FIRST-DEGREE MURDER AND THE DEATH SENTENCE IN GEORGIA

by

Arnold Barnett

FIRST-DEGREE MURDER AND
THE DEATH SENTENCE IN GEORGIA

ABSTRACT

We propose a classification procedure for first-degree
murder cases, aimed at illuminating how Georgia judges,
juries, and prosecutors decide which convicted killers should
be sentenced to death. The scheme is empirical in origin,
arising from the study of over 500 actual cases. We consider
what the classification rules imply on such subjects as pro-
portionality review, allegations of racial bias in sentencing,
and the relationship between sentencing behavior and statutory
guidelines.

ARNOLD BARNETT
Sloan School of Management
Operations Research Center
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

ACKNOWLEDGEMENTS

INTRODUCTION

If "proportionality review" was at one time among the less visible issues in the capital punishment debate, its obscurity ended abruptly on October 4, 1983. On that day, J.D. Autry, already strapped down and injected with sedatives, was about to be executed with lethal drugs by the State of Texas. At literally the last moment, the U.S. Supreme Court halted the execution, explaining that it would have to decide whether Texas was constitutionally required to conduct tests of "proportionality" of its capital sentences.

The Supreme Court has since held that proportionality reviews, while certainly permissible, are not always mandatory.[1] But the decision did not render the issue moot. Over 25 states provide for some kind of proportionality review in their death-penalty procedures.[2] And the kinds of comparative judgments embodied in such reviews are important in varied investigations about the workings of the death penalty (e.g. whether race affects sentencing).

The notion animating proportionality review--one that has been explicitly endorsed by the Supreme Court[3]--is that death sentences cannot be distributed in an arbitrary manner. It is considered objectionable if a given defendant is put to death while, in adjacent counties (or adjacent courtrooms), several defendants in virtually the same situation are all given prison terms.

---

1. Pulley v. Harris, 52 U.S.L.W. 4141 (January 23, 1984). The Court stated, however, that proportionality review might be required if alternative checks on arbitrariness in a state's death sentencing were inadequate.

2. These states include Alabama, Connecticut, Delaware, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, Missouri, Montana, Nebraska, Nevada, New Hampshire, New Mexico, North Carolina, Oklahoma, Pennsylvania, South Carolina, Tennessee, Virginia, Washington, and Wyoming. The relevant statutes are listed in Baldus, Pulaski, and Woodworth, Comparative Review of Death Sentences: An Empirical Study of the Georgia Experience, 74 Journal of Criminal Law and Criminology p. 661 (1983).

3. See e.g. Zant v. Stephens, 103 S. CT. 2733, 2744 (1983), Garden v. Florida 430 U.S. 349, 361 (1977) (plurality opinion), Jurek v. Texas, 428 U.S. 262, 276 (1976); Lockett v. Ohio, 438 U.S. 586, 601 (1978).

Pursuant to that view, most states have taken steps meant to ensure uniformity in the imposition of the death penalty. Legislatures have prepared lists of aggravating and mitigating factors that judges, juries, and prosecutors must review in their decisions on homicide sentences. And, as a further precaution, murder trials are often broken into two phases: the first to determine guilt or innocence, and the second to set punishment for those convicted. Proportionality review is a retrospective test of whether such procedures are in fact avoiding capriciousness. When a death sentence is handed down, one looks at the outcomes in a series of similar cases; unless death was the penalty in an appreciable fraction of these, the present sentence is deemed excessive (or "disproportionate").

Defining the word "similar" in this context, however, is a most difficult task. Even two cases that coincide on the primary factual dimension (e.g. the robbery-killing of a grocer) might differ substantially on others (e.g. the prior criminal record of the defendant). Whether a given death sentence seems disproportionate can depend crucially on which cases are held comparable to the one at hand.

One could partition homicide cases into "similar" clusters on the basis of some theory of jurisprudence. But the clusters would inevitably reflect value judgments about the function of the death penalty and the relative culpability of various defendants. Thus, two thoughtful individuals could well wind up devising vastly different systems of categories.

In such a situation, it is useful to examine how capital sentencing guidelines are actually being interpreted. Observing the distinctions that juries and others are making might suggest an operational definition of "similarity" that, if nothing else, would at least have the virtue of reflecting "contemporary community standards." And, to the extent that one cannot rationally distinguish those cases that evoked death sentences from those that did not, one substantiates the fear that capital punishment cannot be applied consistently.

A major ancillary benefit of such an empirical exercise might be greater understanding of the racial patterns in death sentencing. In overall statistics, the fraction of white-victim slayings that end in death sentences is considerably higher than the comparable fraction for blacks.[4] But does this discrepancy reflect racial prejudice or instead legitimate distinctions that are coincidentally correlated with race? This important question might be easiest to consider within a broader analysis of sentencing behavior.[5]

This paper tries to identify the primary stimuli to death sentences in present-day Georgia. Data and narrative summaries about hundreds of first-degree murder cases -- all of them tried under Georgia's current death-penalty statute -- were prepared under the supervision of Professor David Baldus, who made them available to the Proportionality Review Project of the National Center for State Courts. We will consider the circumstances and verdicts of over 500 such cases, trying to infer a set of classification rules that, roughtly speaking, divide the cases into "homogeneous" subjects within which all killings are viewed as equally about "death-worthy".

Our methods of analysis and various findings will be described in detail in the remainder of the paper. To put it briefly, we did devise a classification scheme that fairly well describes Georgia sentencing behavior in first-degree murder cases. But to summarize the scheme in a few words risks oversimplifying that behavior; thus we will not try to do so here. Nor will we risk diminishing the problems of this endeavor by racing through them in an Introduction.

---

4. See Gross and Mauro, "Patterns of Death: An Analysis of Racial Disparities in Capital Sentencing and Homicide Victimization", (1983), Stanford University School of Law (unpublished manuscript), Table 1 (p. 43) and Table 30 (p. 93).

5. "The question of potential racial bias in sentencing was acknowledged by the United States Supreme Court last December when it granted a stay of execution to a prisoner in Georgia so the issue of discrimination in Georgia's sentencing could be studied."- The New York Times, July 9, 1984, p. A8.

We start our work in the next section with a brief description of the Georgia homicide statute. Then we discuss the rationale for and details of our "scoring rules" for murder cases (Sections III and IV). In Section V, we show the relationship between case scores and the chances of death verdicts. Next we focus several sections on "Category 3", the subset of cases in which sentencing behavior seems least consistent. Among other things, we consider whether racial bias is the source of the inconsistency (Section VII).

The paper ends with a discussion of the limitations of our analysis (Section X), with a comparison of this study with some other death-sentencing research (especially the related paper of Baldus, Pulaski, and Woodworth),[6] and with various tests of the viability of our classification model (Appendix C). Appendix A provides the scoring rules in their entirety; Appendix B provides several illustrations on their use.

---

6. Baldus, Pulaski, and Woodworth, supra note 2.

II. SOME HISTORICAL BACKGROUND

The present era in death-sentencing began on June 29, 1972 when the U.S. Supreme Court, in a 5-4 decision, struck down the capital statutes of Georgia (and, by implication, those of all other states).[7] While two justices found that capital punishment is inherently unconstitutional, the three others in the majority found that, as applied, the death penalty violated the Eighth Amendment prohibition against cruel and unusual punishment. The relatively few murderers put to death were chosen in so "freakish" and "arbitrary" a manner, and the discretion allowed to judges and juries seemed so unbridled that, to Justice Potter Stewart, "this death penalty is cruel and unusual in the same way that being struck by lightning is cruel and unusual."[8]

In the next four years, over 30 states revised their capital punishment laws so as to reduce the element of randomness in the distribution of death sentences. But on July 2, 1976, the Supreme Court struck down a North Carolina statute that provided mandatory executions of convicted first-degree killers (Woodson v. North Carolina).[9] The Court held that the law was too inflexible to be just, because it barred juries from considering "compassionate or mitigating factors" stemming from the "diverse frailties of humankind." The Court speculated that, rather than sentence certain defendants to death, juries would acquit them even if convinced of their guilt.

Having ruled out both too little discretion and too much, the Court indicated what kind of "intermediate" statute it would find acceptable. In Gregg. v. Georgia[10]

---

7. Furman v. Georgia, 408 U.S. 238 (1972).

8. Id. at 306-10.

9. Woodson v. North Carolina, 428 U.S. 280 (1976).

10. Gregg v. Georgia, 428 U.S. 153 (1976).

(also decided on July 2, 1976), the Court upheld the revised death sentencing laws of the state of Georgia. With approval, the Court noted that juries would be "permitted to consider any aggravating or mitigating circumstances," yet would be required to "find and identify at least one statutory aggravating factor" before sentencing a murderer to death.

The Georgia law[11] that the Court allowed to stand states that a defendant is "death-eligible" only if at least one of the following conditions applies[12]:

(1) The offense of murder was committed by a person with a prior record of conviction for murder, armed robbery, kidnapping or rape, or the offense of murder was committed by a person who has a substantial history of serious assaultive criminal convictions.

(2) The offense of murder was committed while the offender was engaged in the commission of rape, armed robbery, or kidnapping, or aggravated battery, or the offense of murder was committed while the offender was engaged in the commission of burglary or arson in the first degree.

(3) The offender by his act of murder, knowingly created a great risk of death to more than one person in a public place by means of a weapon or device which would normally be hazardous to the lives of more than one person.

(4) The offender committed the offense of murder for himself or another, for the purpose of receiving money or any other thing of monetary value.

(5) The murder of a judicial officer, former judicial officer, district attorney or solicitor or former district attorney or solicitor during or because of the exercise of his official duty.

(6) The offender caused or directed another to commit murder or committed murder as an agent or employee of another person.

(7) The offense of murder, was outrageously or wantonly vile, horrible or inhuman in that it involved torture, depravity of mind, or an aggravated battery to the victim.

(8) The offense of murder was committed against any peace officer, corrections employee or fireman while engaged in the performance of his official duties.

(9) The offense of murder was committed by a person in, or who has escaped from, the lawful custody of a peace officer or place of lawful confinement.

(10) The murder was committed for the purpose of avoiding, interfering with, or preventing a lawful arrest or custody in a place of lawful confinement, of himself or another.

---

11. See GA. CODE ANN. #27-2534-1 (1983).

12. The statute describes treason and hijacking as always susceptible to the death penalty. But we will consider no cases involving either.

