of the total number of observations).

support for the deterrence hypothesis. These authors use neither the Ehrlich model nor the Lott and Mustard data, which means they thereby avoided a number of serious data and specification problems.

While DW raised substantial concerns about the DRS, MG, and Zimmerman papers and found the KLS paper more reliable, not everyone agrees: for example, David Muhlhausen of the Heritage Foundation, testifying on June 27, 2007, before the Subcommittee on the Constitution, Civil Rights, and Property Rights of the Committee on the Judiciary of the United States Senate, reviewed the work of Ehrlich and these three pro-deterrence papers:

> the recent studies using panel data techniques have confirmed what we learned decades ago: Capital punishment does, in fact, save lives. Each additional execution appears to deter between three and 18 murders.

Muhlhausen's confident assessment of the recent research was notable in that it ignored both the KLS paper, our own critique of the previously mentioned studies that Muhlhausen found persuasive, and every other study disputing a finding of deterrence for capital punishment (some recent ones prior to Mulhuausen's testimony include Berk, 2005; Fagan, 2006; Fagan et al., 2006; some subsequent articles include Zimring, 2008; Cohen-Cole et al., this issue; Hjalmarsson, this issue; Kovandzic et al., 2009). In fact, the only other panel data study that Muhlhausen referenced in his one-sided review of the literature was Ekelund et al. (2006), which concluded that single-victim homicides were deterred by capital punishment but multiple-victim homicides were not. Muhlhausen did not mention that one of the strongest findings in the Ekelund study was that the presence of a capital punishment regime led to *more* homicides, so in our reading, as we explain further below, the Ekelund study actually undermines rather than supports the deterrent effect of the death penalty.[7]

Moreover, the most comprehensive assessment of the impact of the death penalty using the latest data—a recent paper by Kovandzic, Vieraitis, and Boots ("KVB")—has just concluded that there is "no empirical support for the argument that the existence or application of the death penalty deters prospective offenders from committing homicide." Table 2 provides a capsule summary of the six panel data studies just mentioned, of which

---

7. In their meta-analysis of death penalty studies, Yang and Lester (2008) list the Ekelund study as showing that capital punishment leads to an *increase* in homicides.

Table 2. Six Panel Data Studies on Impact of Executions on Murder Rates

(1) Study	(2) Geographical unit of analysis (time period of analysis)	(3) Original OLS estimates	(4) Corrected or expanded OLS estimates	(5) Original 2SLS estimates	(6) Corrected 2SLS estimates and evaluation	(7) Instruments	(8) Assessment of instrument validity
(1) Dezbakhsh, Rubin, Shepherd (ALER, 2003)	County-level (1977–1996)	None	Inconsistent— we generated OLS estimates based on original DRS data and specification minus the instruments.	Negative (stat. significant)	Negative (not significant, with clustered standard errors). Invalid if the instruments are deemed invalid (see next two columns).	(1) Police spending, (2) Judicial spending, (3) Prison admissions, (4) % voting for a Republican president; all instruments are state-level	Not valid

(continued)

Table 2. Continued

(1) Study	(2) Geographical unit of analysis (time period of analysis)	(3) Original OLS estimates	(4) Corrected or expanded OLS estimates	(5) Original 2SLS estimates	(6) Corrected 2SLS estimates and evaluation	(7) Instruments	(8) Assessment of instrument validity
(2) Mocan and Gittings (ALER, 2003)	State-level (1984–1997)	Negative (stat. significant)	Negative and insig. w/o coding errors; positive and insig. using Zimmerman execution ratio	Negative (stat. significant)	Negative (not significant, with clustered standard errors). Invalid if the instruments are deemed invalid (see next two columns).	(1) Deterrence variables $t-2$, (2) Death penalty law $t-1$, (3) Death penalty law $t-2$	Not valid
(3) Zimmerman (J. Appl. Econ, 2004)	State-level (1978–1997)	Negative (not significant)	Zimmerman reaffirms this finding of no significance in his contribution to this issue	Negative (stat. significant)		Proportion of murders committed at t and $t-1$: (1) by strangers, (2) with nonfelony murder circumstances, and (3) by non-white offenders; (4) indicator for release from death row at $t-1$, (5) indicator for "botched execution" at $t-1$	(1)–(4) not valid; (5) weak

(4) Katz, Levitt, and Shustorovich (ALER, 2003)	State-level (1950–1990)	Mostly negative (some stat. significant)	We expand the data to 1934–2000; results are mixed (but only stat. significant results are positive for 1934–1960.)
(5) Ekelund, Jackson, Ressler, and Tollison (Southern Economic Journal, 2006)	State-level (1995–1999)	Positive for all murders (stat. significant)	Too short a time frame to be reliable, and no state fixed effects
(6) Kovandzic, Vieraitis, and Boots (Criminology and Public Policy, 2009)	State-level (1977–2006)	Mostly negative (never stat. significant)	

Notes: Corrected estimates are based on Donohue and Wolfers (2005) and this paper. "Negative" estimates imply that each execution deters murder. "Positive" implies the death penalty regime is linked with an increase in the number of murders. Shading indicates statistical significance.

The first three studies offer evidence of deterrence with 2SLS models (and, in the case of Mocan and Gittings, with OLS models). These findings are all suspect for the reasons listed. The fourth study by KLS based on OLS models concludes: "there is little evidence in support of a deterrent effect of capital punishment as presently administered." A more extensive exploration in the sixth study by KVB reaches the same conclusion. The Ekelund et al. study purports to show deterrence for single murders and increases in multiple murders, but when correctly interpreted shows the death penalty is associated with *higher* levels of both types of murder. The Ekelund results are likely spurious, owing to the short time frame and the lack of state fixed effects.

three provide support for the deterrence hypothesis and three undermine this hypothesis.

With so much conflicting evidence being bandied about (or ignored when convenient), and with experts telling Congress that studies support the deterrent effect of the death penalty when they actually refute that claim, it may be useful to provide an overall assessment of the latest literature with a view to understanding why different researchers have reached such divergent conclusions while frequently relying on the same U.S. homicide data. Hopefully, this paper can give some guidance both to researchers interested in studying capital punishment and to legislators, policymakers, and academics who may be confused by such conflicting assessments of its likely impact on murder.

It will be helpful to provide a brief roadmap to Table 2. One can see from the second column whether the particular study used state or county data (only DRS relied on county data), and what years were analyzed. Note that of the five studies that examine more than five years of data, only KVB present data after 1997 (their dataset runs from 1977 to 2006). The next column notes that five studies presented OLS estimates, indicating whether these estimates are positive (suggesting antideterrence, as in the Ekelund paper) or negative (suggesting deterrence, as in MG), with significant findings identified by a shaded box. The fourth column provides additional information or corrections to the original OLS estimates, starting with our effort to provide OLS estimates drawing on the structure of the 2SLS DRS models (row 1—these are our estimates since DRS did not present OLS estimates), and continuing down to our efforts to correct some MG coding errors, extend the KLS time period, etc. The bottom line of column 4 supports the column 3 finding of KVB that there is no OLS support for a deterrent effect of capital punishment.

The next four columns of Table 2 (columns 5–8) summarize the three studies that provide 2SLS estimates. Column 5 shows that all three studies present negative and significant estimates (suggestive of deterrence), but column 6 indicates that they all become insignificant if one clusters the standard errors. Column 7 lists the instruments that are employed in the three studies, and column 8 reveals that in all but one case, the instruments are invalid. Of course, invalid instruments cannot be expected to yield valid estimates.

2. Using New York's Death Penalty Experiment to Illustrate Some Modeling Complexities

Many complex issues lurk in the background as one tries to specify the appropriate model for estimating the impact of capital punishment. An example exploring the depth, accuracy, and geographic precision of the information available to potential murderers may be instructive. Crime in the late 1980s and early 1990s rose sharply in New York during the initial crack epidemic and then started to turn down in about 1992. Republican George Pataki managed to unseat Mario Cuomo in 1994 in part on the pledge that he would restore the death penalty to the state, which he succeeded in doing in 1995. This New York death penalty statute remained in effect for a decade until it was declared unconstitutional by the highest state court in 2004. New York's recent dalliance with the death penalty never resulted in an execution. While politicians and think-tank advocates, such as David Frum of the American Enterprise Institute, have made the unsupported claim that the death penalty statute played a substantial role in New York's renowned drop in murders (Frum, 2006), the New York experience offers insight into the modeling choices involved in trying to estimate the true impact of capital punishment.

2.1. Endogenous State Adoption of a Death Penalty Statute?

The New York story immediately raises an endogeneity concern: the state, and indeed the nation, had endured a sharp increase in crime in the late 1980s and early 1990s, which aided Pataki's bid to unseat death penalty opponent Mario Cuomo. (The high national crime rate likely contributed to the 1994 Republican landslide that also defeated Texas Governor Ann Richards, and ushered in a Republican majority to both houses of Congress.) By 1995, of course, the sharp drop in crime was well underway, but there is a danger that a regression will spuriously attribute to New York's 1995 death penalty law this mean reversion in the level of crime. Before trying to assess the best approach to modeling the elements of the death penalty to which potential murderers might respond, note that these choices can be influential in estimating the impact of capital punishment: New York's very substantial murder rate decline continued after 1995, and, since all of the Table 2 regression studies weight by population, whether New York is counted as a treatment state (based on its 1995 law) or a control (based on its

lack of executions) can be important. Obviously, if one models New York's unusually large post-1995 murder rate decline as being influenced by the death penalty, then the deterrence argument will be strengthened.

2.2. What Do Potential Murderers Know?

The standard economic model assumes that individuals respond to prices, which they are able to estimate with reasonable precision. For consumer purchases, price information is readily available, and the cost must be paid with certainty if the product is to be obtained. In such cases, prices can enter the demand equation, and there is no need to introduce the complexities of psychological or information factors.

In the case of the death penalty, however, there is considerable uncertainty about the expected risk of execution. The econometrician needs to employ a proxy for this expected risk that captures the information available to and relied upon by potential murderers. For example, would potential capital murderers be aware of the presence of the New York death penalty statute (and therefore be deterred when the law took effect) or would they only learn of the law when, or fail to credit the possible sanction until, death sentences were handed out or convicts were executed? If passage of the law with great fanfare—recall that this was a major part of Pataki's successful gubernatorial campaign—discouraged criminals from committing murder, then a binary identifier of the state legal capital punishment regime (the law dummy approach) would be appropriate.

But how much else do potential killers know? Would they know that certain county prosecutors in New York opposed the death penalty, rendering the risk of execution in those areas virtually zero? Specifically, Manhattan District Attorney Robert Morganthau and Bronx District Attorney Robert Johnson were adamantly opposed to the death penalty, and made their view quite clear before the death penalty law took effect. Writing in the *New York Times* in February 1995, Morganthau stated:

> People concerned about the escalating fear of violence, as I am, may believe that capital punishment is a good way to combat that trend. Take it from someone who has spent a career in Federal and state law enforcement, enacting the death penalty in New York State would be a grave mistake.

> Prosecutors must reveal the dirty little secret they too often share only among themselves: The death penalty actually hinders the fight against crime.

. . . It exacts a terrible price in dollars, lives and human decency. Rather than tamping down the flames of violence, it fuels them while draining millions of dollars from more promising efforts to restore safety to our lives.

Some crimes are so depraved that execution might seem just. But even in the impossible event that a statute could be written and applied so wisely that it would reach only those cases, the price would still be too high.

It has long been argued, with statistical support, that by their brutalizing and dehumanizing effect on society, executions cause more murders than they prevent. "After every instance in which the law violates the sanctity of human life, that life is held less sacred by the community among whom the outrage is perpetrated."

Despite Morganthau's pleas, New York State went on to adopt a death penalty statute shortly thereafter. As a judge from New York's highest court later noted:

The very same day the legislation was signed into law by the Governor, the Bronx County District Attorney issued a press release purporting to "make [his] policy clear regarding the exercise of [his] discretion" to impose the death penalty. In this statement, the District Attorney expressed deeply felt concerns regarding the effectiveness and administration of the death penalty. He concluded by stating, 'For all these reasons, while I will exercise my discretion to aggressively pursue life without parole in every appropriate case, it is my present intention not to utilize the death penalty provisions of the statute.' On November 2, 1995, the District Attorney was reelected by approximately 89% of Bronx County citizens who voted. *Johnson v. Pataki*, 91 N.Y.2d 214, 691 N.E.2d 1002, 668 N.Y.S.2d 978 (1997), Smith, Judge (dissenting).

Despite this articulated opposition to the death penalty, from 1995 to 2004 the murder rate dropped in Manhattan by 64.4% (from 16.3 to 5.8 murders per 100,000), and in the Bronx by 63.9% (from 25.1 to 9.1 per 100,000).[8] Another New York City borough with the identical laws and police force, and with broadly similar economic, social, and demographic features as Manhattan and the Bronx—Brooklyn—had a top prosecutor who issued the largest number of notices of intention to seek the death penalty (albeit with no executions) (Kuziemko, 2006). Yet Brooklyn experienced only a 43.3%

8. In the rest of New York State (excluding Manhattan and the Bronx), the murder rate fell by only 36.5% (from 6.5% to 4.1%).

decline in murders over this period, from an initial figure (almost identical to Manhattan's) of 16.6 murders per 100,000 in 1995 down to only 9.4 in 2004. Just as Frum's claim that New York's murder drop resulted from the passage of the death penalty law in 1995 is unfounded, one cannot draw strong causal inferences from the fact that Manhattan and the Bronx led the way in the decline in homicides, but certainly, there is not even a hint of a deterrent effect of capital punishment in the crime patterns across these counties with anti-death-penalty prosecutors.

For our purposes, the important points are the modeling complexities: Was any of the two-thirds drop in the Manhattan and Bronx murder rates the result of New York's death penalty law, or did potential murderers understand and rely upon the anti-death-penalty pronouncements of the DAs in both counties? As the next subsection shows, these complexities are often swept under the rug by certain specification choices.

2.3. The Law Dummy Model

Table 2 focuses on studies that have modeled the impact of the death penalty with a variable that in some form counts the number of executions in a given state and year. It is possible, though, that the simple presence of a state law authorizing capital punishment is enough to deter. Consider two panel data studies that have sought to test this proposition by running a regression with a "law dummy" indicating the presence of a capital punishment law. First, using data for 1960–2000, Dezhbakhsh and Shepherd (2006) reported that the coefficient on the law dummy was negative (that is, finding deterrence) and statistically significant. DW showed, however, that this result became insignificant when the standard errors were adjusted by clustering and when year fixed effects were added to the regression (as is standard) (see Table 2 of Donohue and Wolfers, 2005).

Second, as indicated in the first column of Table 3, MG find that death penalty laws have had a statistically significant dampening effect on murder, using the 1977–1997 data. While, as Section 3.2 below reveals, DW established that the MG OLS results claiming *executions* lead to fewer murders go away if MG's coding errors are corrected, this same correction had only a small dampening effect on the negative and statistically significant estimate that MG present in their death penalty indicator model (compare column 1 to column 2 in Table 3). But Table 3 illustrates the importance of New York to MG's apparent deterrence conclusion, since dropping

Table 3. Impact of Legalized Death Penalty: Testing the Sensitivity of Mocan and Gittings' Results (1977–1997)

	(1) Original MG estimates	(2) Corrected estimates	(3) NY observations dropped	(4) NY death penalty indicator set to 0
Death penalty legal	−0.154**	−0.151**	−0.0021	−0.0018
Indicator ($t - 1$)	(0.0061)	(0.0063)	(0.0048)	(0.0073)

Notes: Standard errors are robust and clustered at the state level. Corrected estimates based on the discussion in Donohue and Wolfers (2005). ** indicates statistical significance at the 0.05 level.

While MG's estimate of the impact of a valid death penalty law is negative and significant (column 1), even when corrected for coding errors (column 2), these results are entirely dependent on the state of New York. If New York state is simply dropped from the analysis (column 3), the coefficient drops by two orders of magnitude and becomes insignificant. Column 4 yields results similar to column 3 when New York is treated as not having a death penalty law (the state never executed anyone during or after this sample period, the DAs in Manhattan and the Bronx, which enjoyed enormous drops in crime over this period, were staunch opponents of capital punishment, and the law was ultimately ruled unconstitutional).

New York decreases the estimate by two orders of magnitude and eliminates the finding of statistical significance (column 3). Column 4 shows virtually the same result as column 3 if we code the New York legal dummy as zero for the years 1995–1997 (in essence, positing that potential murderers correctly realized that there was virtually no risk of execution in New York despite the passage of the statute). In other words, the big drop in murders in New York in the mid-1990s, led by Manhattan and the Bronx, whose DAs strongly articulated opposition to the death penalty, is what drives the finding that the presence of a death penalty law is correlated with lower crime.

Finally, KVB has run the dummy variable model on the longest time period of 1977–2006 and has found that the death penalty is never significant at the 0.05 level, although it is significant at the 0.10 level if clustering is not needed (again attributing the New York decline in murders from 1995 to 2004 to the death penalty law).

3. OLS Estimates of the Impact of Executions on Murder

We now turn to the Table 2 studies that model the impact of capital punishment with some measure of the frequency of executions. The table immediately reveals the key methodological difference between the three studies that support the deterrence hypothesis and the three that do not: only the studies that present 2SLS estimates—DRS, MG, and Zimmerman—claim to find deterrence, while the studies that rely only on OLS—KLS,

Ekelund et al., and KVB—show no evidence of deterrence. The second column of Table 2 summarizes the OLS results provided in five of the six studies (DRS did not show OLS results). Column 3 shows that after correction or expansion of the data period, there is no support for the deterrence hypothesis in the OLS estimates across any of the six studies. We discuss the OLS results of the six studies in turn.

3.1. Dezhbakhsh, Rubin, and Shepherd—DRS

DRS did not present OLS results, instead relying on a 2SLS approach that we discuss in greater detail in Section 4. Using DRS data and specifications but without instrumenting, we obtained OLS estimates of the impact of executions on murder rates. The results were unstable and inconsistent. Specifically, for DRS's six models, we generate one OLS estimate that is statistically significant and *positive* (antideterrent) and two that are significant and *negative* (deterrent). The other three estimates are negative but insignificant.[9]

Even if these results were not conflicting, there would be little reason to credit them in light of the fact that the DRS specification (that is, their 2SLS estimation, which we converted into an OLS estimation) does not control for either the number of police or the extent of incarceration in assessing the impact of executions on murder, even though a vast literature suggests that more police and higher levels of incarceration reduce the murder rate. This choice is particularly problematic given the fact that Texas enjoyed an extremely sharp drop in crime in part because of an enormous increase in incarceration, which coincided with the increase in executions. In general, any empirical study of murder that does not control for the incarceration rate risks suffering from major omitted variable bias, and should not be

---

9. Table 1 of DW established that the effort by Dezhbakhsh and Shepherd to examine the effect of death penalty abolitions and reinstatements, which purported to show evidence of crime increases from abolition and crime drops from reinstatement, led to no such conclusion if one compared the changing states (the treated group) with the control group of nonchangers. DW also showed that the claim by Cloninger and Marchesini that the Illinois and Texas death penalty moratoria led to an increase in homicide was the product of their decision to examine *growth rates* in homicide. When we replicated their analysis using *levels* of homicide, the evidence of an unusual jump in murders disappeared. (Indeed, in Illinois this approach suggested that the moratorium was accompanied by a statistically significant *decline* in the murder rate.)

Table 4. Reanalyzing MG's OLS Estimations of the Impact of Executions on the Murder Rate (1984–1997)

| | Corrected MG OLS Results | | | |
| | Dependent variable | | | |
	Murder rate	Robbery rate	Violent crime rate	Burglary rate
Execution ($t - 1$) per death sentence ($t - 7$)	−0.50 (0.34)	−3.14 (7.71)	−4.69 (16.13)	4.60 (19.83)
Execution ($t - 1$) per death sentence ($t - 1$)	0.03 (0.14)	5.11 (4.44)	9.81 (11.70)	2.94 (16.06)

Notes: Standard errors are robust and clustered on the state level. Basic specifications are from MG. Corrections to MG's OLS estimates are outlined in Donohue and Wolfers (2005). MG employed a dummy for Oklahoma in 1995 to control for the Oklahoma City bombings and this indicator is not included in the robbery and burglary regressions. While MG achieves a negative (albeit insignificant) coefficient on the execution variable for the murder rate equation, the fact that similar negative coefficients emerge in the robbery and violent crime equations suggests that executions are proxying for some other impact dampening violent crime. When we follow Zimmerman and lag the death sentence only 1 year, the estimated effect on murder becomes positive, as it is for all other crimes (row 2).

relied upon.[10] While DRS (like Lott and Mustard in their work on guns) followed the original model of Ehrlich in the unfortunate choice of omitting incarceration as an explanatory variable, the DRS study and the Ekelund study are the only two studies in Table 2 marred by this particular flaw.

3.2. Mocan and Gittings—MG

At first glance, the MG paper appears to make a persuasive case for the deterrence hypothesis. MG finds that executions are linked with statistically significantly *lower* rates of murder. MG then seek to confirm that the death penalty is not simply a proxy for some other anticrime influence by showing that when their model is run on other crimes, such as robbery, rape, and two property crimes, the effects are no longer statistically significant. (For the two property crimes the effect is positive, and for the two violent crimes it is negative). But DW found that MG's OLS results were the product of a coding error (see DW, Table 6). When the coding error is corrected, the statistically significant effect in the MG OLS regressions disappears, as indicated in the first column and first row of Table 4.

---

10. While controlling for incarceration, Moody and Marvell (2009) report that executions had no statistically significant impact on murder in their county data analysis for the period 1977 through 2000.

Despite the lack of significance, should the negative sign be taken as some evidence of deterrence? Two reasons undermine even this weak conclusion. First, as the first row in Table 4 indicates, MG followed DRS in specifying the execution variable as a ratio of executions last year to death sentences seven years ago (to predict today's murders). While we find the ratio of executions to death sentences to be an unconvincing explanatory variable (as discussed in Section 4.1), if it is to be used, we agree with Zimmerman's argument that a better measure would be the ratio of executions to death sentences lagged one year (since potential criminals will respond to the most recent evidence on sanctions). As Table 2 indicates and row 2 of Table 4 depicts, when we follow the Zimmerman approach and posit that potential criminals would base their decisions about criminal homicide on more recent data by using the once-lagged ratio of executions to death sentences, the MG murder coefficient estimate turns positive (yet insignificant), consistent with the view that the death penalty *increases* murder (see DW, Table 6, panel C).[11]

Second, as MG showed and we document in the first row of Table 4, the coefficient on the execution variable generated a similar negative sign in "explaining" robbery and violent crime. The similar signs on the execution variables in these violent crime equations suggest that the death penalty is likely proxying for other factors that cause overall violent crime to fall (or which correlate with such declines).[12] Such an effect would be consistent with our discussion above of omitted variables bias in the DRS study. However, when we again follow the Zimmerman approach and lag death sentences only by one year, we see that executions positively correlate with all crimes, including murder (although nothing is statistically significant—see row 2 of Table 4).

---

11. When we employ the same lag structure in the DRS 2SLS model, however, the estimates jump in the opposite direction: the coefficient on predicted executions_t – 1/death sentences_t – 1 is significant at the 5% level in the second stage, and the implied lives saved jump from eighteen to an eye-popping eighty-five. Thus, one sees how sensitive these models are in that identical changes in lag structures lead to very different results in different studies.

12. Conceivably, one might argue that the death penalty could have an indirect dampening effect on violent crime, but this would typically not be a traditional argument based on the criminal responding to the risk of capital punishment when considering whether to commit a crime since no crime other than an aggravated murder would be eligible for the death penalty.

3.3. Zimmerman

Zimmerman concluded that none of his OLS estimates of the impact of executions on murder for 1978–1997 reached statistical significance. In Zimmerman's latest work (in this issue), he adds a lagged dependent variable to his model and again computes OLS results for the same time period. He concludes: "the OLS specifications . . . again provided no evidence to suggest a deterrent effect of capital punishment, which is consistent with Zimmerman's original results."

3.4. Katz, Levitt, and Shustorovich—KLS

KLS only presented OLS results, which were estimated on state data from the period 1950–1990. While the presence of some sporadic statistically significant KLS estimates led Joanna Shepherd and Paul Rubin to testify before Congress that the refereed studies unanimously supported the view that the death penalty deterred murder, KLS concluded their own study as follows: "there is little evidence in support of a deterrent effect of capital punishment as presently administered."

An appealing aspect of the KLS study is that it was designed to probe whether prison conditions influenced crime rates and simply used the execution rate as a control. As a result, there is little reason to fear that the authors selected their regression approach to make a particular point about the deterrent effect of the death penalty. To see whether the KLS conclusion would be robust to various modifications and extensions, we added additional years of data going back to 1934 and coming forward through 2000. Table 5 describes the various permutations, starting with a simple expansion of the exact KLS model to extend an additional ten years through 2000. This data expansion eliminated the sporadic statistically significant evidence of deterrence noted by Shepherd and Rubin.

The first row of Table 5 shows the particular base model, devised for KLS's exploration of the impact of prison harshness, as proxied by death rates in prison. In this row, the impact of executions on murder is estimated using a ratio of executions per prisoner in the state's penal system. The next two panels of Table 5 describe the results when two other specifications of the execution variable were used: (i) the ratio of executions to lagged murders, and (ii) the ratio of executions to population. Not one of the regressions summarized in Table 5 showed a statistically significant deterrent effect,

Table 5. The Impact of Executions on Murder: Donohue and Wolfers' Modifications and Extensions of KLS's OLS Regression Models and Data

Study (modification)	Geographical unit of analysis (time period of analysis)	OLS estimates
Katz, Levitt, and Shustorovich, using Executions per Prisoner (extending end period of 1990 to 2000)	State-level (1950–2000)	Mostly negative (not significant)
Katz, Levitt, and Shustorovich (using Executions per Lagged Murder instead of KLS's Executions per Prisoner)	State-level (1934–2000)[a]	Positive (not significant)
	State-level (1934–1960)[a]	Positive (not significant)
	State-level (1961–2000)	Mostly negative (not significant)
	State-level (1961–2000)[a]	Mostly negative (not significant)
Katz, Levitt, and Shustorovich (using Executions per Capita instead of KLS's Executions per Prisoner)	State-level (1934–2000)[a]	Mostly positive (not significant)
	State-level (1934–1960)[a]	Positive (some stat. significant)
	State-level (1961–2000)	Inconsistent (not significant)
	State-level (1961–2000)[a]	Negative (not significant)

Notes: "Negative" estimates imply that each execution deters at least one murder. Shading indicates statistical significance.
[a]Dropping the KLS controls for infant mortality and the insured unemployment rate owing to data unavailability for the 1934–1949 period.

although for 1934–1960 a statistically significant *positive* effect emerged (suggestive of antideterrence). Based on the additional findings summarized in Table 5, KLS's conclusion that there is little evidence in support of a deterrent effect can now be strengthened: there is *no* statistically significant evidence using the KLS model (and our two primary modifications of it) of any deterrent effect of capital punishment over the period of 1934–2000 (or in any of the subsamples shown in Table 5).