Unlike many other states, Georgia does not accompany the aggravating factors above with a list of statutory mitigating factors. (Mitigating circumstances typically include the defendant's youth or the absence of a criminal record; Massachusetts even includes combat service in Vietnam.) Thus a Georgia jury[13] is under no obligation to treat any aspect of a case as mitigating. On the other hand, the jury is not compelled to give a death sentence even if many aggravating circumstances are present.

In other words, the new law does not foreclose the possibility that the same case could elicit different responses from different juries. In forthcoming sections, we will explore how the Georgia statute is actually being implemented.

III. FRAMEWORK OF THE STUDY

This study focuses on the general question: given the circumstances of a particular murder case, what is the probability the perpetrator will be sentenced to death? We are striving for an empirical answer based on the details and outcomes of a large number of Georgia cases. Given the size of the data base, our analytic approach does not follow the most familiar social science paradigm; we should discuss this circumstance at the outset.

Traditionally, one might create a set of N numerical variables $(x_1, x_2, \ldots, x_N)$ that, taken together, summarize the facts of the case. For example, x_j could be the defendant's number of prior convictions for violent crimes, while x_k could be an indicator variable equal to one if the defendant raped the victim and zero if he did not. Some of the factors would presumably pertain to the strength of the evidence presented in court.

---

13. Throughout this paper, we use the word "jury" as a generic term for the sentencing authorities. In reality, Georgia prosecutors are unusually influential in the process; in a considerable fraction of first-degree murder cases, the state simply waives the death penalty. And judges are important both in the rare cases in which the defendant waives a jury trial and, in more common situations, through their sentencing instructions.

Using some multivariate statistical technique (e.g. logit regression), one would then develop a mathematical formula that estimates p, the probability the defendant gets a death sentence, from the values of the x_i's that describe the case. The formula would be calibrated from data about a set of actual trials. Through scrutinizing the mathematical expression that arises, one can infer how juries are affected by the presence (or absence) of any given circumstance.

But statistical methods, however appealing in the abstract, can be problematic in particular settings. Any multivariate method for which a well-developed theory exists (i.e. any of those on the standard computer packages) entails a series of strong assumptions. Unless these assumptions are accurate, computations that depend on them can yield highly misleading results. One could wind up discarding variables of real importance, while embracing others that are actually irrelevant.[14]

While some potential problems are rather technical, others are not at all abstruse. Two of particular interest in the current endeavor concern statistically-correlated variables and the assumption of independent effects.

In a series of homicide cases, certain features might tend to arise in tandem. Every deliberate drowning, for example, might be preceded by the kidnapping of the victim. In every slaying of a bank teller, the victim might be white. When two variables tend to "move the same way" within the data set, it is hard to tell whether one of them is responsible for their combined effect and, if so, which, or whether both of them contribute and, if so, in what proportions.

---

14. The author has argued that such problems could well have afflicted studies on job discrimination and on the deterrent effect of capital punishment. See Barnett, The Deterrent Effect of Capital Punishment: A Test of Some Recent Studies, 29 Operations Research, p. 346 (1981); and Barnett, An Underestimated Threat to Multiple Regression Analyses Used in Job Discrimination Cases, 5 Industrial Relations Law Journal, p. 156 (1982).

Moreover, simple models assume that the various factors in a case independently exert influence on the jury's decision. But the very weight accorded a particular circumstance could well depend on which others are present. Killing by strangulation, for example, might in general increase the chance of a death verdict. But if a man, arriving home to find his wife in bed with another, proceeds to strangle the intruder, his mode of killing might heighten the jury's belief that he acted in the thrall of uncontrollable passsion. Hence, the very same factor could be deemed aggravating in some situations and mitigating in others.

While there are multivariate techniques aimed at coping with such problems, their success in the present context cannot be guaranteed. If two variables are correlated, one can try to partition their joint effect between them. But the procedure can give unstable answers, and is imperiled by the violation of any of several technical assumptions. And if the effect of two variables taken together differs from the sum of their individual effects (e.g. strangulation and lovers' triangle), the use of "interaction terms" in the model could reflect this. But if there are (say) 40 original variables, there are 780 possible pairwise inter-actions and 9880 possible three-way interactions. Thus the model could become unwieldly and its data requirements enormous.

For these reasons (any many others like them), the formal statistical methods might not be ideally suited to the problem we are studying. We could be in a situation where mathematical complexity and deeper understanding, far from being synonymous, could be negatively correlated. Thus it is not to favor the horse over the locomotive to conjecture that a person, working with actual case summaries and using common sense, _might_ gain greater insight into jury behavior than a computer that processes the data mechanically. With this possibility in mind, we begin discussing in the next section a fairly simple approach for analyzing the case records.

IV. A PROCEDURE FOR CLASSIFYING GEORGIA HOMICIDE CASES .

The data at our disposal consist of narrative summaries of several hundred homicide cases, accompanied by computer-coded information about the defendant, the victim(s), the circumstances of the killing, and the verdict reached.[15] In every case, the accused was charged with first-degree murder[16] under Georgia's present homicide statute adopted in March 1973.[17] We only consider those trials in which the defendant was convicted of homicide (though, in a handful of cases, of a lesser charge than first-degree murder). About one-sixth of these trials ended in death sentences; the rest resulted in prison terms, generally set at life.

To get a sense of how the minority of cases that led to death penalties differed from the majority that did not, the author began reading the summaries of blocks of cases of each type. In the initial overview, no attention was paid to the race of the defendant or the victim, to details about the victim's criminal history, or to the characteristics of the county in which the trial took place. As will become clear, however, these factors later entered the analysis.

The summaries made clear that homicide cases show immense variety. While certain elements are present in numerous records, the full constellation of circumstances in a given case is only rarely reproduced. Proportionality review, therefore, cannot realistically be based on comparisons across cases that have nearly identical facts.

In terms of the sentences meted out, however, certain consistencies did seem to come through. The death cases appeared to differ from the others on three

---

15 . We focus on the verdict in the initial trial, believing that the most useful reflection of prevailing community standards. However, virtually all the death sentences have been appealed; some have been reversed and almost all others are still being considered. Although the appeals and reversals clearly raise troubling questions, we will not consider them in this study.

16. We will not be considering the "filtering" process by which only a minority of solved killings lead to first-degree murder charges. That process could in itself be somewhat arbitrary, but it is not the subject of this inquiry.

17. GA. CODE ANN. #27-2537 (1983).

primary dimensions:

    (1) The Certainty the Defendant Is a Deliberate Killer

    (2) The "Status" of the Victim

    (3) The "Heinousness" of the Killing

These dimensions will be defined shortly.

It seemed useful to classify the various cases under a simple numerical scheme, in which three integers--one for each of the dimensions--reflected what appeared to be the case's most salient elements. To make the procedure as objective as possible, a detailed set of classification rules was prepared; they are presented in their entirety in Appendix A.

Readers will naturally want to know how the objectivity of the classification procedure was established, what predictive power it achieved, and whether its underlying assumptions were proved viable. But first they will want to know just what the procedure is. The remainder of this section is devoted to that concern; issues of reliability are discussed in some detail in Appendix C.

The Certainty the Defendant is a Deliberate Killer

The word "certainty" above refers to the degree of assurance that the accused was, in fact, the killer of the victim. (If substantial doubts existed, the defendant would presumably have been acquitted; the notion is, however, that the threshold of certainty needed for a death sentence is higher than that for a guilty verdict.) "Deliberateness" pertains to whether, even assuming the defendant performed the killing, he acted knowingly to cause the victim's demise.[18]

On this dimension, the case is related either zero (usually low), one (average), or two (unusually high). Zero reflects a relatively weak case in terms of certainty and/or deliberateness. A case based solely on circumstantial evidence, for example, would deserve this rating. Zero would also apply if the defendant was not the

---

18. We use the pronoun "he" because the heavy majority of both killers and victims are male.

triggerman, if he clearly suffered from mental illness, or if its details hint that the slaying was accidental. (As noted, the precise criteria for this and other ratings are set forth in Appendix A.)

A score of two, by contrast, signifies exceptionally strong evidence that the killing was not an isolated, aberrant act of passion or panic. If the defendant plotted the murder extensively, had previously tried to kill the victim, or was implicated in other killings, the case would be classified two.

If neither a zero nor a two is justified under the scoring rules, the case receives a rating of one. If, as happens rather rarely, criteria for both zero and two are satisfied, a score of one is also given. Among recent cases of first-degree murder in Georgia, most seem to warrant this intermediate classification: of the 513 the author rated, 67% scored one. (26% scored zero, and 7% scored two.)

The "Status" of the Victim

The "status" of the victim relates primarily to the relationship between the victim and the accused. Its presence in the classification scheme reflects an observed pattern under which, all other factors being equal, stranger-to-stranger killings are more "prone" to death sentences than those in which the victim knew the defendant.[19] The cases are scored either zero or one on the dimension of "status," with the latter number suggesting a higher chance of a death verdict.

The zero/one dichotomy is close but not identical to the stranger/nonstranger split. A stranger who acted in a highly provocative way just prior to his killing, or was engaged in an illegal enterprise, would call forth a rating of zero. And

---

19. This notion is hardly original with this study. The point is that the case records are in accord with an intuitive and widely recognized pattern.

even if the defendant and victim were acquaintances, the score one would be appropriate if the person was slain in his official capacity (e.g. as the defendant's supervisor in a factory, or a teller in a bank being robbed).

This status dimension has little explicit basis in law. It is hard to avoid speculating that, in killings in which jurors can imagine themselves or their loved ones as victims, death penalties are more likely to be imposed. To say this is not to impute cynicism to the juries; when a case evokes genuine fear, considerations of deterrence may more greatly affect the sentencing decision than otherwise.

Heinousness

On this dimension, the case is scored zero, one, or two, depending on the answers to the questions: "Could the killing be construed as an act of self-defense?" and "Was the killing vile?" The criteria for self-defense are quite stringent, requiring a clear mortal threat to the defendant or his loved ones. Among vile slayings are those with multiple victims, those preceded by psychological torture or sexual abuse, and those involving bizarre weapons or mutilated bodies.

A killing in self-defense that is not vile rates a zero, while a vile murder unrelated to self-defense rates a two. All other homicides are assigned the rating one; a one generally reflects the absence of both vileness and self-defense.