3.5. Ekelund et al.

The Ekelund study, unlike the other studies listed in Table 2, is based on a very short span of data for the highly unusual years 1995–1999. During this period, crime was dropping sharply throughout the nation in ways that are not well captured by the standard crime models, so it is likely that the paper is marred by omitted variable bias. Because they used such a short time period for their analysis, Ekelund et al. did not use state and year fixed effects, thereby foregoing one of the greatest advantages of panel data. Moreover, like the DRS study, the Ekelund study is marred by its omission of a control for the influence of incarceration on crime, and, like DRS, MG, and Zimmerman, it also relies (in five of its eight specifications) on the problematic ratio of arrests to murders, as a pseudo-probability of arrest. Consequently, the estimates to emerge from such an unusual study estimated on a truncated data period will be far less reliable than a study conducted with better controls over a longer period during which one can at least hope the unexplained swings in murder rates will tend to average out. But while the results of this study are likely spurious, it has been cited to the Congress as supporting deterrence, when it actually should either be ignored or taken as evidence against deterrence.

Oddly, while Ekelund et al. conclude that capital punishment deters single murders (but not multiple murders), their paper really suggests on its face that the death penalty leads to *more* murders of all kinds.[13] The

---

13. Even the touted conclusion of the Ekelund study that single-victim homicides were deterred by capital punishment but multiple-victim homicides were not is surprising since most single-victim homicides are not subject to the death penalty while most multiple killings are. Ostensibly unaware that their study largely undermines the deterrent effect of the death penalty, Ekelund et al. try to explain this anomaly by stating that "only the first premeditated murder is subject to penalty—execution. Killings beyond the first are, in effect, free (p. 525)." Again, most individual murders will not be subject to execution,

misinterpretation stems from the authors' failure to appreciate fully that while they found that executions lead to a small decrease in (single) murders, their estimates also showed that capital punishment laws lead to large *increases* in murder. In fact, the large estimated pro-murder effect of the death penalty law outweighs the small execution effect. Specifically, a state would need between seventeen and thirty-nine executions by lethal injection per year just to get back to the level of murder that would have been experienced if the state had no death penalty regime at all![14] Thirty-nine executions would be more than even Texas, as the most active executing state, ever had during the 1994–1999 study period (peaking at only thirty-seven executions in 1997). In other words, the Ekelund study actually undermines rather than supports the deterrent effect of the death penalty.[15]

3.6. Kovandzic, Vieraitis, and Boots—KVB

KVB have provided the latest assessment of the death penalty, with a careful OLS analysis of 1977–2006 state panel data, which is the most

---

but certainly if you have killed more than one individual you are both more likely to get caught and also more likely to be executed, so the underlying theory of the Ekelund study would seem to be flawed.

14. Ekelend et al. show that a death penalty statute leads to a substantial increase in the murder rate (the coefficients range from 0.103 to 0.273), while executions have a small dampening effect of from –0.006 to –0.007 per execution. Since every state with executions also has a death penalty law, the true impact of the death penalty is based on the combined effect of these two variables, and the large *increase* in murders outweighs the small decrease.

15. Perhaps the death penalty looks better in the four states that used electrocution over the Ekelund study period? Again, no. The study offers estimates of the combined impact of a death penalty law, a state using electrocution, and each execution. Since these figures were all *positive* and highly significant in the multiple murders model, electrocutions apparently increase multiple homicides! For single murders the Ekelund estimates of the component effects are mixed, so one needs to do some calculations to assess overall affects. If the estimated coefficients for Ekelund's three models were meaningful, a state would need to *electrocute* twenty-six, twenty, or three inmates (respectively) in a year to no longer show an *increase* in murders relative to states with no death penalty laws. Only two of the four electrocuting states executed as many as three inmates over this period—Alabama executed three in 1997, and Florida electrocuted three in 1995 and four in 1998. This means that in two of the three Ekelund models, electrocution was uniformly associated with *higher* murder rates. The outlier third model, which only required three electrocutions to be suggestive of deterrence, differed from the other two by introducing the flawed "pseudo-probability of arrest" variable, which has murders in the current year (that is, *the dependent variable*) in the denominator of this explanatory variable, leading to clear ratio bias.

recently available data. The key features of their estimation are that they relied on state and year fixed effects with linear state trends, and they clustered their standard errors to correct for serial correlation. Using seven different base specification models to estimate the effect of executions on murder, which were then subject to extensive robustness checks, KVB found no support for a deterrent effect of the death penalty.

KVB then went on with an extensive sensitivity analysis of these seven base models, and across their sixty-six primary regressions estimating the impact of executions, only two were significant at the 0.05 level—one was positive and one negative. The one negative and significant finding (suggesting possible deterrence) was generated using a contemporaneous measure of the level of executions and no controls for state trends, although in five other estimates without state trends (including the model with "executions over lagged homicides" for the key variable) there was no significant effect.[16] Overall, KVB make a powerful case that OLS estimates on state data for 1977–2006 simply do not support the proposition that state capital punishment laws or executions have a dampening effect on homicides.

4. Causal Inference Using 2SLS

The discussion so far has shown that across quite a range of different models and different years, panel data OLS estimates using state data (and, in the case of DRS, county data) consistently show no evidence of a deterrent effect of capital punishment. But a major concern when estimating the impact of a law or policy is that the same factors that led to changes in capital punishment also directly affected the murder rate. To some degree, this problem is mitigated by well-constructed panel data models that have rich controls as well as fixed effects for time and space. Fearing that the death penalty law or the number of executions would still not be conditionally exogenous, scholars such as DRS, MG, and Zimmerman have estimated the impact of the executions using 2SLS.

---

16. The contemporaneous execution model, specified in levels, seems a bit odd, since it suggests that one execution in a massive state would impact murder as powerfully as one execution in a small state. Using a more reasonable specification such as executions per homicide or executions per population, generates no statistically significant impact.

Unfortunately, the three Table 2 2SLS studies all have striking problems. As Section 4.1 will describe, these three studies adhere to the faulty structure of the original Ehrlich study, and indeed magnify the problems of the Ehrlich study when they try to replicate in panel form what Ehrlich had done in his time-series analysis. Section 4.2 then shows that all three studies rely on invalid instruments, which are commonly included in various studies as explanatory variables in the second stage rather than being instruments that are excludable from the second stage. Section 5 then shows that the 2SLS estimates suggestive of deterrence are both highly unreliable and theoretically dubious, which is perhaps not surprising given that their instruments fail to meet the test of excludability.

4.1. The Problematic Specification Used by Ehrlich and His Followers

Although DRS, MG, and Zimmerman use panel data models to estimate the impact of the death penalty, they expressly follow the structure of Ehrlich's time-series model in controlling for the arrest rate, conviction rate, and execution rate. There are three problems with this approach: (i) Ehrlich's various ratios, which were designed to reflect the relevant deterrence probabilities, are inaccurate and in fact are *not* probabilities since they commonly exceed one or are undefined; (ii) Ehrlich's murder conviction rate variable, though flawed and arguably inadequate for his national time-series analysis, is still vastly superior to the "death sentences to murder arrests" variable that DRS, MG, and Zimmerman used as a proxy for the conviction rates; and (iii) the estimates are sensitive to the complicated and unpersuasive lag structures used to create the deterrence "probabilities." These issues will be discussed in turn.

4.1.1. The Ehrlich deterrence ratios.   Ehrlich stated that a criminal contemplating murder would be interested in three probabilities, which he proxied with national data on the murder arrest rate, the conviction rate for murderers who are arrested, and the execution rate for those convicted. These rates are intended to reflect the probabilities of adverse outcomes facing potential murderers, but instead they are three linked ratios, with complex lag structures, of arrests/murders, convictions/arrests, and executions/convictions. This approach has created considerable difficulties for the various studies that have followed Ehrlich, who, for data availability

reasons, now proxy the second two ratios with death sentences/arrests, and executions/death sentences.

The ratio of arrests to murders is immediately problematic. First, the ratio is undefined if there are no murders in a given year. Should this case be treated as zero arrest rate or a perfect arrest rate? This is a particular problem for DRS, since their ratio of "county arrests to murders" will frequently have a zero denominator. Moreover, temporal mismatch between year of arrest and the year of the crime can also improperly elevate the measured arrest "probability" in years of crime decline (or depress it in years of increasing crime).

Second, the ratio of arrests to murders fails as an explanatory variable when there are multiple victims. For example, take the case of someone who murders a number of individuals before killing himself. Since this incident would lead to no arrest (the perpetrator killed himself) and many dead victims, what Ehrlich and his followers would deem to be a zero probability of arrest is linked with lots of murders. But the idea that a mass killer would believe he faced no chance of arrest seems quite wrong. Of course, the more multiple murders there are, the worse the arrest rate looks, even if the murderers are all apprehended.[17]

Third, the converse problem with the "arrest ratio" occurs when there are multiple offenders per murder. This falsely makes the arrest rate look better since multiple arrests bump up the ratio of arrests to murders (frequently beyond one, underscoring that Ehrlich and his adherents are *not* using true probabilities). Indeed, erroneous or improper arrests also serve to artificially inflate Ehrlich's arrest "probability." In DRS's study, 8727 county–year arrest rate observations are greater than one (27.1% of the nonmissing observations), and another 8944 of these observations are exactly 1 (28% of

---

17. Similarly, while Tim McVeigh faced a high probability of arrest and execution for bombing the Oklahoma City federal building, his case involved one arrest and 168 deaths, suggesting (in those studies following the Ehrlich approach) a low probability of arrest for murder for Oklahoma county in 1995. Since DRS measured these variables contemporaneously, the 1 in 168 "probability" of arrests is deemed to explain the 168 murders. In general, one would expect negative ratio bias to influence the correlation between murder rates and arrest rates since murders appear simultaneously in both the numerator of the left-hand side variable and the denominator of the right-hand side variable. (Note that MG dummy out Oklahoma in 1995.)

the nonmissing). Overall, the unweighted mean arrest rate in DRS's county sample is 1.01, with 10% of the observations being two or greater.[18]

4.1.2. The lack of conviction data.   Ehrlich states that his conditional probability of conviction given a murder charge "is estimated by the fraction of all persons charged with murder who were convicted of murder in a given year as reported by the FBI UCR." This is not entirely correct: rather than reporting a national total for this variable, the FBI's Uniform Crime Reports (UCR) only reported conviction data for a relatively small and changing sample of cities. For example, for 1953, the UCR only reported convictions for 197 cities with populations of over 25,000, for a total population of roughly 25 million.[19] Since the UCR has now stopped reporting these conviction numbers, Ehrlich's followers (DRS, MG, and Zimmerman), who in any event needed panel data on this variable, have replaced Ehrlich's problematic "probability of conviction if charged with murder" variable with an even worse measure: the ratio of death sentences to arrests. Again, there will be a problem whenever the denominator is zero, creating the issue of whether this should be treated as zero, missing, or something else (perhaps looking back to the prior positive value if any)? But more fundamentally, this ratio does not well capture deterrence pressures on criminals: one can imagine that a state that convicts and sentences every arrested murderer to a term of years but never uses capital punishment would generate much more deterrence than a state that convicts only a small fraction of its murderers while sentencing a number to death. Ehrlich tried to control for this probability of conviction (albeit poorly), but the followers of Ehrlich have no effective control for conviction rates.

4.1.3. The complex lag structure of the deterrence "probabilities."   Both DRS and MG (unlike Zimmerman) impose a complex lag structure on their

---

18. The problem with the arrest rate is also severe in the state data—in the MG data, the arrest rate is larger than 1 for 26.3% of the nonmissing observations.

19. The number of cities included in the convictions reported by the UCR ranged from 13 cities (all 100,000 or more population for a total population of 9,369,010) in 1936—the first year for which the UCR included information on persons found guilty—through 3025 cities in 1970, representing a population of 68,897,000.

arrest rate, death sentence, and execution ratios. For example, rather than simply using a contemporaneous or once-lagged ratio of "executions to death sentences" (as Zimmerman does, although we would prefer a more meaningful ratio such as executions to murders or executions to population), DRS and MG instead use the ratio of last year's executions divided by the number of death sentences seven years ago to predict murder rates today.[20] As we noted in Section 3.2 above, the DRS and MG estimates jump wildly (and in opposite directions!—see footnote 11) when we follow the Zimmerman/KLS approach and posit that potential criminals would base their decisions about criminal homicide on more recent data by using the once-lagged ratio of executions to death sentences (see DW, Table 6, panel C).

4.1.4. Summary on the deterrence ratios.   The conceptual and practical problems that the three studies stumble into in trying to mimic Ehrlich's model are so daunting that it is hard not to conclude that it is unwise to go down that path. For example, a deterrence measure that is more conceptually appropriate than the ratio of arrests/murders might be the clearance rate for murder, which is available by state from the FBI over this period. This measure tries to correct for the extreme problems caused by multiple offenders and victims, and also corrects for situations where the case is solved but the offender is killed or commits suicide (which would prevent arrest but should not be taken as a failure of the criminal justice system as the arrest ratio would imply). But murder clearance rates have been trending down over the last thirty years in a way that is probably adequately captured by linear state trends, so the flawed arrest rate variable (used by DRS, MG, Zimmerman, and Ekelund) should probably be retired. Moreover, since conviction data are not available by state, the "death sentences to murder arrests" variable should also be jettisoned.