In summary, each case is classified with three separate numerical ratings. We will use the notation (i,j,k) where 'i' is the score on "certainty," 'j' on "status", and 'k' on 'heinousness." 'i' and 'k' can take on any of three values, while 'j' takes on two; thus there a total of 3x2x3=18 possible classifications. The individual ratings are arranged so that, the higher the score on a given dimension, the greater seems the empirical risk of a death sentence.

The least "deathworthy" score would be (0,0,0), which could arise if the defendant knew the victim, and killed him in self-defense in a manner seeming somewhat accidental. At the other extreme is a (2,1,2) case, such as the murder-for-hire of three police officers. Most robbery killings of a merchant would be classified (1,1,1); if, however, the victim took out a gun and fired at the defendant, (1,1,0) would probably be appropriate. We illustrate how cases are rated in Appendix B, where we classify four actual Georgia homicides.[20]

In the next section, we present the probability of a death verdict as a function of a case's classification. But various details, implications, and limits of the classification scheme are not apparent from the preceding brief description; for continuity, we postpone discussing them to Sections IX-XI.

V. DEATH SENTENCING RATES BY CASE CLASSIFICATIONS

The tables in this and subsequent sections are based on 513 Georgia homicide cases, all tried between 1973 and 1980 and all involving a first-degree murder charge. They pertain to about 20% of the killings in Georgia during that period, and a considerably larger fraction of the first-degree murder trials. (A small proportion of the cases in the Baldus file were not considered in this analysis; see Appendix C.) In all, the data concern an unusually large random sample from the relevant population.

Some death-sentencing rates appear in Table 1, where the cases are partitioned based on their (i,j,k) vectors. For convenience, we place in the same column all classes of cases with the same value of i+j+k. This arrangement is consistent with a general expectation that, as i+j+k goes up, so does the chance of a death sentence. But we are NOT suggesting that the content of the three ratings can be represented

---

20. The reader might do well at least to skim Appendices A and B now, lest the classification procedure seem needlessly obscure.

by their simple sum. A (1,1,1) killing differs in substantial respects from a (1,0,2), and there is no theoretical reason for assuming that their death-sentencing rates will be equal.

Table 1 reflects only an initial "sort" of the first-degree murder cases. As we will see, other data allow us to refine and clarify certain provisional numbers. Even without such elaboration, however, the table conveys a good deal of information.

TABLE 1:   DEATH-SENTENCING RATE AS A FUNCTION OF CERTAIN DETAILS OF THE KILLING

		(0,0,2) 0 (22)	(2,0,1) .17 (12)		
(0,0,0) 0 (8)	(1,0,0) .02 (42)	(1,1,0) 0 (10)	(1,1,1) .27 (48)	(1,1,2) .85 (53)	(2,1,2) .80 (5)
	(0,0,1) 0 (47)	(0,1,1) 0 (44)	(0,1,2) .20 (15)	(2,0,2) .53 (17)	
	(1,0,1) .01 (159)	(1,0,2) .21 (29)			

NOTE: Numbers in parentheses refer to the number of cases in the category. In two possible categories -- (2,1,0) and (2,1,1) -- there were no cases at all. (2,0,0) and (0,1,0) had only one case apiece; neither ended in a death verdict.

If the classification procedure had no discriminatory power, one would expect the death rates in all categories to hover around 16% (the overall rate for the cases considered). But Table 1 indicates that, as anticipated, there is clear positive relationship between i+j+k and the risk of a death verdict. Of the 334 cases in which i+j+k (hereafter defined as s) does not exceed two, a mere two of them elicited death verdicts. The death rate rose to 23% for the cases for which s=3, and when s≥4, fully 77% of the defendants were condemned to death.

Indeed, it is instructive to start at the category (1,1,2) and then, in three separate maneuvers, to reduce each of i, j, and k by one while holding constant the other two variables. The result, reiterated in Table 2, is a consistently drastic drop in the capital punishment rate. Such statistics lend strong support to our hypothesis that each of the three dimensions is of major importance in its own right.
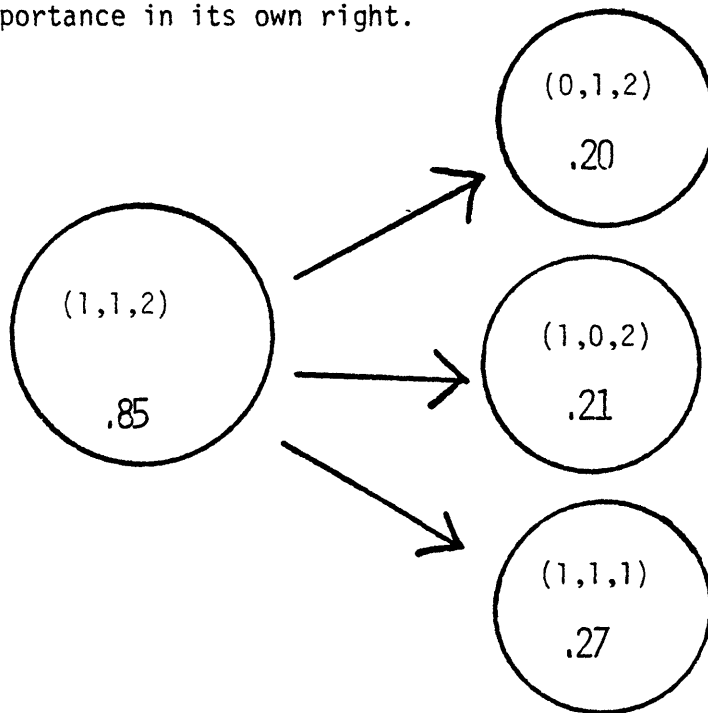

TABLE 2: DROP IN THE DEATH-SENTENCING RATE WHEN ONE OF THREE CASE RATINGS IS REDUCED

In Table 3a, j and k are held constant while i spans the values from zero to two. Table 3b displays some analogous data when k is varied across its range. The steady rises in death rates justify the decision to let i and k take on three values rather than just two.

(3a) $(0,0,2)$ 0 $\longrightarrow$ $(1,0,2)$ .21 $\longrightarrow$ $(2,0,2)$ .53

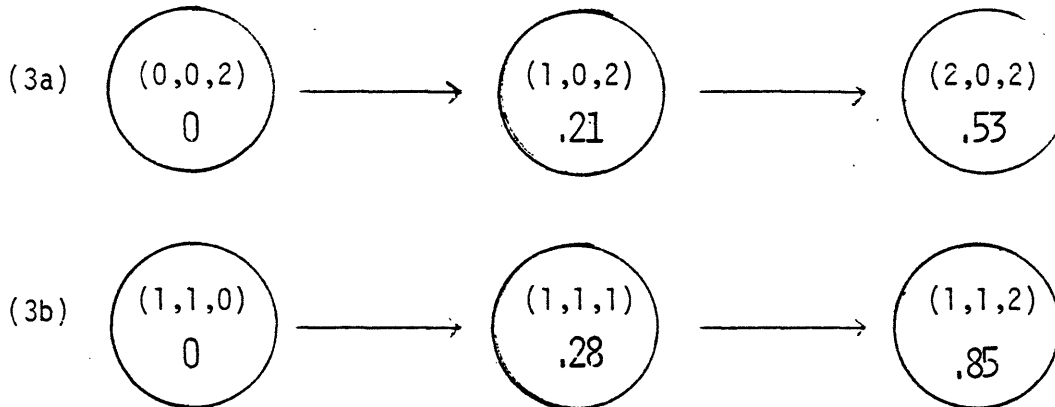(3b) $(1,1,0)$ 0 $\longrightarrow$ $(1,1,1)$ .28 $\longrightarrow$ $(1,1,2)$ .85

TABLE 3: THE EFFECT OF VARYING i AND THEN k WHILE HOLDING FIXED THE OTHER TWO PARAMETERS.

Of course, the classification rules were largely developed from the very data on which they were tested. Thus their success in identifying key patterns is not altogether surprising. But it was not at all foreordained that a simple numerical scheme would prove so effective. That outcome was contingent upon (and, indeed, shows the existence of) a considerable degree of regularity in the Georgia sentencing decisions.

The entries in Table 1 serve to focus our attention on Category 3: those classes of cases for which s=3. There the death rates, although low, are nonetheless well above zero. Before discussing this category further, we might do well to remind ourselves what kinds of killings it contains. Typical cases in the subdivisions of Category 3 might be:

$(1,1,1)$ The killing of an unarmed grocer with a single shot during a robbery.

$(2,0,1)$ The extensively plotted--though not especially sanguinary--killing of one's spouse for economic motives.

$(1,0,2)$ The killing of a long-term personal enemy through holding his head below water.

$(0,1,2)$ The kidnap-murder of a stranger, in which the defendant hit the victim but did not fire the shot that killed him.

Given the differences just outlined, it was not obvious that the death-sentencing rate would be fairly stable across Category 3. Yet the variations around the category-wide average of 23% are small and, given the small samples involved, nowhere close to statistically significant. This outcome is something of a "stroke of mathematical luck"; it is an empirical finding we did not anticipate and NOT a condition imposed in advance on the analysis.

More important, an obvious question arises about Category 3. Why is it that, for every such killing that leads to a death sentence, there are three others that do not? The 23% figure could reflect oversimplications in the classification rules, genuinely inconsistent behavior by different juries, or invidious distinctions (e.g. by race) that effectively divide otherwise homogeneous classes of murders. But the explanation could also be more benign, tied to the defendant's prior record or to regional differences in the adjudication of capital cases.[21]

We will explore these varied possibilities in the next few sections, starting with geography and criminal history.

VI. THE ROLE OF REGION AND OF CRIMINAL RECORD

From the Census Bureau's description of the county in which it took place, Professor Baldus classified each murder trial as either "urban" or "rural". The underlying idea was that, if attitudes on the death penalty do vary across localities, it would most likely be apparent on an urban/rural dimension. The differing viewpoints between such regions have been prominent in many a statewide election; one might suspect

---

21 . In Williams v. Maggio, 679 F. 2nd 381 (5th cir. 1982), the Fifth Circuit sustained the consitutionality of Louisiana's system of circuitwide -- rather than statewide -- proportionality review. See also Maggio v. Williams 104 S. Ct. 311 (1983) (order vacating stay of execution).

a similar divergence on a punishment issue that, in recent years, has often been correlated with more general political views.