---

20. It is not clear that potential murderers would have any information on death sentences, whether last year or seven years ago. Conceivably, though, they could pick up this information from notorious trials and the sentences would remind them that the death penalty is being used in their state. This might suggest that the sentencing variable should appear as a separate explanatory variable and not as an odd—yet influential—pseudo-probability. Note that not all executions are covered in local papers, so if executions are more newsworthy than death sentences, it may be that information on death sentences would only be available to those in prison (or to acquaintances of the condemned). (Hjalmarsson, this issue.)

Finally, the "executions to death sentences" variable makes little sense, whether one dresses it up with complex lags or not. For example, a state that each year sentences twenty to death and executes ten would presumably be a much greater threat to murderers than a state that sentences one and always executes that one. Yet for this key "executions to death sentences" variable, the first state has a "deterrence" measure of 0.5 and the second state has a measure of 1, which is the opposite of what one would expect. We would recommend following the lead of the KLS study, which avoids the problems of the highly mismeasured pseudo-probabilities of DRS, MG, Zimmerman, and (in the case of the arrest/murder variable) the Ekelund study.

4.2. The Invalid Instruments of the 2SLS Studies

A growing literature has emphasized that invalid or weak instruments can lead to highly unreliable estimates, and we suspect the current 2SLS death penalty studies have confirmed this unfortunate lesson. The great difficulty in 2SLS estimation is to find a valid instrument that is both correlated with the number of executions (the first stage), but also "excludable" from the murder rate equation (the second stage). This second requirement means the instrument does not directly influence murder (except through its influence on executions) and is not a proxy for a variable that should be included in the murder rate equation but is not. Finding such instruments is never easy, and for the reasons summarized below, in our view, none of the instrumental variables employed by DRS, MG, or Zimmerman—which are set forth in column 7 of our Table 2 summary table—is convincing.

In addition to the problematic deterrence ratios and the invalid instruments, it is also important to notice that the results in all three of the 2SLS studies are statistically insignificant if one clusters the standard errors to account for serial correlation. DW showed that the DRS and Zimmerman results were not statistically significant if one adjusted their standard errors by clustering, and MG reported clustered standard errors, which also eliminated the statistical significance of their results at the 0.05 level.[21]

21. DRS not only failed to adjust for serial correlation by clustering by state but they also did not generate robust standard errors to deal with heteroscedascticity, which explains why their standard errors are so low and thus their estimated *t*-statistics are so high. Indeed, the extremely high *t*-statistics on the DRS execution variables, which ranged up to 19.5, provided an immediate tip-off that their standard errors were too low. Half

There is currently some debate as to whether such clustering is required, with Bertrand et al. (2004) insisting that it is, while others suggesting it is not (National Research Council, 2005). While most applied researchers are clustering the standard errors today in state and county panel data studies and virtually all researchers believe that some correction for serial correlation in panel data is needed, the issue of whether clustering generates approximately correct standard errors or in fact overcorrects the standard errors is currently being investigated.[22]

Some specific comments on these three 2SLS studies, and why we strongly doubt the validity of their instruments, are provided below.

4.2.1. Dezhbakhsh, Rubin, and Shepherd—DRS.

DRS relied exclusively on a 2SLS approach to argue that each execution saves many lives. Implementing the two-stage approach requires at least one exogenous variable in every equation one estimates. But consider the first DRS equation in this system that tries to estimate the probability of arrest that a murderer would face in a given *county*. The alleged exogenous variable here is *statewide* police expenditures (*not* adjusted for inflation and *not* per capita), which is sometimes criticized as exactly the sort of endogenous variable that the two-stage approach is trying to circumvent. In the second DRS equation estimating the *statewide* likelihood of receiving a death sentence once arrested, the alleged exogenous variables are the *statewide* Republican presidential candidate's vote in the most recent election, which may not be exogenous (when murders go up, Republicans tend to do better in the elections), and the raw *number* of *state* prison admissions, which is hardly likely to be exogenous to the level of crime. The third equation, which estimates the *statewide* likelihood of execution given a death sentence, relies on the supposedly exogenous variable "*statewide* expenditures on the judicial and legal systems" (*not* adjusted for inflation and *not* per capita). But the evidence suggests—see both

---

of their primary six regressions lost statistical significance with the robustness correction alone.

22. Our own research examining the impact of laws permitting the carrying of concealed weapons indicates that failure to cluster in county panel data analyses for 1977–2000 generates Type 1 error about 45–75% of the time with random assignment of these laws. Clustering reduces this error rate down to about 8–22% for these randomly assigned placebo laws. This would suggest that clustering is desirable and does *not* overinflate the standard errors in panel data crime models (Donohue et al., 2009).

Cook (in this issue) and Roman et al. (in this issue)—that states with lots of executions have to spend a great deal on appellate appeals (and death penalty trials), which means that the likelihood of getting a death sentence is influencing the expenditures instead of the proper causal pathway for a truly exogenous variable, in which the expenditures influence the likelihood of getting the death penalty.

DW discussed how the DRS instruments seemed not to work either in the first stage (see Table 9 of our original paper) or in the second stage, leading us to question whether these are credible instruments. Our intuition is supported by a Hausman test for overidentification, which rejects the validity of these instruments. While Rubin has tried to justify the DRS instruments on the grounds that they "have been used in numerous empirical papers" (Rubin, 2006), we have emphasized that the previous use of these instruments in fact underscores their *invalidity*:

> In separate papers, Rubin and coauthors have used the same instruments (or subsets of them) as providing variation in truth-in-sentencing legislation, firearms right-to-carry laws, sentencing guidelines and California's three-strikes law. It cannot be the case that these previous papers were correct in positing that these instruments affect homicides only through that array of channels, and that Rubin is correct that these instruments influence homicides through their effect on execution policy to the exclusion of other pathways. Yet without valid instruments, one cannot generate reliable results nor offer useful policy recommendations from an instrumental variables estimation. Their results turn out to be extremely fragile to the inclusion or exclusion of particular instruments (Donohue and Wolfers, 2005).

To illustrate once again how the DRS instruments can generate seemingly powerful yet nonsensical results, consider the following exercise in which we run three separate regressions, in each one using all the DRS instruments to instrument for one of three explanatory variables that are at times used in murder rate regressions: NRA membership rate, three-strikes laws, and state assault weapons bans, which are all measured at the state level.[23] This exercise leads to some strange and conflicting results on guns: state assault weapons bans led to major *increases* in murder (with a whopping *t*-statistic

---

23. In these regressions, which are based on DRS's preferred model four, we simply treat the arrest rate, sentencing rate, and execution rate as exogenous variables that appear in both the first and second stages of the four regressions.

of 11.24), but a 10% increase in NRA membership rate would lead to an increase of 0.7 murders per 100,000 people (backed up by a *massive t-*statistic of 6.06). In other words, the DRS instruments and model would imply that increasing national NRA membership by 10% would raise the 2006 murder rate from 5.7 to 6.4 per 100,000, implying an increase of over 2000 murders! These results raise further concerns that instruments that fail the excludability requirement can generate ostensibly strong, yet wholly unreliable, results for a range of explanatory variables.

One further indication of the unreliability of the DRS models is provided by a subsequent paper by Shepherd (2005). Using the identical approach to the DRS paper, but using the state-specific estimates (instead of the aggregate estimate of the effect of executions across all states), Shepherd reached a conclusion in considerable tension with the DRS study. Specifically, Shepherd found that executions deterred murders in six states, had no impact on murder in eight states, and *increased* murder in thirteen states. In other words, Shepherd's single change of generating state-specific estimates while using the identical data and otherwise identical specification of DRS contradicted the DRS finding of statistically significant deterrent effects in twenty-one of the twenty-seven states examined.[24] Consequently, the DRS finding that one can conclude, with 95% confidence, that each execution leads to roughly eighteen fewer murders should be taken as highly implausible.

4.2.2. Mocan and Gittings—MG. MG provide 2SLS estimates on the rationale that the presence of a capital punishment law may be endogenously determined by the past homicide rates within the state. MG instrument for the presence of a capital punishment law "with twice-lagged deterrence variables and two lags of capital punishment law." This approach is problematic. The basic assumption of their paper is that criminals are deterred by the once-lagged values of the "deterrence variables" and by the presence of a capital punishment law, which means that the once-lagged variables directly influence the rate of murder according to MG's approach. Thus, MG's identifying assumption is that the second lag turns these deterrence variables

---

24. Shepherd describes her results as follows: "The results are striking. Executions deter murder in a few states, have no impact in a few more, but *increase* murders in many more states than the number where there is deterrence" (Shepherd, 2005, p. 229, emphasis in original).

into valid instruments that are correlated with the presence of a capital punishment law but don't influence murder except through their effect on capital punishment law. But this strong assumption is hard to justify, particularly in light of MG's own description of their Table 4 models "in which the deterrence variables enter with two lags to allow richer dynamics" of how potential murderers evaluate the subjective probabilities of apprehension, conviction, and execution. Consequently, the MG instruments appear to be invalid since they would not be excludable from the second-stage regression.

4.2.3. Zimmerman.   Zimmerman concluded that none of his OLS estimates of the impact of executions were statistically significant, but his 2SLS estimates showed statistically significant evidence of deterrence.[25] But in addition to the flawed deterrence pseudo-probabilities, Zimmerman's 2SLS estimates rest on instruments that, though quite creative, are likely to be flawed. Specifically, Zimmerman's instruments included an indicator for whether an offender was released from death row in the previous year; an indicator of whether there was a botched execution in the previous year; and both contemporaneous and once-lagged values of the proportion of murders committed by strangers, by nonwhites, and under nonfelony-related circumstances.[26] Note the conflict between MG and Zimmerman on the appropriate treatment of the release of a death row convict: MG deemed that factor to be an important explanatory variable in the second-stage regression and thus not excludable, while Zimmerman used it as an instrument, which required it to be excludable. In addition, if, as one might well suspect in the current political environment, courts or governors are sensitive to

---

25. Zimmerman (2004, p. 190) estimates "that each state execution deters somewhere between four and twenty-five murders per year (fourteen being the average)," but states that "[t]he results appear to be highly sensitive to functional form."

26. With respect to the "stranger murder" instrument, Zimmerman (2004, p. 175–76) states that "this instrument is likely to be exogenous to the structural crime equation since the *rate* of per-capita murders in a given state will not influence the proportion of murders that are committed by strangers." This variable may well correlate with the rate of arrest as Zimmerman posits, so the first requirement of an instrument is met, but the requirement of excludability seems dubious: crime waves tend to correlate with stranger murders, so an increasing proportion of murders that are committed by strangers likely will be linked to higher murder rates. Thus, this variable does not influence murder rates only through its influence on arrest rates. We agree with the concern that Zimmerman articulates on p. 174: "Bound et al. (1995) show that 2SLS estimates can be highly biased in small samples if the instruments are weak and have even a small correlation with the structural error term."

the murder rate in making decisions that release convicts from death row, then the release variable would violate the exogeneity requirement for valid instruments.[27] Moreover, one might expect certain classes of homicides to vary more than others. If so, their share of total homicides will be directly correlated with the homicide rate, invalidating Zimmerman's use of these murder-proportion instruments.

Among the entire array of instruments in the three 2SLS studies we discuss, only one (and remember these studies all need at least three because they are instrumenting for three endogenous deterrence variables) seemed plausible—Zimmerman's indicator of whether there was a botched execution in the previous year. While this variable ostensibly comports with the definition of an instrument in that it might be correlated with executions but not influence the murder rate except through its effect on executions, it is still potentially problematic. We have emphasized that the rate of executions in the post-moratorium period that these studies examine has been rather low, which complicates the estimation process. Given the extreme infrequency of *botched* executions—Zimmerman counts only twenty-three such cases across his entire data period—there is considerable danger that this would be an extremely weak instrument. Moreover, the identifying assumption that botched executions would depress executions in the following year would seem to be undermined by the fact that the one jurisdiction that had two botched executions in a single year—Texas in 1992—actually saw an increase in executions from twelve in 1992 to seventeen in 1993. Finally, as if the small number of botched executions was not enough of a problem, Zimmerman erroneously codes the case of Tommie J. Smith, who was executed by lethal injection in July 1996 in *Indiana*, as a botched execution in *Illinois*.

5. The Direction of Any Endogeneity Bias

This brief discussion of the DRS, MG, and Zimmerman 2SLS regressions underscores the key point that the requirements for valid instrumentation

---

27. One dramatic illustration of the power of this issue involved the criminal Willie Horton, who was released from a Massachusetts prison on a weekend furlough program while serving a life sentence for murder, without the possibility of parole. He did not return, fled to Maryland and committed a rape and armed robbery in 1986, which became a major political liability for Massachusetts Governor Michael Dukakis during the 1988 Presidential race.

must be carefully addressed before one can hope to achieve useful results from a 2SLS approach. Bad instruments cannot resolve an endogeneity problem, so if valid instruments cannot be found, researchers must assess how acute the endogeneity issue is. If debilitating, then the researcher must conclude that valid point estimation of the impact of capital punishment is not possible. On the other hand, if inclusion of the full set of explanatory variables in a panel data framework renders executions conditionally exogenous, then OLS estimates should be acceptable.

5.1. The Sources of Endogeneity

The most obvious argument in favor of a 2SLS approach to estimation of the impact of capital punishment is that the risk of execution for murder endogenously rises during periods when states experience surges in murder rates. This might be thought to be analogous to Steven Levitt's argument that instrumentation is needed to estimate the impact of police on crime because high crime leads to the hiring of more police, thus biasing panel data estimates against finding that police dampen crime (Levitt, 2002).