One might measure the defendant's prior record by some complex mathematical function of the number and nature of his past offenses, as well as the time he spent in prison. But trying to devise an appropriate formula is uninviting, especially because it is hard to imagine that any juries actually used it. We therefore adopt a simpler approach.

For the present purposes, we reduce the issue of prior record to a simple yes/no question. The accused is said to have a serious prior record if he had been convicted for any violent crimes or if, while his only convictions were for property offenses, he served time in prison. The rationale for the latter condition is that, given the lenient treatment generally accorded early offenders, those incarcerated for property crimes presumably committed many of them. The defendant is said to have NO serious prior record if (as happens quite frequently) he has no convictions or if, although he does, they involve neither violence nor imprisonment.

With such definitions in hand, we can try to assess the role of region and criminal history in the death-sentencing decision. These variables cannot be pivotal when $s \leq 2$ because, regardless of where such a case stands in these respects, a death verdict is exceptionally rare. More detailed analysis is necessary in Categories 3 and 4 (i.e. s=3 or 4), which have both appreciable death-sentencing rates and nontrivial sample sizes.

As noted, the death rate is largely invariant across the subdivisions of Category 3. Thus it is not to blur salient distinctions to use overall Category 3 statistics in this discussion. We cannot proceed the same way in Category 4, for there is a large difference in the death-sentencing rates for (1,1,2) and (2,0,2) killings. The latter class has too few cases for the present purpose (there are only two urban/prior record (2,0,2) killings); hence we restrict our Category 4 study to the (1,1,2) class.

TABLE 4:   DEATH SENTENCING RATES BY LOCALITY AND DEFENDANT'S HISTORY:
CATEGORY 3 HOMICIDES

	URBAN	RURAL
PRIOR RECORD	.27(22)	.33(24)
NO PRIOR RECORD	.14(22)	.19(36)

(Sample sizes in parenthesis.)

TABLE 5:   DEATH SENTENCING RATES BY LOCALITY AND DEFENDANT'S HISTORY:
CLASS (1,1,2) HOMICIDES

	URBAN	RURAL
PRIOR RECORD	.86(7)	1.00(19)
NO PRIOR RECORD	.67(12)	.79(14)

Tables 4 and 5 present the relevant data. In all four geographic comparisons that were made (record/no record, Category 3, class (1,1,2), the rural areas had higher death-sentencing rates than their urban counterparts. The pattern is summarized by a (surprisingly) simple rule: when x is the urban death rate, the corresponding rural rate is about 1.2x. This difference, however, is relatively small. Hence, even if rural juries are "tougher" than urban ones, they only appear so to a limited extent.

By contrast, prior record seems to play a major role in the sentencing decision. As Table 6 (next page) points out, those convicted of Category-3 homicides who have serious records run twice the risk of a death sentence as the others. And in (1,1,2) killings, a death sentence is almost certain unless the defendant had no record of serious trouble. The numbers in Table 6 suggest that, among the eight (1,1,2) killers not sentenced to death, about 3/4 have their "clean" prior records to thank.

Although of interest on their own, these findings resolve only partially the "puzzle" of Category 3. In every contingency depicted in Table 4, somewhere between two and six times as many defendants avoid death sentences as receive them. We next inquire whether considerations of race might explain the discrepant outcomes.

TABLE 6: THE EFFECT OF PRIOR RECORD ON THE DEATH-SENTENCING RATE
IN OTHERWISE COMPARABLE CASES

CATEGORY 3   CLASS (1,1,2)

R .32   R .96

↑   ↑

NR .17   NR .73
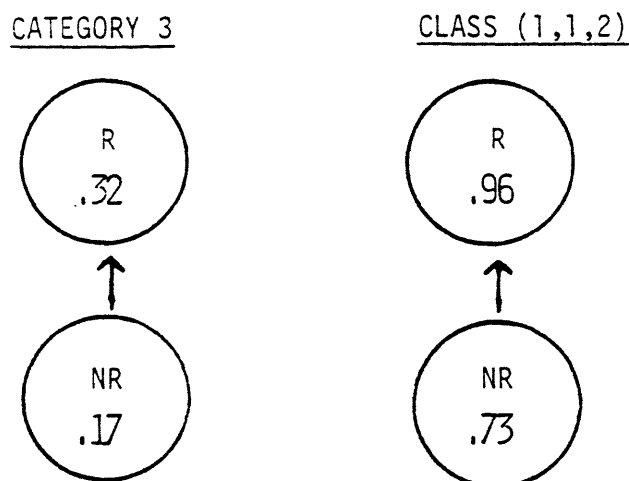
R = Prior Record; NR = No Prior Record

(Table 6 combines the urban and rural data from Tables 4 and 5.)

VII. THE ROLE OF RACE

Our interest in race as a factor in death sentencing is heightened by a
number of recent studies[22], almost all of which treat the race of the victim as a
major explanatory variable. Some of these investigations concern the very period
and state that we are examining. Gross and Mauro have estimated that, in post-
Furman Georgia, killers of whites face 7.2 times the death-sentencing risk of the
killers of blacks.[23] Baldus, Pulaski, and Woodworth concurred that "Georgia is

22. These studies include Baldus, Pulaski, and Woodworth supra note 2, Gross and
Mauro, supra note 4, Zeisel "Race Bias in the Administration of the Death
Penalty" The Florida Experience", 95 Harvard Law Review p. 456 (1981), Radelet
"Racial Characteristics and the Imposition of the Death Penalty", 46 American
Sociological Review p. 918 (1981), and Bowers and Pierce "Arbitrariness and
Discrimination under Post-Furman Capital Statutes", 26 Crime and Delinquency
p. 563 (1980).

23. Gross and Mauro, supra note 4, at p. 69. This multiplier of 7.2 is already
adjusted for certain differences in case characteristics.

operating a dual system, based on the race of the victim, for processing homicide cases. Georgia juries appear to tolerate greater levels of aggravation without imposing the death penalty in black victim cases; and ... the level of aggravation in black victim cases must be substantially greater before the prosecutor will even seek a death sentence."[24]

We might begin the present review with some macroscopic statistics. In the 513 cases examined, the defendant was black in 294 (or 57.3%). This last fraction was almost identical to the proportion of blacks among those defendants sentenced to death (58.1%). Thus, there is no immediate evidence of bias against black defendants. But, while 40.2% of the cases involved black victims, only 17 of the 84 capital cases did so. From these last two figures, we deduce that a factor of 2.6 separates the death-sentencing rate in white-victim cases from that in the others.

It is useful to disaggregate the murder cases according to their s-values which, as we have seen, are clearly related to the rate of death verdicts. Starting with a partition suggested by Table 1, we observe the following patterns:

TABLE 7: DISTRIBUTION OF MURDER CASE CLASSIFICATIONS BY RACE OF DEFENDANT[25]

s-VALUE	PERCENTAGE OF BLACK DEFENDANT CASES	PERCENTAGE OF WHITE DEFENDANT CASES
0-2	64.6	68.5
3	21.9	16.2
4-5	13.5	15.3

---

24. Baldus, Pulaski, and Woodworth, supra note 2, at p. 710.

25. In these calculations, we deleted a small number of cases in which the information about race was unclear.

TABLE 8: DISTRIBUTION OF MURDER CASE CLASSIFICATION BY RACE OF <u>VICTIM</u>[24]

s-VALUE	PERCENTAGE OF BLACK- VICTIM CASES	PERCENTAGE OF WHITE- VICTIM CASES
0-2	80.4	58.1
3	14.2	22.6
4-5	5.3	19.2

Note that the distribution of cases by s-value is almost independent of the
<u>defendant's</u> race. While there is a small excess of black defendants in Category 3,
that is counteracted by a smaller discrepancy of the opposite kind in the "death-prone"
Categories 4 and 5. Thus, we would not have expected an overall correlation between
the chance of a death sentence and the race of the accused and, indeed, we found none.

For race-of-victim, however, the situation is different. As proportions of
their respective total numbers, there are 1.6 times more white than black victims
in Category 3 killings. And in Categories 4 and 5, the corresponding multiplier
jumps to 3.6. Therefore, whites are disproportionately victimized by those kinds of
killings that most often evoke death verdicts. Perhaps these statistics form the embryo
of a <u>nonracial</u> explanation of the apparent importance of the victim's race.

Before pursuing this line of thought further, however, there is a possible
parallel we should consider. In some job discrimination cases, questions have been
raised about the value of performance ratings devised by employers. The objection
was made that such ratings, far from being neutral measures of employee achievement,
might be reflections of the very bias that was the subject of inquiry. In the present
context, the various distinctions made by judges, prosecutors, and juries, even if
expressed in terms unrelated to race, could still be manifestations of conscious or
unconscious racism.

---

26. In these calculations, we deleted a small number of cases in which the information
about race was unclear. Multiple killings with victims of both races were
exceedingly rare.

The author is no more qualified to make judgments on this matter than are readers. He would simply suggest a careful review of Appendix A, with special attention to whether the various criteria induce race-related effects, or vice-versa.

Delving into the data a bit further, we obtain the chart:

DEATH SENTENCING RATE BY RACE OF VICTIM:[27]
CATEGORIES 3 AND 4

	BLACK VICTIM	WHITE VICTIM
CATEGORY 3	.15(33)	.28(69)
CATEGORY 4:		
(1,1,2)	.86(7)	.84(45)
(2,0,2)	.57(7)	.50(10)

The victim's race seems rather unimportant in Category 4 but, in Category 3, the situation seems less clear. There the killers of whites are about twice as likely as the killers of blacks to get capital sentences. A more detailed breakdown of the Category 3 murders yields the following statistics:

TABLE 9: DEATH SENTENCING RATES BY RACE AND DEFENDANT'S RECORD: CATEGORY 3 HOMICIDES

	BLACK KILLS BLACK	WHITE KILLS BLACK	WHITE KILLS WHITE	BLACK KILLS WHITE
PRIOR RECORD	.20 (15)	NO CASES	.36 (11)	.44 (18)
NO PRIOR RECORD	.11 (18)	0 (1)	.25 (24)	.13 (15)

_____

27. Of the 104 cases in Category 3, two were deleted because of doubt about the race of the victim.

We see from these numbers that the higher death rates for the killers of whites are not explained by their worse prior records. Once normalized for the defendant's history, all rates in the table for black-victim slayings fall below their counterparts for white-victim cases.

There is some value to paying particular heed to the half of Category 3 murders that are classified (1,1,1). As murders go, such crimes are not especially bestial; nor do their perpetrators seem exceptionally cold-blooded. A typical (1,1,1) killing is of the homeowner during a burglary, of the taxi-driver during a robbery, or of the policeman trying to arrest the defendant. Self-defense, narrowly construed, is rarely an element in such cases, and rarely are the murders vile. (Vile robbery-killings of course exist, but they would generally be classified (1,1,2).)

Interestingly, the race-of-victim effect, roughly a factor of two for all Category 3 killings, is a factor of three in the (1,1,1) cases. And in those (1,1,1) killings committed by convicted felons, the ratio climbs to the value of five. To be more specific, there were 18 white-victim (1,1,1) cases in which the defendant had a serious prior record; 9 of them led to death sentences. But of the 10 such black-victim cases, only one led to the death penalty. Despite the small sample sizes, this disparity is statistically significant if viewed in isolation.

While these findings increase the sense that race is important, they are not wholly unequivocal in their message. If race-of-victim seems far more important in prior-record (1,1,1) killings than in Category 3 as a whole, race is therefore of diminished importance in the remainder of the category. The subset of cases in which the factor -of-five prevails contains only about 5% of the 513 murders studied. Even if race were irrelevant throughout the data set, apparently strong effects could arise in small subsets by chance alone.

But while not without some force, such an argument might not be compelling here. The (1,1,1) murders committed by prior felons are not just a random subset of the universe of homicides. They bring forth anger and a great deal of fear and, because they are more "rational" than other homicides, they might more plausibly be deterred by capital punishment. Yet they are not so gruesome as to be hideous murders per se. Against such conflicting pressures, the race of the victim might attain greater importance than in more "clear-cut" situations.

There could be at least two different sources of such a racial pattern. Zeisel has suggested that, because blacks are far more hostile to the death penalty than whites, their insistence on it in black victim cases might nowhere approach the comparable white attitude.[28] Especially in a "borderline" category like (1,1,1), such a difference might influence prosecutors pondering whether to seek the death penalty. And, as Gross and Mauro have noted, predominantly white prosecutors and juries might feel special sympathy for white victims resembling themselves.[29] The limited data we have studied cannot in themselves ratify any such theories.[30]

A summary of our findings about race might go as follows: Salient differences in the details of the killings of blacks and whites could to a considerable extent, explain the higher rate of death sentences in the white-victim cases. But in a limited fraction of cases--exemplified by the robbery-killing of a merchant--the race of the victim might matter a great deal. Thus, while it could idealize the murder trials to call them colorblind, it might caricature them to speak of pervasive, virulent racism.

---

28. Zeisel, supra note 20, p. 467.

29. Gross and Mauro, supra note 20, at p. 115.

30. Zeisel and Zimring, Eigen, and O'Malley (supra note 42) also mention "social distance" theories under which the slayings of "high-status" people by those with lower status would bring forth the harshest punishment. But in class (1,1,1), the social status of the victims (taxi drivers, merchants, etc.) is neither exceptionally high nor correlated with race.

VIII. THE CATEGORY 3 DILEMMA

Despite our efforts in the last few sections, we have been unable to uncover even one subdivision of Category 3 in which a majority of trials ended in death sentences. While we will not discuss formally the legal ramifications of this outcome, we might do well to speculate briefly on how it arose.

Even when acting under statutory guidance, individuals will differ in their assessments of when death sentences are warranted. Research summarized by Kadane, for example, has suggested that only 1% of Americans would impose a death verdict in every sentencing context in which the option exists.[31] Most people apparently have thresholds that separate the "deathworthy" situations from the others.

In Georgia, there is overwhelming agreement that death is not the appropriate penalty for killings with s-values lower than 3. And there is a clear consensus that death _is_ appropriate in the classes (1,1,2) and (2,1,2). (The consensus in a bit shaky in the class (2,0,2), but the "halfway" death rate of 50% is passed.) In Category 3, which contains 20% of the killings and 30% of the death sentences, matters are far less settled. It could be that many of the personal thresholds that divide "life" from "death" cases fall within that category's boundaries.

Yet the outcomes of the Category 3 trials, viewed collectively, might define a public position on such killings far more clearly than do general statues or vaguely worded opinion polls. And that position seems decisively to _reject_ the use of the death penalty in the cases in Category 3. Those perpetrators of such murders sentenced to death can, with some justification, view themselves as unlucky: they received harsher sentences than the heavy majority of their "peers."

---

31. Kadane, "Juries Hearing Death Penalty Cases: Statistical Analysis of a Legal Procedure," 78 Journal of the American Statistical Association p. 544 (1983).

To be sure, the people might prefer a more complex sentencing strategy than an "all or nothing" approach. While they might recoil from putting to death every robbery-killer, they might nonetheless support the occasional execution of such a murderer; such sporadic acts, it might be reasoned, might keep alive a flicker of deterrence. But such deliberate caprice in the sentencing policy would seem to offend present legal doctrine, and is precisely what proportionality review is meant to prevent.

Suppose, for argument's sake, that all Category 3 death sentences in Georgia were vacated. If our classification scheme makes sense, this act <u>alone</u> might greatly reduce the arbitrary element in the Georgia death sentencing.[32,33] (The two aberrant death sentences in Categories 1 and 2 would also presumably be vacated.) Yet Georgia would still have a death penalty rate of 11% for first-degree murders; this would entail something like eight executions per year. Georgia might be able to satisfy the requirements for proportionality review, therefore, without coming anywhere close to abolishing capital punishment.

We would reiterate that, according to their lights, Georgia's prosecutors, judges, and juries seem to behave with a fairly high degree of consistency. The variance of outcomes in the trials of "similar" first-degree murders is probably far less than the dispersion of attitudes among the citizenry of Georgia. But we should say more about the norms around which this consistency takes place, and do so in the next section.

---

32. Again, we remind readers that we have not considered possible anomalies in the process under which murder charges are filed.

33. Note that, if the Category 3 death sentences were vacated, the racial disparities suggested in the last section would also diminish greatly.

IX. SOME GENERAL POINTS

The classification rules in Appendix A contain a large number of qualifying

remarks, such as:

(i)   "If the only evidence for self-defense is the defendant's
      uncorroborated claim, assume its absence..."

(ii) "Neglect references to insanity if the defendant has no apparent
      medical history."

(iii) "The killing has an 'accidental' touch about it, because ...
      (Three specific conditions are set out)..."

(iv) "Give a rating of zero on 'certainty' if it seems clear that the
      defendant neighter ordered the killing nor was the triggerman.
      (Note that this differs from the weaker statement that it is
      uncertain whether the defendant was the triggerman.)"

In a substantial fraction of homicide cases, the accused contends that he

did not kill the victim or, that, if he did, the killing was an accident, and

act of self-defense, or the consequence of temporary or longer-term mental illness.

What (i)-(iii) imply is that juries tend to greet such claims with skepticism, and

to give them little weight unless strong evidence supports them.

The caveat in (iv) deserves elaboration. Suppose three accomplices commit

a robbery-homicide, but it is unclear which one killed the victim. We are suggesting

that juries will disregard this indeterminacy in setting punishment for any one of

the defendants. Perhaps there is an inchoate fear that to do otherwise would allow

the co-perpetrators, through a strategy of collusion, to reduce the total punishment

that they receive. It is noteworthy in this connection that, in 1982, Texas executed

a participant in a robbery-killing even though it was never established whether he

or his partner fired the fatal shot. [34]

In some respects, the rating procedure is as revealing for what it leaves

out as for what it includes. It makes no reference to the age of the defendant,

---

34. See The New York Times, December 6, 1982, p. A16, and December 7, 1982
    p. A1 and A19.

to his being retarded or killing while intoxicated, or to his showing remorse after the act. Given that Georgia has no statutory mitigating factors, the disregard of such details might not be inherently improper. Still, their apparent insignificance in the punishment decision--coupled with the exacting definitions of self-defense, accident, and mental illness--suggest a pattern: given flexibility about how to interpret the word "mitigating," prosecutors and juries incline towards narrow rather than expansive solutions.

The circumstances that increase the risk of a death sentence are related to, but do not coincide with, the aggravating factors in the Georgia homicide statute. Unlike the statute, the classification rules distinguish whether the felony accompanying a homicide is a robbery or a rape. And while the statute imparts special significance to the killing of a policeman, the rules treat as equally aggravating the slaying of any person in his professional capacity. The statutory reference to "great risk to a large number of persons" seems to have little practical importance; the phrase "wantonly vile," by contrast, seems exceedingly significant. Indeed, the rules could be construed as providing a detailed definition of the concept.

Certain provisions of the classification rules do not have their roots in the statutory aggravating factors. For killings in a private home, for example, the statute makes no reference to the number of victims. But so long as some statutory aggravating factor is present (e.g. the element of robbery, the defendant's prior record), the jury could cite it to justify a death verdict that it really believes is warranted for a broader set of reasons. Thus differences between our rules and the statute reflect not contradictions, but rather the exercise of a certain flexibility that is built into the law.

We should stress that the classification scheme is intended to model the general tendencies in the data. Thus, we are not suggesting, for example, that there was never a case in which a jury felt mercy towards an intoxicated defendant. But this caveat does not absolve us of the need to verify the assumptions embedded in the scoring rules, a task we attempt in Appendix C.

X. LIMITATIONS OF THE ANALYSIS

Several potential problems pose threats both to our specific findings and to any generalizations that might be drawn from them. The hazards deserve explicit recognition.

First of all, there are questions about the accuracy of our perceptions on individual cases. Professor Baldus and his associates did an exemplary job of gathering data, but narrative summaries and computer variables cannot possibly illuminate everything of interest. Brief case descriptions can only crudely portray the credibility of the various witnesses, the eloquence of the opposing attorneys, and the emphasis placed in court on the various elements of the case. While perfection in this regard is unattainable, one cannot deny that its absence could have adverse consequences.

Moreover, we have only considered the minority of homicide cases in which the defendant was charged with first-degree murder. If, in the setting of charges, prosecutorial discretion is exercised in an arbitrary manner, Georgia's overall punishment structure might be far less comprehensible than the one we have depicted.[35] Should a large number of cases we would have classified (1,1,2), for example, have elicited only manslaughter charges, our 85% death rate in that class could be highly misleading.