This argument is a serious one and merits consideration. Figure 1 suggests one way in which it might operate: when the murder rate is high, support for the death penalty and belief in its efficacy as a deterrent increases. If judges, prosecutors, and juries think (and act) along these lines, then the murder rate would have an impact on the frequency with which the death penalty was sought, handed down, and sustained on appeal. This tendency may have been important during the 1930s through the early 1960s when falling crime dampened the enthusiasm for the death penalty (see Figure 1). Figure 2 corroborates that the rate of executions and the rate of murder both fell very sharply over this period. Under these circumstances, OLS estimates would tend to *mask* any deterrent effect of capital punishment and would be biased toward an antideterrent finding.

But note that the Table 2 studies all look primarily (or in five of six cases, exclusively) at the post-moratorium period. Figure 1 shows a similar pattern of declining enthusiasm for the death penalty when murder rates fall, but unlike in the pre-moratorium period, executions do not closely follow the murder rate after 1977 (see Figure 2). Indeed, the rise in executions begins almost 20 years *after* the rise in murder rates and continues unabated for almost the next 20 years. This is not surprising given that a defining trait of the use of capital punishment in the United States over the last 30 years is
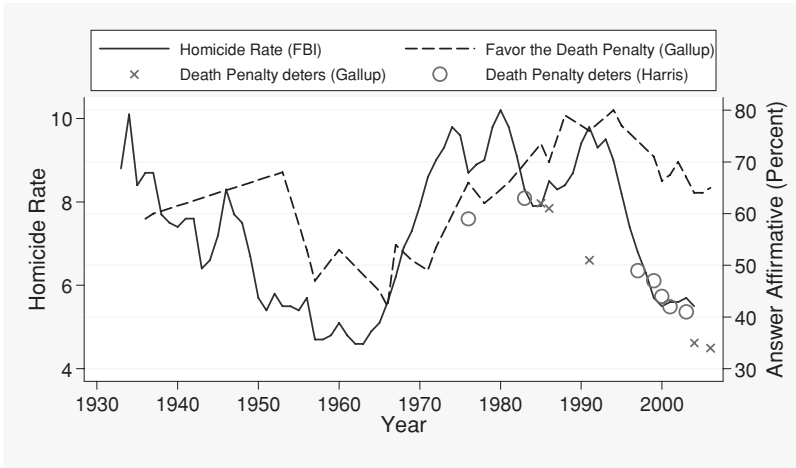
Figure 1. U.S. Homicide Rate and Public Opinion: 1930–2006.
Notes: Affirmative answers are based on the following public opinion polling
questions: "Are you in favor of the death penalty for a person convicted of
murder?" (Gallup). "Do you feel that the death penalty acts as a deterrent to the
commitment of murder, that it lowers the murder rate, or not?" (Gallup). "Do you
feel that executing people who commit murder deters others from committing
murder, or do you think such executions don't have much effect?" (Harris).
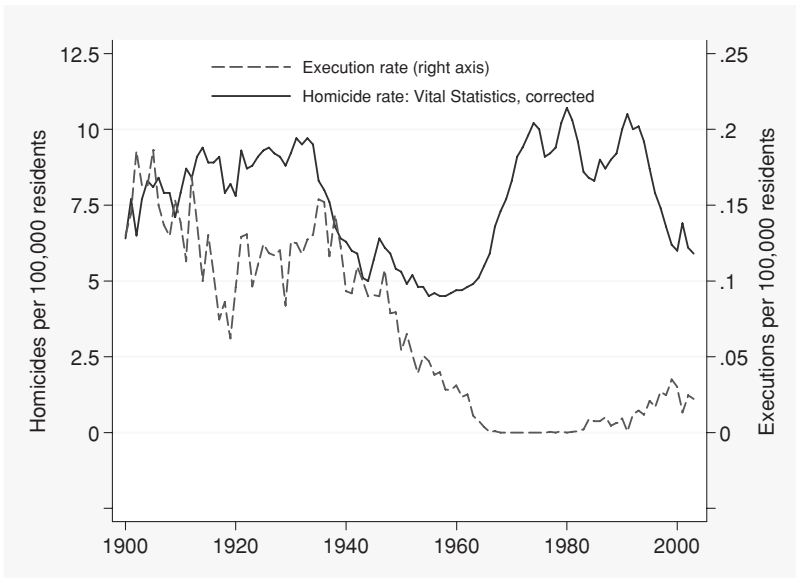

Figure 2. Homicide and Execution in the United States.

that there is almost invariably a lengthy lag between the commission of a capital murder and the execution of the defendant. If executions don't take place until long after the crime, then the simultaneous link between murder rates and execution risk is broken. This fact leads KVB to conclude that "current-year execution risk is an exogenous event having little or nothing to do with current-year homicide rates." KVB rely on this point to buttress their recent OLS evaluation.

If this type of casual time-series analysis could convince us that there was some bias against finding deterrence in the earlier period but no such bias in the post-moratorium, then one would take comfort in simply relying on the OLS estimates in the latter period. Certainly, this effect would be attenuated at the very least during the post-moratorium period if the main source of its operation would be jury decisions to more frequently hand down the death penalty in high crime periods, given the long lags between sentence and execution.[28]

But even if this primary source of endogeneity is not a concern during the post-moratorium period, other factors might be operating that would also suggest that a 2SLS approach could be desirable. First, high crime rates prompt get-tough responses, including not only more executions but also more aggressive policing, harsher prison conditions, longer criminal sentences in general (through truth in sentencing, three-strikes, gun penalty-enhancement laws, and less use of parole and prison leaves) and substantially more resort to life without parole in particular. Moreover, if the police or the public becomes more aggressive in shooting criminals when crime increases, one might observe a correlation between the prevalence of executions and those lawful shootings. Indeed, the number of criminals who are justifiably shot by the police or others is substantially higher than the number of citizens who are executed. Specifically, in 2002, a year in which there were

---

28. One could also imagine a contrary bias emanating from a tendency of stochastic jumps in the murder rate to induce greater resort to capital punishment. If these stochastic jumps tend to be followed by mean reversion in murder rates, OLS regression might improperly attribute to the executions the subsequent mean reversion in the murder rate. (Similarly, stochastic drops in murder rates could dampen executions, thereby making the subsequent reversion to the higher mean murder levels incorrectly appear to be causally induced by the drop in executions.) Unlike the previously mentioned endogeneity concern that would mask a deterrent effect, this mean-reversion effect would tend to *overstate* any deterrent impact.

seventy-one executions, roughly 200 individuals were justifiably killed by other citizens and about 350 were killed by police, of whom about 200 reportedly had attacked the police (Bonczar and Snell, 2003; Bureau of Justice Statistics, 2004).[29]

Ideally, the econometrician wants to know what potential murderers perceive to be the full array of negative consequences associated with an attempt to commit murder. Since death penalty studies typically estimate the murder rate, as opposed to the death-eligible murder rate, any of these justifiable killings of murderers can be thought of as immediate "executions," which would presumably have an even greater deterrent effect on murder by virtue of their speedier application and greater frequency.[30] Because all of these factors are omitted from the murder equations in all the studies we have discussed, their contribution to reductions in murders will be attributed to executions (for those factors that are positively correlated with executions), thus *overstating* any deterrent impact of executions. One way to establish the presence of such omitted variable bias is to run the regression that estimates the effect of executions on a crime, such as burglary, that does not carry the death penalty. Thomas Kovandzic shared with us his regression showing that executions do correlate with a lower rate of burglary, which is likely the result of some omitted variable capturing the harshness of the criminal justice system that correlates with executions.

If all the factors omitted from the regressions that correlate with executions also correlated with lower crime then the OLS estimates would only be biased in favor of a finding of deterrence. Again, if this were the full story, we could sign the endogeneity and at the least think of the OLS estimate as representing an upper bound to any deterrent effect. This would be powerful evidence against deterrence given the nature of the OLS estimates we have already seen.

But, of course, other types of omitted variables are possible. If one subscribed to the argument that abortion legalization dampened crime (Donohue and Levitt, 2001), then one would need to see whether executions

---

29. If not all of these killings were of potential or actual murderers, though, one would ideally want to adjust the total number downward.

30. On the other hand, some of these individuals may have been suicidal (consider the implications of attacking the police), in which case the speedy termination of life might actually have served as an inducement to crime.

were more common in the high-abortion states, thereby wrongly suggesting that executions deterred crime. Conversely, if executions were more common in the low-abortion states, then any deterrent effect would tend to be masked.

If future research were to find that the above forms of endogeneity bias are small or nonexistent in the post-moratorium period and that the primary omitted variable bias tended to involve factors that correlated with executions while themselves dampening crime, then OLS estimates on post-moratorium data would seem to be *unambiguously biased in favor of deterrence*. Identifying likely pro-deterrence bias in the post-moratorium data could therefore aid in our interpretation of the various post-moratorium OLS estimates, which tend to be negative (suggesting deterrence), but small and statistically insignificant.

5.2. Pro-deterrence Bias during the Post-moratorium Period

We have already discussed how high crime in New York stimulated a successful Republican effort to enact a death penalty statute in 1995, and comparable forces operating at roughly the same time turned Texas into the most active executing state in America. Although the homicide rate peaked in Texas in 1991 at 15.3 per 100,000 during Ann Richards' first term as Governor and fell every year of her term (down to 11 per 100,000 in her last full year of 1994), "the key campaign issues in the [1994] Texas gubernatorial election were mainly crime and gun control; Richards suffered when her stances on both issues became viewed as weak."[31]

Governor George Bush responded to the concern of voters by encouraging the greater use of capital punishment: while only fifty executions occurred during Ann Richards' four-year tenure as Governor, 152 murderers were executed during Bush's six years as Governor, with almost half (75 of 152) coming in the last two years. Thus, increased executions in Texas came in response to a political decision, not as part of a mechanical application of the death penalty to a fixed percentage of murders, which is the vehicle for simultaneity endogeneity bias. In contrast, during the period from 1933 until 1962, the coincident declines in both homicide rates (down 48%) and executions (down 80%) suggest that both the mechanical link

---

31. http://en.wikipedia.org/wiki/Ann_Richards (accessed on June 18, 2006).

between murders and executions, as well as the declining enthusiasm for capital punishment suggested in Figure 1, might be biasing the estimated impact of the death penalty *against* deterrence during these pre-moratorium years. Indeed, this is exactly what we found when we estimated KLS-type models over the period of 1934–1960: even though this was the period when executions were more frequent, harsher, and more swiftly imposed, our estimates, summarized in Table 5, reveal a positive and statistically significant relationship (suggesting antideterrence).

The pro-deterrence omitted variables and mean-reversion endogeneity concerns, however, are evident in the post-moratorium Texas data. As the three panels of Figure 3 show, in 1992 Texas and New York had very similar murder rates and incarceration rates (New York had no death penalty statute in 1992 but, again, no one was executed even after one was passed in 1995, while Texas had five executions in 1991 and 12 in 1992). Texas responded to the crime peak not only with a revival of executions, but also with a massive increase in incarceration rates (Butterfield, 2004).[32] Spelman conducted a county-by-county assessment of the crime drop in Texas and concluded:

> Texas' prison buildup was massive: 100,000 more prisoners, 5,000 more jail inmates, at an estimated direct cost of $1.5 billion per year for Texas taxpayers. The increase was much larger, on both a percentage and an absolute basis, than the prison expansion of any other state. It was Texas's principal response to the crime problem (2005).[33]

Researchers, such as DRS, who do not control for a variety of inhibiting influences on murder (for example, the magnitude and harshness of

---

32. KLS found that harsh prison conditions tended to have a deterrent effect on crime, which might also be relevant to the crime drop, given the conditions that prevailed in Texas prisons in the 1990s during Bush's tenure as governor. According to one report: "The corrections experts say that some of the worst abuses have occurred in Texas, whose prisons were under a federal consent decree during much of the time President Bush was governor because of crowding and violence by guards against inmates. Judge William Wayne Justice of Federal District Court imposed the decree after finding that guards were allowing inmate gang leaders to buy and sell other inmates as slaves for sex."

33. Spelman concluded that executions had no effect on violent crime in Texas, and that virtually all of the drop was attributable to increased incarceration. Spelman's conclusions about the ineffectiveness of capital punishment are not directly on point for our analysis since Spelman focuses on the effect of the death penalty on the vastly larger category of violent crime, rather than on murder alone.
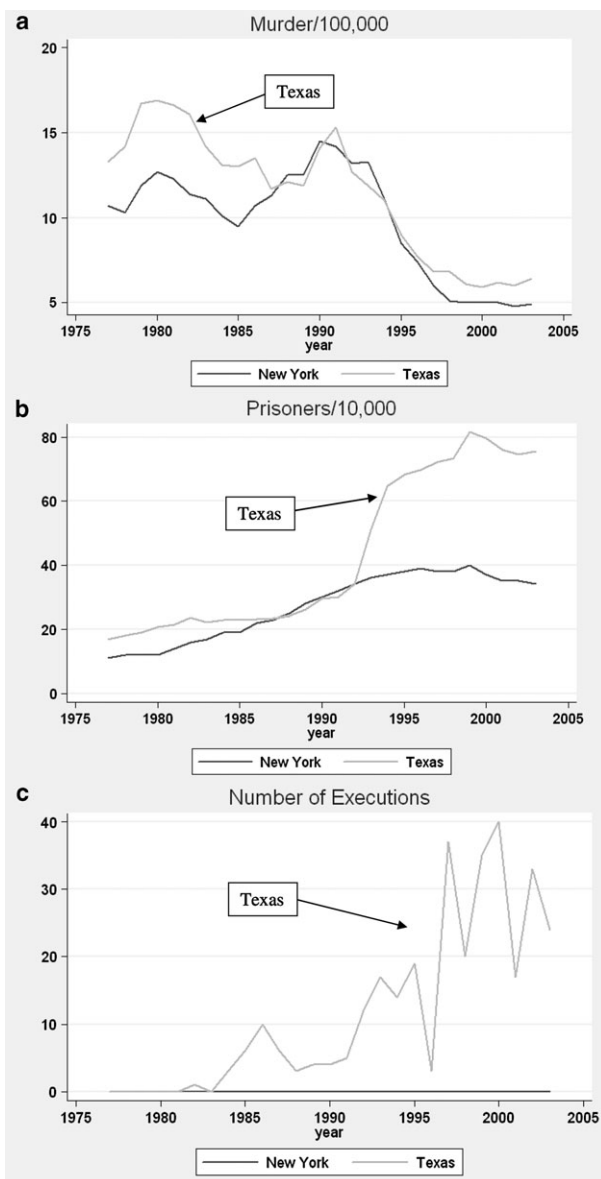
Figure 3. (a) Murder Rate in New York and Texas: 1977–2003. (b) Incarceration Rate in New York and Texas: 1977—2003. (c) Executions in New York and Texas: 1977–2003.
Note: Data on prisoners under state or federal jurisdiction are from the Bureau of Justice Statistics.

incarceration) may be attributing the large decline in the murder rate of the 1990s to the execution rate rather than these other factors. Similarly, any mean reversion in the murder rate would also be incorrectly attributed to "deterrence variables" in studies that do not properly instrument for executions. Even though from 1992 to 2003 New York had no executions and only a fraction of the increase in the rate of incarceration, the state of New York actually enjoyed a steeper decline in homicide rates of 62.9% versus the decline of 49.6% in Texas over the same period.[34]

In summary, if further research could confirm that the likely omitted variables and endogeneity bias in post-moratorium estimates of the impact of capital punishment on murder operated to exaggerate the deterrent effect, then valid instrumenting would be expected to *lower* the estimated deterrent effect of the death penalty.

Two important conclusions would follow from this observation. First, any 2SLS estimates showing *greater* deterrence than the corresponding OLS estimates in the post-moratorium period when the OLS estimates are biased in favor of deterrence are presumptively invalid. Second, the inability to find adequate instruments for executions during the post-moratorium period is not wholly debilitating because we at least know that OLS estimates of the impact of executions will be *biased in favor of the deterrence argument*. If we don't find a deterrent effect in the post-moratorium OLS regressions (or do find evidence of antideterrence), we can assume that valid instrumentation would only strengthen the finding of no deterrence (or antideterrence). In this event, the fact that every (corrected) OLS estimate of the post-moratorium effect of capital punishment referred to in Table 2 found no statistically significant effect would be taken as evidence *against* deterrence as the death penalty was practiced in the United States during the post-moratorium period.

---

34. Over this same period, both states had only modest changes in the number of police state-wide, although New York maintained a considerably higher police presence throughout. Specifically, in New York, full-time law enforcement per 1000 residents rose from 4.0 to 4.1 as sworn officers dropped from 3.1 to 3.0. In Texas, full-time law enforcement per 1000 residents rose from 3.4 to 3.5 while sworn officers stayed constant at 2.2. Source: FBI annual UCR publication.

"Progress in economics comes . . . not from maintaining priors in the face of weak evidence but from obtaining new evidence and adjusting priors to new knowledge" Richard Freeman (2005).

6. Conclusion

What conclusions about econometric evaluation of the deterrent effect of the death penalty can be drawn from our prior work, a review of the studies summarized in Table 2, and our current discussion? There is a clear pattern in the panel data studies: OLS estimates show no evidence of a deterrent impact and three 2SLS studies find that a deterrent effect exists.

Clearly, all of these studies face challenges in trying to generate plausible estimated effects of executions on murder. First, as Berk (2005) persuasively illustrated, the studies conducted to date on the post-1977 panel data have had less variation in the application of capital punishment than a researcher would ideally like in trying to provide reasonably accurate estimates of the effect of the death penalty. Second, many choices have to be made in the process of specification of these panel data models, at times on points about which our knowledge is limited. There is considerable uncertainty about the precise mechanism of informational acquisition about the relevant sanctions facing potential murderers, as well as about how and whether the death penalty coarsens the social environment and thereby negatively impacts those with potential criminal propensities.

Third, the long lag between murder and execution (and the general uncertainty in application of the death penalty) in the post-moratorium period makes it hard to know the correct model to capture the risk of execution. Linear models in the number of executions, even with the array of lag choices that have been employed, may not well capture the more impressionistic sense of risk that a potential murderer might develop over that period.

Fourth, data limitations are constraining in that it is not easy to classify "capital murders"—those that can lead to the death penalty (and these definitions are not uniform across death penalty states and have changed substantially over time within states, which further complicates the picture). Given the likely degree of knowledge by potential murderers about the precise details of capital punishment, it is not clear that we should construct our models on the objective reality of what murders are capital crimes. It may be preferable to try to explain "all murders" instead of "capital murders" (the

only ones that in fact can lead to a death sentence) if information among the potential criminal element is poor and/or the costs of trying to ascertain the subjective understanding of this element is too high (given the likely costs of error and benefits from greater precision).

While endogeneity is a concern in trying to evaluate the impact of the death penalty on murder, there are reasons to think that the bias during the post-moratorium period may primarily be in the direction of exaggerating the deterrent effects of executions. If future research can confirm this suspicion, then the state OLS panel data may provide a valuable upper-bound on the potential for the death penalty to deter murder in the United States over the last 30 years. Of course, since the OLS results show no evidence of deterrence, if they in fact represent an upper bound, then the case for a deterrent effect would be severely undermined.[35]

Signing the endogeneity in this way would also establish that the 2SLS estimates are incorrect since instrumentation would then be expected to reduce the size of the estimated effects relative to OLS estimates and in the three studies presented, the 2SLS estimates all greatly increased the apparent deterrence of the death penalty. It is not surprising to us that these 2SLS studies cannot yield reliable conclusions.

First, all of the 2SLS estimates lose statistical significance if one clusters the standard errors. While such an adjustment for serial correlation has now become standard practice in the applied econometric literature, the debate over the need for, and a proper approach to make, this adjustment continues (Donohue et al., 2009).[36]

Second, the series of linked "deterrence" ratios create major, largely self-inflicted difficulties for all three 2SLS studies. Future authors of studies in this domain would do well to look to the simpler and more compelling structure of the OLS studies rather than the convoluted approach of mismeasured pseudo-probabilities that Ehrlich initiated. In particular, the key ratio of

---

35. Might measurement error in the endogenous variables (executions) explain the large 2SLS estimates? This seems unlikely since the instrumentation is for a variable that is already perfectly measured. The problem is that the endogenous variable (executions) may not be fully capturing the knowledge and fears of potential criminals about a harsh sanction. None of the instruments is designed to address this measurement problem.

36. Zimmerman (this issue) argues that clustering may overinflate the standard errors, and that using other forms of adjustment for serial correlation, his 2SLS estimates remain robust.

executions to death sentences, whether dressed up with a fancy lag structure or not, is not well designed to capture a meaningful risk of execution.

Third, perhaps credible instrumentation could overcome any endogeneity biases that exist, but the requisite instruments have not been forthcoming to date. The instruments employed in the 2SLS studies are either invalid on their face or very weak, and thus it should not be a surprise that they cannot yield reliable estimates. In our opinion, this is the final fatal flaw in the various 2SLS studies.

Given the insurmountable difficulties of the poor structure of the key deterrence pseudo-probabilities and the invalid instruments that plague the 2SLS studies, we doubt that efforts to model average (Cohen-Cole et al., this issue) or conduct meta-analyses (Yang and Lester, 2008) based on any of these studies will generate helpful estimates, although these tools will certainly be useful if they can be applied to sounder regression frameworks.

As far as we know, no one has grappled with the problem of spillover effects across geographic lines, which would certainly complicate the process of estimation. Moreover, capturing any brutalizing effect of the death penalty requires a deft assessment of how people identify with governmental action. Presumably, executions in Saudi Arabia do not brutalize the population of Vermont, but one could imagine that executions in Texas might have that impact within the United States in ways that are not easily captured in the existing statistical models. Indeed, if any brutalization effect operated nationally from the presence of executions within the United States, its presence could never be picked up from a panel data analysis.

At the end of the day, the fact that all these analyses over such a long period of time using plausible data and models generate so little evidence of deterrence suggests that any effect is likely to be small, and that one should be highly dubious about "new" claims that strong evidence of deterrence exists. Clearly, a functioning system of criminal justice that exposes criminals to a sizeable risk of arrest, conviction, and punishment will deter all crimes including murder.[37] Beyond that, there is no statistically significant evidence of additional deterrence from either the period of more frequent and quicker application of capital punishment (prior to 1960) or in the post-moratorium

---

37. Our discussion only refers to situations in which a functioning criminal justice system is in place, since all of the analyses were conducted on American data after 1933. This analysis cannot say anything about reliance on the death penalty in societies where the criminal justice system operates at a very low level of efficacy.

phase when incremental tweaks through small numbers of executions are handed out in a highly unpredictable manner and implemented only after a decade or more of appeals. At this point, we are left to struggle with the difficult question of whether these results can be taken to show that the death penalty has had no effect on murder in the United States or that we simply can't detect the likely small effect (positive or negative) given the crudeness of our data, our models of potential criminal conduct, and our improving but still imperfect econometric tools.

In this regard, it is worth considering a fascinating recent study by Zimring and Fagan (forthcoming) that compares Hong Kong, which has no death penalty, with the very comparable state of Singapore, which in the mid-1990s experimented with a policy of swift and prodigious use of capital punishment. Since these two "cities" are roughly equal in size, share many population characteristics, and had similar murder rates before Singapore decided to ramp up its rate of executions, it would seem to be a perfect natural experiment of the impact of a major use of the death penalty. As the authors note:

> Singapore had an execution rate close to 1 per million per year until an explosive twentyfold increase in 1994–95 and 96 to a level that we show was probably the highest in the world. Then over the next 11 years, Singapore executions dropped by about 95%. Hong Kong, by contrast, has no executions all during the last generation and abolished capital punishment in 1993. Homicide levels and trends are remarkably similar in these two cities over the 35 years after 1973, with neither the surge in Singapore executions nor the more recent steep drop producing any differential impact.

While Singapore and Hong Kong are obviously different from the United States, it is striking that Singapore's massive increases in executions shortly after Hong Kong formalized a long practice of not using capital punishment generated no apparent change in the path of homicide. As advocates for transparent approaches to illustrate the impact of capital punishment, we commend this study highly.

Some are surprised at the thought that the death penalty would not deter murder. They might themselves instinctively feel distinct unease at the thought of doing anything that could lead to a sentence of death. Yet this represents the wrong calculus, since virtually all such individuals would feel tremendous unease at doing something that could put them in a cage for the rest of their lives.

The important arithmetic of the death penalty is that it can only have a possible useful effect on a very small number of individuals—those that would *not* be deterred by the prospect of life without possibility of parole but *would* be deterred by the presence of the death penalty. In other words, if we look at New York—a state with no capital punishment (as of 2004), a large population (19,300,000) and a relatively low murder rate (4.77 per 100,000 people)—we find that 921 murders occurred in 2006. Assuming that 921 roughly represents the number of murderers in New York in 2006, then this represents the maximum number of individuals whose behavior could have been changed in a socially acceptable manner by the presence of a death penalty law (at least under a rational actor model). But against these 921 murderers who might potentially have been deterred by capital punishment, there were about 19,299,000 individuals in New York who were *not* deterred by the threat of capital punishment (since it was nonexistent and yet they still did not kill). This number is roughly 20,000 times as great as the number of murderers in New York in 2006. If the death penalty has a brutalization effect, then we at least have to think about whether any of the 19,299,000 current nonmurderers might be subject to a malign influence of capital punishment that would work in opposition to any possible benign influence that could potentially influence only 921 individuals.

Appendix: Some Additional Econometric Issues

A. Specification and Data Issues

This section explores the array of critical specification and data choices that researchers should consider in trying to estimate the impact of capital punishment on murder. Section A.1 begins with a discussion of the choice of a dependent variable, and Section A.2 explores how the choice of the model concerning the relationship between the death penalty and crime, as well as the type of data—county, state, or national—that is to be used in estimation, depends on what potential murderers know and respond to at the time they contemplate committing crimes and on various types of informational and crime spillovers across geographic boundaries. Section A.3 discusses some of the unfortunate modeling choices, initiated by Isaac Ehrlich, that have plagued a variety of studies that tried to adhere to those choices, such as the work by DRS, MG, and Zimmerman.

A.1. The Dependent Variable.   The six Table 2 studies specified the dependent variable as the murder rate, measured as murders per 100,000 population. MG also ran OLS state panel data regressions using the log murder rate as the dependent variable in estimating the impact of executions. Although when corrected for coding errors, neither of MG's execution models proved to be statistically significant (see DW, Table 6), MG's log murder rate dependent variable produced even weaker results for the deterrence hypothesis than were found using the levels of the murder rate.[38]

While the panel data studies typically seek to explain some measure of the aggregate murder rate, one might argue that "capital murders"—those that potentially expose the perpetrator to the death penalty under the relevant law—would be a conceptually superior dependent variable. Fagan et al. (2006) explored exactly this question, and concluded that there was no evidence that capital murders were any more responsive to the presence of the death penalty than overall homicides.