But to acknowledge these problems is not to conclude that they are genuinely damaging. We sampled, for example, from the full set of first-degree murders, not just those that led to a penalty trial. Thus, for a (1,1,2) killing to have escaped our attention, it would have to have been "downgraded" past scores of "lover's quarrel" and "barroom brawl" slayings (involving both white and black victims). That such distortions occur with regularity is not self-evident.

---

35. This is a special case of the problem of "sample selection bias," which is discussed in Berk, "An Introduction to Sample Selection Bias in Sociological Data," 48 American Sociological Review p. 386 (1983).

It is unclear how far one can generalize either our particular results or the method by which they were obtained. It should be remembered that the scheme is exploratory in nature, emerging from the Georgia data rather than some a priori theory. For this reason, the classification procedure itself--and not just the conclusions arising from its use--might not survive a transplant to another state. (That possibility could, of course, be the subject of further investigation.)

Indeed, there are reasons to suspect that Georgia is not fully comparable to (say) Pennsylvania, Kansas, California, or even Louisiana, Its relatively high death-sentencing rate might alone suggest less flagrant inconsistency than in other states where capital verdicts are far rarer. And the asymmetric manner in which its statutes treat aggravating and mitigating circumstances (seeming to deemphasize the latter) could induce different sentencing behavior than other, more "balanced" laws. To be prudent, we should bound the region in which our findings apply with the Georgia state line.

But there is another limitation of this work that transcends such practical problems. We have tried to depict how Georgia's judges, juries, and prosecutors actually behave, which is quite separate from the issue of how they should be behaving. In the language of social science, this is a descriptive study and not a normative one. As such, it is not meant as some simple blueprint for proportionality review or anything else.

We have suggested that Georgia sentencing behavior is somewhat consistent. But even if consistent, the distinctions made by juries could well, in Kalven and Zeisel's phrase, be demeaningly trivial compared to the differences in punishment that they entail. In any case, consistency need not assume some preeminent status in the distribution of punishment. Suppose (to put the matter graphically) that juror choice in homicide cases were restricted to death or probation.[36] There

---

36. This example was suggested by Professor Frank Zimring, to whom I am grateful for many sharp observations about the uses and limitations of this work.

would probably be clear regularities in the sentencing outcomes, but they could never make defensible the monstrous choice imposed on the jurors.

This paper did not aspire to -- nor could it ever -- provide a justification for death verdicts. While we have furnished empirical evidence concerning certain attributes of just sentencing, about numerous others we have nothing useful to say.

XI. RECENT RESEARCH

This article is not the first on the actual patterns in post-Furman death sentencing. Here we briefly review some of the others and discuss this paper's relationship to them. After surveying several papers, we focus on the article by Baldus, Pulaski, and Woodworth, and then consider the present manuscript in the context of such literature. All the other studies to be cited are both thoughtful and worthwhile; to the extent that our brief summaries tend to emphasize imperfections, it is as part of the explanation why the present paper is not superfluous.

Broadly speaking, empirical studies of capital sentencing can be described as either "classical" or "exploratory." Classical efforts begin by specifying in advance the possible determinants of sentencing behavior. Then, through some statistical method, they calculate the frequency of capital verdicts as a function of these determinants. Implicitly, such studies estimate the effect on the sentencing outcomes of each of the explanatory factors considered.

The exploratory studies give the data more latitude to "speak for themselves." In a manner that is initially highly unstructured, their authors peruse the summaries of a variety of homicide cases. Then they try to characterize how the relatively few trials that led to death sentences differed in their details from the others.

(i) Seven Papers

Among the classical papers on post-Furman patterns, those of Radelet, Zeisel, Bowers and Pierce, and Gross and Mauro are prominent.[37] All of these scholars were concerned primarily with racial disparities in the imposition of capital punishment. Statistics from the state of Florida were presented in all four papers; Bowers and Pierce also had data for Georgia, Ohio, and Texas, while Gross and Mauro studied a total of eight geographically-dispersed states.

Beyond considering the races of the defendant and the victim, the authors subdivided their cases on a few aggregate dimensions. Bowers and Pierce and Zeisel only inquired whether the slaying was a felony killing (i.e. one committed in connection with a separate felony, like robbery). Radelet employed a victim-status dichotomy, based on whether the person killed was a close friend, lover, ex-lover, or family member of the accused. After several cross-tabulations of death-sentencing rates by race and one other variable, Gross and Mauro culminated their work with a logit regression analysis that included five nonracial factors[38]:

(1) whether the homicide was committed in conjuction
     with another felony

(2) whether the victim and defendant knew one another

(3) whether there were one or several victims

(4) whether there were any females among the victims

(5) whether the killing was committed by gun

---

37. See supra note 20.

38. The analysis did allow for the possibility of some interactions among
     these variables.

These papers uniformly suggested that, as applied, capital punishment was
largely a response to the taking of white lives. Certain secondary patterns arose
in some of the papers more than others. Two articles, for example, implied that
black defendants faced higher death risks than similar white killers (Zeisel,
Bowers and Pierce).

In assessing such studies, we should not overlook the sweep of their
simplifying assumptions. Even in the comparatively-detailed Gross and Mauro analysis,
a case gets the same rating whether the felony in a felony killing was robbery or
firebombing, whether the victim was the defendant's spouse or the defendant's
clergyman, or whether the deceased was stabbed with a knife or mutilated with an
icepick. None of the researchers considered the defendant's prior record or the
possibility of self-defense. And strength-of-evidence is deemphasized to the point
that even cases in which the defendant was acquitted remain with unknown consequence
in the data bases.

One naturally wonders about the cumulative effect of such problems. Gross
and Mauro concede that killings against whites are on average more "aggravated"
than those against blacks.[39] Is it not conceivable that, within the broad categories
of homicides they lump together as similar, those with white victims are likewise
more aggravated? Radelet acknowledges[40] that "the strength of the racial disparities
observed in this study will fluctuate as other potentially relevant variables are
introduced." But until such factors have been given appropriate weight, how can one
be sure that their effects are mere "fluctuations"?

---

39. Gross and Mauro, supra note 20, at page 115.

40. Radelet, supra note 20, at page 926.

In summary, these studies raise serious and troubling questions about racial bias in sentencing. But that they unequivocally answer such questions seems less clear.

Among the exploratory studies, some that are noteworthy concern the pre-Furman era. Kalven and Zeisel[41] considered 35 homicide cases from the early 1960's in which the judge and/or the jury favored a death sentence. In 21 of the 35, the judge and jury split on whether the defendant should be executed.

The case reviews suggested that a consensus for death was especially likely if a killing involved multiple victims, sexual torture, or a bizarre weapon. Factors that evoked leniency in the disputed cases included the defendant's not being the actual killer, a lover's triangle aspect, an earlier noncapital verdict for another participant in the slaying, and a "worthless" victim.

To Kalven and Zeisel, the patterns just cited are less important than another. When judges and juries hearing the same case under the same statute so often disagree on the life-or-death decision, consistency in the selection of those to be executed might be a farfetched concept.

Zimring, Eigen and O'Malley[42] studied the Philadelphia homicides in the first few months of 1970. Only three such killings led to death sentences; all involved the slaying of whites by black strangers. Two of these killings were felony-murders, one was committed by grotesque means (a hacksaw and a sledgehammer), and one was clearly a manifestation of racial hatred. The authors recognized that these murders

---

41. Kalven and Zeisel, "A Somber Postscript: Decisions on the Death Penalty," in The American Jury Little-Brown, p. 434, (1966).

42. Zimring, Eigen, and O'Malley, "Punishing Homicide in Philadelphia: Perspectives on the Death Penalty," 43 U. Chicago Law Review, p. 227 (1976).

were different from most others, but noted an inconsistency between their death verdicts and the outcomes for other aggravated killings (e.g. multiple murders) and also those for other defendants in the three cases themselves.[43]

In the post-Furman period, an exploratory study of Dade County, Florida was performed by Arkin.[44] He compared ten felony killings that led to capital sentences (three of them given to the same defendant) with 44 felony killings that resulted in prison terms. The death cases, he argued, were especially aggravated; they were clearly distinguishable from a majority of the others and at least somewhat different from 38 of the 44. Thus Arkin saw considerable (though not perfect) selectivity in Dade County's death sentencing and stated that any inference of racial discrimination "collapses"[45] under scrutiny.

A potential problem with these exploratory studies relates to their small sample sizes. The danger is that the discussion of the cases, however illuminating, could be construed as more speculative than systematic. For example, Gross and Mauro, having noted that Zimring et al had "only three death sentences,"[46] went on to contradict Arkin, contending[47] that "because of the small size of his sample -- ten death penalties in all -- no definite conclusion could be reached on the existence of racial discrimination."

Taken together, these seven classical and exploratory efforts suggest an unwritten convention in this literature. When the number of cases is small -- and statistical procedures would founder on the paucity of data -- the intuitive

---

43. The authors also compared cases that led to life imprisonment to the more numerous others that yielded far shorter terms. They discerned a race-of-victim effect, and contrasted the great disparities in sentence lengths with the seemingly lesser differences in the homicides themselves.

44. Arkin, "Discrimination and Arbitrariness in Capital Punishment: An Analysis of post-Furman Murder Cases in Dade County, Florida 1973-76," 33 Stanford Law Review p. 75 (1980).

45. Id.,at p. 88.

46. Gross and Mauro, supra note 20, at page 20.

exploratory approach is viewed as legitimate. But as the sample size reaches into the hundreds, more formal methods involving broad, prespecified categories are usually brought to bear. The result is something of an inverse relationship between the number of cases studied and the level of detail per case, which can lead to the persistent sense that "something is missing."

(ii) The Baldus-Pulaski-Woodworth Paper

Against the backdrop of the research just cited, the work of Baldus et al, plays an important role. For various Georgia killings from the 1970's that led to first-degree murder charges, the authors gathered both narrative summaries and data that concern roughly 200 variables. With this vast information -- which was obtained for several hundred cases -- they sought to describe Georgia sentencing behavior.

At the center of their efforts were a series of multiple regression analyses.[48] These regressions assigned a numerical score to each case meant to relate systematically to the chance the defendant would receive a death sentence. One major scoring rule that emerged was based on about two dozen variables: starting at zero, one added or subtracted a series of numbers depending on which factors were relevant to the case. An insurance motive, for example, increased the score by .17, the kidnapping of the victim by .10, and a murder-by-drowning by .12. In contrast, the score dropped by .08 if the defendant was not the triggerman, and by .03 if the victim had "low status." Nineteen other circumstances would also affect the score.

This approach to processing the data certainly possesses statistical power. Any qualms about it are traditional ones concerning the limits of regression analysis.

---

48. A good introduction to regression analysis appears in Wonnacott and Wonnacott, Introductory Statistics for Business and Economics chapters 12-13, Wiley (1972).

In Section III, we discussed some potential troubles (correlated variables, the assumption of independent effects); other difficulties could involve definitions of variables, certain linearity assumptions, and the tendency of ordinary-least-squares methods to weigh some data points more than others.[49,50] The authors also note the danger of "overfitting" (i.e. the inclusion through chance of noncausal variables in the regression equation).[51]

The study itself provides evidence that such problems might not be mere technicalities. The scoring formula cited arose from averaging three regression analyses that were calibrated from overlapping data. Three regressions were necessary, the authors explained, "because of the tendency of each analysis to produce a unique solution which omitted obviously important and relevant variables. For example, the following variables were omitted from one or more of the models: number of prior felony convictions, victim a hostage, number of convictions for violence, defendant not the triggerman, defendant created great risk in a public place, and insurance motive."[52] Moreover, some variables appeared in initial regressions with coefficients whose signs seemed absurd.[53, 54]

Quite beyond such matters, there is the issue of how to interpret the regression results. The scoring rule identifies some major influences on sentencing and suggests

---

49. See Belsley, Kuh, and Welsch, Regression Diagnostics: Identifying Influential Data and Sources of Collinearity Wiley (1980).

50. The authors also use logistic regression analysis to duplicate certain results; this technique produces scoring rules that, roughly speaking, are multiplicative rather than additive. Though perhaps more natural than linear regression for the estimation of probabilities, it is subject to variants of all the problems cited.

51. Baldus, Pulaski, and Woodworth, supra note 2, note 106 at page 695.

52. Id., note 98 at page 689.

53. Id., note 98 at pages 689-90.

54. Because regression results are subject to an analogue of sampling error, these problems are not inherently devastating. But they might be warning signs that something is amiss and, if that is the case, averaging several regressions together is unlikely to eliminate the trouble.

their relative importance, but it leaves far less transparent what death rates result from the interplay of various circumstances. For example, the authors define Index Category IV -- those first degree cases with death risks near the overall Georgia average -- as all cases with final scores between .23 and .36. There are literally thousands of combinations of elements that would assign to a case a score in this range. Thus a simple characterization of the killings with "average" death risk would seem very hard to devise.

Indeed, two cases with the same score (e.g. .25) need have no overlap at all in the contributions to their common rating. One case could involve a web of aggravating and mitigating factors, while the other is comparatively nondescript. Thus statistical similarity need bear no clear relation to conceptual similarity. The scoring rule might do well at predicting jury behavior without making clear the interaction of forces that generates that behavior.

These statements do not imply that the scoring method is terribly flawed, nor do they indicate that this paper's approach -- which has problems of its own -- is clearly superior. But they do suggest that another try at analyzing the Georgia data was not utterly redundant.

(iii) The Present Paper

This manuscript tries to bring the spirit of exploratory data analysis to a large scale data base. With no prior structure having been imposed on the analysis, the case summaries were read individually. Ultimately, lots of single observations were merged into a succinct theory (though the detailed implications of that theory, as Appendix A suggests, might not be especially succinct).

The approach used here is not without genuine hazards. It is farfetched if not preposterous to suggest that anyone comes to the Georgia cases with no prior

conceptions whatsover. And there is the danger of impressionism, of giving excessive weight to a few unusual cases, or of such pedestrian problems as lapses in memory. There is even the unsettling possibility that the order in which the cases were read influenced the conclusions reached about them.

Accompanying such dangers, however, is the opportunity for a more textured, more nuanced view of sentencing behavior. There are enough cases that there is a good prospect of grasping underlying patterns and, to some extent, of testing hypotheses about them (See Appendix C). And one retains an alertness to unanticipated "signals" from the data, an alertness that can suffer when the records are processed en masse with some predetermined method. For these reasons, certain weaknesses of both the classical studies and small-scale exploratory work might be diminished in the present effort.

The partition of cases that resulted from this exercise is conceptually simple. A particular class like (1,1,2) or (1,0,0) is fairly easy to describe, and the description is sufficiently "tight" that the class's member cases differ relatively little. Thus the calculated death risks by class (sometimes amplified by prior record, region, and race-of-victim) suggest with some clarity how juries respond to each juxtaposition of primary factors.

The detailed results suggest that conceptual simplicity was not achieved at the expense of explanatory power. A small set of dimensions identifies some killings as having a high likelihood of a death sentence, and others for which that outcome is almost inconceivable. And, identified by these same dimensions as falling between these two "poles," one finds the important minority of Georgia cases in which the facts lead to death sentences only one time in four. Hence the "margin" for death decisions--as determined by opposite sentencing patterns on opposite sides of it -- turns out to be the place where sentencing outcomes are, in fact, least consistent.

How do our major findings compare to those in the other articles? Interestingly, almost all death cases described in the earlier exploratory work seem to warrant s-values of 3 or greater. Thus killings beyond the reach of the death penalty in today's Georgia (i.e. those with $s \leq 2$) seem similarly so in the other places and times. And, as in Georgia, some earlier cases falling in Categories 3 and 4 did not evoke death verdicts. Terseness in case descriptions and small sample sizes preclude a more precise comparison of the earlier patterns with our own.

The post-Furman classical studies, as we have noted, attached great weight to the race of the victim. While not contending that race is immaterial, the present paper nonetheless accords it a more restrictive role. Part of the difference may arise because we could not search for racism in the filing of homicide charges.[55] On the other hand, our findings -- as well as those of Arkin -- hint that as one delves deeper into the details of the cases, race might often emerge as less important than aggregate statistics might suggest. ("Often" is not meant here as a euphemism for "always"; we remind readers of the factor-of-five that arose in the (1,1,1) prior record cases.)

Our results and those of Baldus, Pulaski, and Woodworth are generally in agreement. The two papers tend to concur on what factors, viewed themselves, the juries find mitigating and aggravating. This is not to say the results are identical: we gave greater weight to self-defense and weakness of evidence, and we offered a somewhat more specific assessment of where racial discrepancies might be concentrated. But given their dependence on very different methodologies with different limitations, the degree of consensus the studies achieved is noteworthy. In a sense, each paper serves to increase the credibility of the other.

---

55. Radelet (supra note 20) concluded that there are greater racial disparities in the setting of homicide charges than in the courtroom proceedings that follow. But, beyond race, he only considered whether a close relationship existed between the defendant and the victim. Thus, he neglected the vast bulk of the information that the prosecutor had at hand in making the charging decision.

FINAL REMARKS

The analysis described in this paper suggests that, broadly speaking, Georgia's first-degree murders fall into three groups. In the first and largest, death sentences are so rare as to resemble the "strikes of lightning" that led to Furman v. Georgia. In another, comprising about 1/7 of the cases, it is those spared a sentence of death whose treatment appears unusual. But about 1/5 of the cases fall into an intermediate range, in which death verdicts are neither especially rare nor especially common.

It appears that the general indicators of a defendant's death risk are the certainty that he killed and did so deliberately, the victim's status with respect to the accused, and the extent to which the slaying was heinous in its details. In murders whose facts make a death penalty a serious possibility, the defendant's prior record seems important, as does the kind of county where the case was tried. While such variables(and, sometimes, the race of the victim)explain a good deal about sentencing, there remain some situations in which death cases are outnumbered by close counterparts with other outcomes. Especially in such settings, the concerns about equitable treatment that animate proportinality review could well be warranted.

An empirical exercise like this one cannot yield incontrovertible truths. Rather it presents a theory on how Georgia selects these homicide convicts it condemns to death. The soundness of any such theory depends on its innate plausibility, on the appropriateness of the data set that yielded it, and on its explanatory power within that data set.

Readers will have to make their own assessment of the soundness of the theory presented in this paper. And if persuaded of its viability, they must face a deeper question: is the article more comforting or is it more disturbing in what it suggests about post-Furman death sentencing?

APPENDIX A: CLASSIFICATION RULES

I. THE CERTAINTY THE DEFENDANT IS A DELIBERATE KILLER

Score the case either 0, 1, 2 on this dimension, applying the following criteria:

(i) The case is rated 0 if any of the following circumstances pertain:

   (1) The narrative indicates the evidence in the case seemed weak
        (e.g. "case based solely on circumstantial evidence").

   (2) The narrative mentions evidence that worked against the view
        that the defendant was guilty. (e.g. tests for residue on
        defendant's hand from firing a gun were negative.)

   (3) It seems clear that the defendant neither ordered the killing
        nor was the triggerman. (Note that (3) differs from the weaker
        statement that it is uncertain whether the defendant was the
        triggerman.)

   (4) The killing has an "accidental" touch about it, because

        (a) a fairly long period (perhaps a week or more) elapsed
             between the incident and the victim's death, or

        (b) the death was caused by a shot fired somewhat randomly
             (e.g. through a door), or

        (c) the death was caused by a beating similar to previous
             beatings of the victim by the defendant

   (5) There is reason to doubt that the defendant's actions in
        themselves would have caused the victim's death (e.g. (i) the
        defendant beat the victim, but it was a co-perpetrator's
        stabbing that killed him or, (ii) the defendant's beating of
        the victim induced a heart seizure.)

   (6) The narrative mentions that the defendant was previously treated for
        mental problems. (e.g. institutionalized) Neglect references to
        insanity if the defendant has no apparent medical history.

(ii) The case is rated 2 if any of the following elements were present:

(1) The killing was a murder-for-hire. and the defendant was either the instigator or the executioner.

(2) The defendant plotted to kill the victim (e.g. a wife and her lover arrange to murder her husband). If, however, the defendant was one of several plotters, and clearly not the actual killer, assume (2) is not satisfied.

(3) The narrative mentions that the defendant was implicated in other killings.