The choice about the appropriate dependent variable depends on the extent of knowledge of potential murderers. For example, Hjalmarsson (this issue) reports on murders in Houston between 1999 and 2004, of which 20% are classified as capital murders under Texas's capital murder statute, which extends to killing of police and firefighters, intentional murder while committing certain felonies, multiple murders, killing of a child under age 6, and murder while in or escaping from prison. Do potential murderers realize that perhaps 80% of murders are not death eligible? The Texas categories are fairly intuitive, so with any degree of awareness about the death penalty, potential murderers may have at least some sense of the factors that can aggravate a murder to a capital offense. But this also raises the possibility that, at least in the post-moratorium world, any deterrent effect may largely be limited to a relatively small percentage of all murders. In this event,

---

38. One could also conceivably try to estimate numbers of murders rather than murder rates. While this approach sounds unappealing since large states have much higher numbers of murders than small states, that issue can be addressed with the combination of state fixed effects and a control for population, although one needs to think carefully then about how to specify the key execution variable. Note that keeping population in the denominator of the dependent variable will not limit the attendant measurement error to the left-hand side variable, since explanatory variables such as prisoners per capita and police per capita (and executions per capita in some cases) already have the population measure in the denominator on the right-hand side variable. The potential problem of measurement error is real.

greater precision in estimation might result from making capital murders the dependent variable.[39]

Unfortunately, the readily available data—particularly if one wants to go back in time prior to the moratorium period (1972–1976)—is limited to a count of murders and non-negligent homicides. One might expect criminals to appreciate that multiple murders and killings of police are more likely to bring a death sentence, and Ekelund et al. (2006) show that such data can be used in a state panel data analysis (although their finding that capital punishment *increases* the number of multiple-victim homicides is likely spurious for the reasons discussed in Section 3.5). One complication is that multiple murders and police killings may correlate with the rise and subsequent stabilization of illegal markets, such as crack. Such factors that affect murders in certain states for limited time periods have plagued panel data estimations in the post-moratorium period since they are not easily controlled for either explicitly or with state and year fixed effects.[40]

A.2. Modeling the Impact of Capital Punishment.   Panel data analyses of capital punishment rest on interrelated modeling choices about whether the death penalty law or the number of executions (or both) influence murder (or "capital" murder) within certain geographical boundaries. These modeling choices turn on substantive issues about the information available to potential murderers and their responses to such information, as well as pragmatic considerations about data quality and availability. The effect of the death penalty is typically modeled by using one or both of the following key explanatory variables: (i) a dummy variable indicating the presence of an operable death penalty statute (a law dummy model); or (ii) some measure of

---

39. The supporters of the deterrence hypothesis assume that some rational potential murderers who would not be deterred by the prospect of life in prison would refrain from committing murder if they faced the death penalty. The logic of this claim would suggest that such potential murderers might also respond to a new death penalty statute by simply killing their victims in a somewhat less egregious fashion, thereby avoiding a capital murder conviction (without equivalently reducing total murders).

40. For an effort to create a variable that could capture the statewide criminogenic influence of crack, see Fryer et al. (2005). Problems with using the statewide crack measure in a county panel data analysis are discussed in Ayres and Donohue (2009).

the frequency of executions (an executions model).[41] The executions models have been specified in a number of different ways: KVB try to capture the approaches of a number of studies in presenting estimates (with various lags) based on (i) the number of executions (following Dezhbakhsh and Shepherd, 2006); (ii) the ratio of executions to either population or murders (following DW); (iii) the ratio of executions to prisoners (following KLS); or (iv) the ratio of executions to death sentences (following DRS, MG, and Zimmerman).

The appropriate measure would be whatever best captures the *perceived* threat of execution for the pool of potential murderers (Robinson and Darley, 2004; Anderson, 2002).[42] Of course, if these perceptions are highly accurate, then potential murderers would know the risk of execution by county and for particular types of murders. Geographically precise information in the hands of potential murderers would be a factor favoring the use of county data (as DRS did in an incomplete fashion),[43] but countervailing factors have led the majority of scholars (MG, Zimmerman, KLS, KVB, Ekelund et al.) to use state data.

A.2.1. Contaminating influences.   The studies using state data posit that the death penalty (and other measured explanatory variables) in state i will influence murders in state i but not in any other state j (which may or may not have a death penalty regime in place at that time). This assumption may be a prudent modeling choice, but it is also an approximation because the geographical units employed by the researcher are not hermetically sealed. Instead, there will be contaminating influences across jurisdictional boundaries from informational and crime spillover effects. The effect of

---

41. MG present some models that include both of these factors as explanatory variables. There has been some debate over whether it is appropriate to include both variables, as discussed in Cameron (1994).

42. One study of 159 inmates incarcerated in Kentucky and North Carolina in the late 1990s found that 76% of the sample and 89% "of the most violent criminals either perceive no risk of apprehension or are incognizant of the likely punishments for their crimes." The evidence is not ideal for our purposes, because it is not focused on *potential murderers* but rather is from a sample of *incarcerated felons* (who conceivably may be less likely to be knowledgeable and deterrable).

43. DRS use *county* crime data, but use *state* estimates of the likelihood of being sentenced to death or executed. Again, if criminals are aware of the different risks of receiving death sentences across counties, then county execution data would be preferable.

these influences must be carefully considered in the process of modeling as well as in interpreting the resulting estimates.

a. Informational spillovers. How do potential murderers acquire information about the death penalty? Ideally, we would like to know whether potential criminals at the time they deliberate on criminal choices are influenced by the presence of a death penalty statute; the frequency of executions in their state or substate unit (whether county or prosecutorial district); or the frequency of references to executions that they hear—whether real or fictitious (perhaps crime shows can be important here, which may or may not be correlated with legal status). Obviously, their knowledge will not be perfect: a potential murderer in a death penalty state may not even know he faces a risk of execution and his counterpart in a nondeath penalty state may incorrectly assume he faces a risk of execution. If potential murderers have poor information about the geographic contours of death penalty law or execution rates, panel data estimates of any deterrent effect of the death penalty will be biased toward zero (Hjalmarsson, this issue).[44]

Moreover, any antideterrent effect—recall the theory that executions *incite* violence by brutalizing the population—would tend to be masked if the presence of executions generates brutalizing effects outside the executing jurisdiction. Note, then, that informational spillovers can bias the estimated effect of the death penalty toward zero by masking both deterrent and antideterrent effects.

b. Crime spillovers. Locking up a criminal in neighboring New Jersey may dampen crime in New York City if the criminal would otherwise be spending considerable time in, say, Manhattan. This problem is not *directly* related to the presence of a death penalty statute, because if a capital murderer were caught and convicted, he ordinarily would be incapacitated from committing murder regardless of whether the crime occurred in a

---

44. Hjalmarsson's study of death row inmates in Texas found that half of those on death row had never been in prison before their murder conviction. Thus, one mechanism for learning details about capital punishment in a state or county would not be open to a substantial number of potential capital murderers.

In the extreme case where potential murderers have literally no idea about the local death penalty situation but derive their views of the risk of execution from the national media, Ehrlich's time-series approach would actually be a sensible one (although one would still have to contend with the inherent limitations of a national time-series analysis to properly control for other factors influencing homicides, given the limited degrees of freedom).

death penalty state. But the choice of geographic unit will be important if arrests or incarcerations in one county or state are influencing crime in other counties or states since such spillovers will compromise the estimates of other control variables in the panel data models that may be correlated with capital punishment. Relatedly, if criminals respond to the risk of execution by traveling to commit their murders in counties or states where there is no such risk, then the death penalty would be altering criminal behavior but the panel data models would overstate the magnitude of the drop in homicides caused by the death penalty.

A.2.2. The relative advantages of state data.   In general, various spillover effects will be less problematic the larger the geographic unit for which the death penalty effect is estimated (the nation for Ehrlich's time-series analysis or the state or county in the various panel data studies). Conversely, if potential murderers have precise information and all of the effects of the death penalty are localized (that is, there are no spillovers across county lines), then county data would have some strong advantages. As DRS note, the use of county data enables them to more precisely control for any enduring but unmeasured crime features with roughly 3000 county fixed effects, as opposed to fifty state fixed effects.

In the end, we think the advantages of using county data are outweighed by the disadvantages (Ayres and Donohue, 2003; Zimmerman, 2004), and five of the six papers summarized in Table 2 relied on state data (DRS is the sole exception). First, there is a greater concern over the quality of the county data, which is especially problematic for those papers adhering to Ehrlich's use of the murder arrest rate as an explanatory variable (DRS, MG, Zimmerman, Ekelund et al.).[45] While the following section will discuss the conceptual and practical problems with the arrest rate variable, the state data are at least more accurate than the county data.

Second, by using county data, DRS have a very large proportion of zero values for murders by county per year, which creates problems for the standard regression techniques that these followers of Ehrlich employ.[46] Third, the most important DRS variables—such as number of executions,

---

45. Michael D. Maltz, "Bridging Gaps in Police Crime Data," September 1999, NCJ 176365.
46. The DRS county dataset has zero values for murder by county by year in 45% of the nonmissing values.

number of death sentences, and DRS's various instrumental variables—
are measured at the *state* level, so the DRS study is not a pure county
data analysis. Conceivably, their blending of state deterrence data with less
accurate county crime and arrest data may give us the worst of all worlds
by combining the poorer quality county data with the state data's inability
to link execution risks and crime changes at the county level.

A.3. The Problematic Specification Used by Ehrlich and His Followers

A.3.1. The "Number of Executions" model.   Unlike Dezhbakhsh and
Shepherd (2006), KLS used only a measure of executions (rather than a law
dummy) to capture the influence of the death penalty. By so doing, they are
testing the behavioral model that the impact of the death penalty is zero until
executions occur and that executions only impact the murder rate for the year
following the execution. Although the KLS results were estimated on 1950–
1990 data and thus did not cover the period of New York's death penalty
statute, we reestimated the KLS models on data through 2000, which did
cover this period. Given the complete lack of executions, the KLS approach
treats New York as a control state rather than a treatment state. The results
from our expanded KLS model (not shown, but available from the authors)
provide *no* statistical support for the deterrence hypothesis.

A.3.2. Ehrlich ratio models.   The studies following the Ehrlich
approach—DRS, MG, and Zimmerman—basically posit that the mere fact
of a legal death penalty statute has no impact on potential murderers, but
such potential criminals will be influenced by the ratio of death sentences
to arrests, and the ratio of executions to death sentences (as well as the ratio
of arrests to murderers). Accordingly, in these studies, some of the large
drop in New York's murder rate *will* be attributed to the *death sentences*
that were handed down throughout the state over the relevant decade even
in the models based on executions. But note once again the implications of
various decisions about specification and data. We know that DRS, MG, and
Zimmerman implicitly assume that potential murderers know when death
sentences are handed down and executions occur anywhere within their
state.[47]

---

47. Interestingly, because DRS analyze UCR data (which they obtained from John
Lott), they could not have implemented a fully accurate county data analysis for New York
State even if they had wanted to. The reason is that the UCR simply used a single murder

If potential killers perceive different risks across county—for example, they know that prosecutors in Manhattan and the Bronx opposed capital punishment, while the Brooklyn district attorney did not—then this might be an argument for using county data. On the other hand, since DRS use *state* rather than county death sentences and execution counts, they have not exploited this one potential advantage of county data (and DRS had the wrong county crime data for New York City in any event).

A.3.3. The high expense of the death penalty.   The considerable expense of having a death penalty regime is underscored by the fact that during the decade of 1995–2004 when New York had a valid death penalty law in effect, the state spent roughly $170 million to administer its death penalty system, yet not a single person was executed (Death Penalty Information Center, 2004).[48] Also, see both Cook (in this issue) and Roman et al. (in this issue) for estimated costs associated with the death penalty in North Carolina and Maryland, respectively. The high expense of the death penalty raises both economic and econometric issues.

First, while some authors have suggested that if the death penalty deters then it is worth having—see Becker (2006), Posner (2005), and Sunstein and Vermeule (2005)—this reasoning is incomplete in that it would assess a policy based on estimated benefits without considering costs. In the context of the New York death penalty law, an economist would naturally be interested in whether that $17 million per year was being drained away from other murder-reducing expenditures (or other life-saving measures?), whether public or private. In sum, we need to know the opportunity cost of the considerable resources that are consumed in running a death penalty system in the United States.

Second, the high cost of the death penalty raises an interesting econometric issue. In estimating the effect of the death penalty on crime, will the regression estimates reflect the stimulus to crime that the death penalty

count for the entire city of New York and then allocated the decline proportionally (by population) across counties. Needless to say, this compromises DRS' county data analysis. (The murder rate data we have presented in Section 2.2 for New York counties were obtained from the New York State Division of Criminal Justice Services.)

48. Death Penalty Information Center. "Experts Question 'Quick Fix' to New York's Death Penalty" citing a letter to the New York state legislature cosigned by U.S. Attorney Zachary Carter and former First Deputy Police Commissioner of New York City John Pritchard.

regime induces by crowding out other crime-fighting measures? For example, if the expenses imposed by New York's death penalty law ended up dampening police hiring, the attendant crime stimulus will *not* be attributed to the death penalty in econometric models that control for the number of police.[49] Conversely, if the $17 million per year crowds out some crime-reducing expenditures that are *not* controlled for in the relevant regression, then the estimated coefficient on the death penalty law dummy would reflect this assumed murder stimulant. The result is that one would be estimating not the effect of the death penalty with everything else held constant, but the net effect of the death penalty plus the assumed crime stimulant caused by the crowding out of anticrime measures for which the regression does not control. This would likely make the death penalty look *worse* than its own independent influence.

The treatment of New York is, of course, different if the death penalty is modeled by executions rather than a law dummy, since New York had no executions. With this modeling choice and again assuming that we don't have adequate controls for the crowded-out anticrime measures, the postulated uptick in murders ascribed to the wasted $17 million in annual crime-fighting expenditures would actually count as a murder *reduction* for states having executions because, in the difference-in-difference analysis, New York (with its presumed increase in murders) would be a *control* state to be contrasted with the executing *treatment* states. In this event, the death penalty would actually appear to be *better* than its independent influence.

A.3.4. Possible omitted factors in the various regression studies.   Ideally, the murder rate regressions would control for a full array of criminal justice variables, reflecting the degree of incarceration (to capture any incapacitative impact on crime), the nature and extent of policing, as well as typical time served for murderers who are not executed, and the lengths of delays for those who are. These controls are particularly crucial, as it seems likely that jurisdictions that are "tough on crime" not only implement capital punishment vigorously, but also increase social sanctions against criminal behavior, increase policing, enhance sentence lengths, increase time served,

---

49. See Levitt (2002) and Donohue and Siegelman (1998) for discussions about the value of resources spent on hiring police, increasing incarceration, or other crime-fighting measures.

and are less concerned with prisoners' rights. Thus, in the absence of such controls, one might wrongly attribute the effects of these other (omitted) factors to the imposition of capital punishment. A particularly pressing concern in current studies is the absence of data on sentence lengths and time served for murder, reflecting difficulties in obtaining such data. Only two studies listed in Table 2—KLS (2003) and KVB (2009)—include a control for prison conditions, albeit a rather crude one (the prison death rate).

Beyond these punishment variables, it seems likely that education and the inculcation of moral or religious precepts against violence would inhibit murders while the expansion of illegal markets and increases in racial or ethnic tensions, social instability, the percentage of young men (particularly those not engaged in productive activities),[50] and a macho or violent gun culture would tend to exacerbate the murder rate. Typical crime models include crude controls for demographics (typically age and race), but it seems likely that these broader social forces affect both execution policy and homicide rates. If these time-varying forces do not operate uniformly across states, then the year fixed effects will not be able to control for them.

References

Anderson, David. 2002. "The Deterrence Hypothesis and Picking Pockets at the Pickpocket's Hanging," 4 *American Law and Economics Review* 295.

Ayres, Ian, and John Donohue. 2003. "Shooting Down the 'More Guns, Less Crime' Hypothesis," 55 *Stanford Law Review* 1193.

———. 2009. "Yet Another Refutation of the More Guns, Less Crime Hypothesis— With Some Help From Moody and Marvell," 6 *Econ Journal Watch* 35–59.

Becker, Gary. 2006. "On the Economics of Capital Punishment," 3 *The Economists' Voice* art. 4. Available at http://www.bepress.com/ev/vol3/iss3/art4.

Berk, R. 2005. "New Claims about Executions and General Deterrence: Deja Vu All Over Again?" 2 *Journal of Empirical Legal Studies* 303–30.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-in-Differences Estimates?" 119 *Quarterly Journal of Economics* 249–75.

---

50. For example, the percentage of white males 18–55 that is not employed, not in school, and not in the military, is about 10%, and it is roughly twice that percentage for blacks (rising from 20.6% in 1980 to 23.9% in 2000). See Raphael and Stoll (2009, p. 5).

Bonczar, Thomas P., and Tracy L. Snell. U.S. Department of Justice. Bureau of Justice Statistics. 2003. "Capital Punishment, 2002," 8 *Bureau of Justice Statistics Bulletin* 1–17.

Bowers, William, and Glenn Pierce. 1975. "The Illusion of Deterrence in Isaac Ehrlich's Research on Capital Punishment," 85 *Yale Law Journal* 187–92.

Butterfield, Fox. 2004. "Mistreatment of Prisoners Is Called Routine in the U.S.," *New York Times* May 8.

Cameron, Samuel. 1994. "A Review of the Econometric Evidence on the Effects of Capital Punishment," 23 *Journal of Socio-Economics* 197–214.

Cohen-Cole, Ethan, Steven Durlauf, Jeffrey Fagan, and Daniel Nagin. 2009. "Model Uncertainty and the Deterrent Effects of Capital Punishment," *American Law and Economics Review* 10.1093/aler/ahn001.

Cook, Philip J. 2009. "Potential Savings from Abolition of the Death Penalty in North Carolina," *American Law and Economics Review* 10.1093/aler/ahp022.

Death Penalty Information Center. 2004. "Experts Question 'Quick Fix' to New York's Death Penalty." Available at http://www.deathpenaltyinfo.org/article.php?did=1111&scid=.

Dezhbakhsh, Hashem, Paul Rubin, and Joanna M. Shepherd 2003. "Does Capital Punishment Have a Deterrent Effect? New Evidence from Postmoratorium Panel Data," 5 *American Law and Economics Review* 344.

Dezhbakhsh, Hashem, and Joanna M. Shepherd. 2006. "The Deterrent Effect of Capital Punishment: Evidence from a Judicial Experiment," 44 *Economic Inquiry* 512–35.

Dieter, Richard, and Death Penalty Information Center. 1995. "On the Front Line: Law Enforcement Views on the Death Penalty," Poll. Washington, DC: Death Penalty Information Center.

Donohue, John, and Steven D. Levitt. 2001. "The Impact of Legalized Abortion on Crime," 116 *Quarterly Journal of Economics* 379–420.

Donohue, John, and Peter Siegelman. 1998. "Allocating Resources Among Prisons and Social Programs in the Battle against Crime," 27 *Journal of Legal Studies* 1.

Donohue, John, and Justin Wolfers. 2005. "Uses and Abuses of Empirical Evidence in the Death Penalty Debate," 58 *Stanford Law Review* 791.

———. 2005. "The Death Penalty: No Evidence of Deterrence," 3 *The Economists Voice* art. 3. Available at http://www.bepress.com/ev/vol3/iss5/art3.

Donohue, John, Abhay Aneja, and Alexandria Zhang. 2009. "The Impact of Right to Carry Laws and the NRC Report: Lessons for the Empirical Evaluation of Law and Policy." Yale Law School Working Paper, 1–34.

Ehrlich, Isaac. 1975. "The Deterrent Effect of Capital Punishment: A Question of Life and Death," 65 *American Economic Review* 397–417.

———. 1977. "Capital Punishment and Deterrence: Some Further Thoughts and Additional Evidence," 85 *Journal of Political Economy* August 741–88.

Ehrlich, Isaac, and Zhiqiang Liu. 1999. "Sensitivity Analysis of the Deterrence Hypothesis: Let's Keep the Econ in Econometrics," 42 *Journal of Law and Economics* 455–88.

Ekelund, Robert, John Jackson, Rand Ressler, and Robert Tollison. 2006. "Marginal Deterrence and Multiple Murders," 72 *Southern Economic Journal* 521–41.

Fagan, Jeffrey. 2006. "Death and Deterrence Redux: Science, Law and Causal Reasoning on Capital Punishment," 4 *Ohio State Journal of Criminal Law* 255–99.

Fagan, Jeffrey, Franklin E. Zimring, and Amanda Geller. 2006. "Capital Punishment and Capital Murder: Market Share and the Deterrent Effects of the Death Penalty," 84 *Texas Law Review* 1803–67.

Freeman, Richard. 2005. "Labour Market Institutions Without Blinders: The Debate Over Flexibility and Labour Market Performance," 19 *International Economic Journal* 129.

Frum, David. 2006. "Why Australia Needs the Death Penalty," *American Enterprise Institute for Policy Research*. Available at http://www.aei.org/publications/pubID.23966/pub_detail.asp (accessed June 9, 2006).

Fryer, Roland, Paul Heaton, Steven Levitt, and Kevin Murphy. 2005. "Measuring the Impact of Crack Cocaine," National Bureau of Economic Research Working Paper 11318, 1–65.

Hearing before the Subcommittee on Crime, Terrorism, and Homeland Security to consider H.R. 2934, the Terrorist Penalties Enhancement Act of 2003, to strengthen penalties for terrorist offenses. 108th Congress, Second Session, Apr. 21, 2004, p. 24.

Hjalmarsson, Randi. 2009. "Does Capital Punishment Have a 'Local' Deterrent Effect on Homicides?" *American Law and Economics Review* 10.1093/aler/ahn004.

Katz, L., S. Levitt, and E. Shustorovich. 2003. "Prison Conditions, Capital Punishment, and Deterrence," 5 *American Law and Economics Review* 318.

Kovandzic, Tomislav, L. Vieraitis, and D. Boots. 2009. "Does the Death Penalty Save Lives? New Evidence from State Panel Data, 1977 to 2006," 8 *Criminology and Public Policy* 803–43.

Kuziemko, Ilyana. 2006. "Does the Threat of the Death Penalty Affect Plea Bargaining in Murder Cases? Evidence from New York's 1995 Reinstatement of Capital Punishment," 8 *American Law and Economics Review* 116–42.

Levitt, Steven D. 1998. "Why Do Increased Arrest Rates Appear to Reduce Crime: Deterrence, Incapacitation, or Measurement Error?" 36(July) *Economic Inquiry* 353–72.

Levitt, Steven. 2002. "Using Electoral Cycles in Police Hiring to Estimate the Effects of Police on Crime: Reply," 92 *American Economic Review* 1244–50.

Liu, Zhiqiang. 2004. "Capital Punishment and the Deterrence Hypothesis: Some New Insights and Empirical Evidence," 30 *Eastern Economic Journal* 237–58.

Mocan, H. Naci, and R. Kaj Gittings. 2003. "Getting Off Death Row: Commuted Sentences and the Deterrent Effect of Capital Punishment," 46 *Journal of Law and Economics* 470.

Moody, C. E., and T. B. Marvell. 2009. "The Debate on Shall-Issue Laws," 5 *Econ Journal Watch* 269–93.

Morgenthau, Robert. 1995. "What Prosecutors Won't Tell You," *The New York Times* (Feb. 7 (1995). Available at http://query.nytimes.com/gst/fullpage.html?res=990CEFDA1130F934A35751C0A963958260&sec=&spon=&pagewanted=1.

Narayan, Paresh, and Russell Smyth. 2006. "Dead Man Walking: An Empirical Reassessment of the Deterrent Effect of Capital Punishment using the Bounds Testing Approach to Cointegration," 38 *Applied Economics* 1975–89.

National Research Council. 2005. *Firearms and Violence: A Critical Review*. Washington, DC: The National Academies Press.

Nerlove, Marc. 2002. *Essays in Panel Data Econometrics*. Cambridge, UK: Cambridge University Press.

Radelet, Michael L. 2008. "Some Examples of Post-Furman Botched Executions." Available at <http://deathpenaltyinfo.org/some-examples-post-furman-botched-executions> (accessed February 25, 2009).

Radelet, M., and R. Akers. 1996. "Deterrence and the Death Penalty: The Views of the Experts," 87 *Journal of Criminal Law and Criminology* 1–16.

Radelet, M., and T. Lacock. 2009. "Do Executions Lower Homicide Rates?: The Views of Leading Criminologists," 99 *Journal of Criminal Law and Criminology* 489–508.

Raphael, Steven, and Michael Stoll, eds. 2009. *Do Prisons Make Us Safer? The Benefits and Costs of the Prison Boom*. New York: Russell Sage Foundation.

Robinson, Paul H., and John M. Darley. 2004. "Does Criminal Law Deter? A Behavioral Science Investigation," 24 *Oxford Journal of Legal Studies* 173–205. Available at http://ssrn.com/abstract=660742.

Roman, John K., Aaron Chalfin, Aaron Sundquist, and Carly R. Knight. 2009. "Reassessing the Cost of the Death Penalty Using Quasi-Experimental Methods: Evidence from Maryland," *American Law and Economics Review* 10.1093/aler/ahp023.

Rubin, Paul. 2006a. "Reply to Donohue and Wolfers on the Death Penalty and Deterrence," 3 *The Economists' Voice* article 4. Available at http://www.bepress.com/ev/vol3/iss5/art4.

———. 2006b. "Statistical Evidence on Capital Punishment and the Deterrence of Homicide." Testifying on Feb. 1, 2006 in hearings on "An Examination of the Death Penalty in the United States," Senate Judiciary Committee. Available at http://mcadams.posc.mu.edu/blog/TestimonyPaulRubin.pdf.

Shepherd, Joanna. 2005. "Deterrence versus Brutalization: Capital Punishment's Differing Impacts among States," 104 *Michigan Law Review* 203–56.

Spelman, William. 2005. "Jobs or Jails: The Crime Drop in Texas," 24 *Journal of Policy Analysis and Management* 133–65.

Sunstein, C., and A. Vermeule. 2005. "Is Capital Punishment Morally Required?" 58 *Stanford Law Review* 706.

U.S. Department of Justice. Bureau of Justice Statistics. 2004. Homicide trends in the U.S., "Justifiable Homicides." Available at http://www.ojp.usdoj.gov/bjs/homicide/justify.htm.

Yang, Bijou, and David Lester. 2008. "The Deterrent Effect of Executions: A Meta-analysis Thirty Years after Ehrlich," 36 *Journal of Criminal Justice* 453–60.

Zimmerman, Paul R. 2004. "State Executions, Deterrence, and the Incidence of Murder," 7 *Journal of Applied Economics* 163–93.

———. 2009. "Statistical Variability and the Deterrent Effect of the Death Penalty," *American Law and Economics Review* 10.1093/aler/ahp003.

Zimring, Franklin E. 2008. "*Criminolog*y and Its Discontents: The American Society of Criminology 2007 Sutherland Address," 46 (May) *Criminology* 255–66.

Zimring, Franklin, Jeffrey Fagan, and David Johnson. Forthcoming. "Executions, Deterrence and Homicide: A Tale of Two Cities."