(4) The narrative mentions that the defendant had tried previously to kill the victim.

(5) The defendant announced in advance to a third party an intention to kill the victim. (In a "lover's triangle" case, neglect this condition.)

(iii) If the killing warrants neither a 0 nor a 2, give the case a rating of 1.

If the killing satisfies conditions for both 0 and 2, also rate it 1. Most "common" slayings such as killings during armed robberies or during barroom fights would warrant this intermediate classification. Indeed, a 2 reflects unusually clear evidence of premeditation, while a 0 reflects unusually large doubt that the defendant knowingly acted to cause the victim's demise.

II. THE STATUS OF THE VICTIM

On this dimension, the score is either 0 or 1.

Give a score of 0 if:

(1) The victim was a relative of the defendant (even his or her child).

(2) The victim was a friend of the defendant. (Interpret the word "friend" loosely; if, for example, two people of similar age are riding together voluntarily in a car, consider them friends. HOWEVER, the mere fact that two people know each other is not sufficient. Neighbors of vastly different ages, or the bank teller and the depositor, are not assumed friends barring other evidence of social ties.)

(3) The victim was an enemy of the defendant, though not the defendant's employer. (Interpret the word "enemy" loosely; if, for instance, the victim and defendant vied for the affections of the same woman, if the victim had harassed one of the defendant's loved ones, if there was a feud of some sort that turned violent, assume enmity existed. If, however, the victim could be viewed as the defendant's employer--whether as (say) his supervisor in a factory or the person who hired him to perform some chores--do not give a score of 0 under (3).)

(4) The victim, although a stranger to the defendant, acted in a highly provocative manner just prior to the killing (e.g. racial taunts).

(5) The victim was engaged in an illegal or often-disapproved activity at the time of the killing (e.g., a drug dealer, a prostitute or prostitute's customer, owner of a homosexual bathhouse, etc.).

If the case does not warrant the rating '0', give it the score '1'. '1' is the appropriate rating for most stranger-to-stranger killings and those in which the defendant only knew the victim in the latter's official capacity (e.g. as employer, or attendant in a local gas station). If there are several victims, give the case a '0' if any of those slain qualify for it.

III. THE "HEINOUSNESS" OF THE MURDER

There are two aspects to this dimension: the question whether self-defense motivated the killing and how "gruesome" it was.

SELF-DEFENSE is an element in the case under any of the following circumstances:

(1) The victim had at hand a deadly weapon at the time of the killing. (Merely having a gun in the store or house does not satisfy (1).)

(2) The victim was killed with his own weapon. (This is taken to imply (1) is satisfied even if the narrative does not explicitly say so.)

> NOTE: If the victim was a police officer, do NOT invoke SELF-DEFENSE did (1) or (2) UNLESS the officer fired shots before the defendant did.

(3) The victim had threatened to kill the defendant or one of the defendant's loved ones.

(4) The victim had attacked the defendant at the time of the killing.

If none of the above conditions existed, self-defense was NOT a mitigating circumstance in the homicide.

> NOTE: If the only evidence for self-defense is the defendant's uncorrobated claim, assume its absence EVEN IF any of (1)-(4) is alleged.

A homicide is classified as VILE if:

(1) it was accompanied by rape, or sexual abuse, either against the victim or someone in the company of the victim

or

(2) there were at least two homicide victims

(3) the deceased was a kidnapping victim at the time he was slain

(4) psychological torture preceded the killing (e.g. Russian roulette, a sustained period of terror)

(5) the victim was shot several times in the head at close range

(6) the killing was execution-style (i.e. victim forced to kneel or squat, then shot in head)

(7) the death was caused by strangulation, or arson

(8) the death was caused by a drowning in which physical force kept the victim below water

(9) the killing involved ten (10) or more shots or stab wounds, except when the murder wapon was a penknife or other small cutting instrument

(10) the physical details of the killing are unusually repulsive (e.g. the victim drowned in his own blood)

(11) the body was mutilated, or otherwise grossly disfigured (except in an attempt to conceal the homicide)

(12) the killing was performed with a bizarre weapon (e.g., a hacksaw, a hammer, an icepick)

(13) The defendant apparently derived pleasure from the very act of killing. (This is distinct from his believing the victim deserved to die, and taking pleasure on that account.)

(14) The crime was specifically described as extremely bloody in the narrative.

Absent all these circumstances, the homicide is categorized as NOT VILE.

Despite the length of the list above, most "simple" shootings, stabbings, and beatings would not be classified as vile under these rules.

APPENDIX B: FOUR ILLUSTRATIVE CLASSIFICATIONS

Example 1

The defendant (a 27 year old female) fatally shot her husband. The defendant was angry that the victim was messing around with another woman. The defendant asked a friend for a ride for the purpose of "locating and cutting up" the other woman. The friend refused and took her home. After her return home, the victim was leaving the house when a shot was heard. Witnesses saw the victim fall to the ground and the defendant emerged from the house carrying a gun. Witnesses testify that she said, "I told him if I ever caught him messing around with another woman, I'd kill him." Defendant told police, "I don't know why I shot him, I just did." Victim had a history of beating his wife.

Ratings:

"Deliberateness"   1

"Status of Victim"   0

"Heinousness":   Self-defense   NO
                                          } Thus heinousness rating was 1.
                   Vileness   NO

Explanation of Ratings:

Deliberateness: No basis for either a zero or a two, a circumstance
                   that dictates a score of one.

Status of Victim: The victim was the defendant's spouse.

Heinousness: There is no serious evidence that self-defense was a factor;
                   while the victim "had a history of beating his wife," he was
                   leaving the house at the instant she shot him. None of the
                   criteria for "vileness" are satisfied in this case.

A killing like this one, incidentally, almost never leads to the death penalty. (Nor did it in this particular case.)

Example 2

Defendant and co-perpetrator (both teenage males) lived close to victim (a 55 year old female) and had conspired to rob her for some time. On the night of the murder, the co-perpetrator entered the house first and then forced the victim to let the defendant in. There was evidence that they raped the victim and to stop her from screaming, the co-perpetrator repeatedly stabbed her with the defendant's knife. The defendant made three different statements to the police following the crime showing varying degrees of involvement in the murder.

Defendant was found guilty of rape and murder.

Ratings:

"Deliberateness"   1

"Status of Victim"   1

"Heinousness":   Self-defense   NO
                                        } The heinousness rating was 2.
                  Vileness   YES

Explanation of Ratings:

Deliberateness: No clear evidence of a prior intent to kill (as opposed to rob). While "there was evidence" that the co-perpetrator stabbed the victim, it was not overwhelming: the knife was owned by the defendant and his statements to the police about his guilt were contradictory. As noted earlier, uncertainty whether the defendant was the killer is a weaker condition than clear indications that he was not; it is the latter circumstance that justifies a '0' on deliberateness.

Status of Victim: Clearly a '1'.

Heinousness: No self-defense; the rape in itself makes the killing vile.

This defendant was sentenced to death.

Example 3

Defendant was a 23 year old military man. Victim was a male.

Apparently, defendant got drunk with victim's nephew before the offense and the nephew told defendant that the victim had a large amount of money in his house. It is not clear if this man specifically encouraged defendant. Defendant, while still intoxicated, went into the house and was intent on robbing the victim. Defendant claims that victim came out from his bedroom and fired a shotgun in the direction of the defendant. Defendant returned the fire with a pistol and killed the victim. Defendant then fled with some money he found. When defendant was questioned by police, he made a full confession, gave them the money and led them to where he had hidden the victim's shotgun.

Ratings:

"Deliberateness"   1

"Status of Victim"   1

"Heinousness":   Self-Defense   NO
⎫
⎬ Thus heinousness rating was 1.
⎭
                      Vileness   NO

Explanation of Ratings:

Deliberateness: The killing satisfies none of the conditions that would warrant a '0' or a '2'. (Premeditation to rob is not equivalent to premeditation to kill.)

Status of Victim: While the defendant knew the victim's nephew, he apparently did not know the victim.

Heinousness: The ratings are appropriate because: (i) under the stated criteria, the killing is not especially vile, and (ii) there is no compelling evidence of self-defense, only the defendant's uncorroborated story. (Had a bullet from the victim's gun been found (say) in the wall, the self-defense claim would be more credible and the rating on that subject changed to YES.) The fact that the defendant left the house with the victim's shotgun does not prove the victim actually tried to use it.

This defendant was sentenced to death. The author suspects that, had the self-defense claim been stronger, the outcome of the trial would have been a prison term.

Example 4

    Defendant (a 21 year-old male) was the next door neighbor of
the victim (a 49 year-old male). The defendant's girlfriend went
to the victim's house and argued with his wife over a cake plate.
Defendant went over and argued with victim and his wife also. Defendant
claimed that the victim opened up a knife and threatened him. Defendant
went home and later the victim went to his porch with the knife.
Defendant shot three times. The third shot was fatal.

Ratings:

"Deliberateness"   <u>1</u>

"Status of Victim"   <u>0</u>

"Heinousness: Self-Defense   <u>YES</u> } Thus heinousness rating was <u>0</u>.

                     Vileness   <u>NO</u> }

Explanations of Ratings:

Deliberateness: No reason for a zero or a two; thus a one.

Status of Victim: The victim and defendant clearly were acquainted;
                    the notion of "dispute turned lethal" fits the
                    spirit of II-(3) in Appendix A.

Heinousness: The killing was not vile. But self-defense apparently is
                    an element, given that the victim arrived at the defendant's
                    house with knife in hand.

    The defendant in this case was sentenced to prison.

APPENDIX C: SOME TESTS OF THE CLASSIFICATION MODEL

We have presented the full text of the classification rules (Appendix A, augmented in Section V) and some data about their explanatory power. But there are still some questions including:

(1) Is the scheme "objective" in the sense that two different people, applying the scoring rules in a given case, would probably reach the same ratings?

(2) Several potential important factors are not considered in the rating procedure (e.g. age of victim). Is there empirical evidence supportive of such exclusions, and other kindred simplifications?

(3) We have evidence that the scheme is fairly powerful within the set of cases that led to its formulation. But how well would it <u>predict</u> the sentencing decisions in other Georgia cases from the same period?

Below we describe some tests we performed related to each of these questions.

Objectivity

"Objectivity" was explored in an experiment conducted by Mary Elsner of the National Center for State Courts. She asked student volunteers from William and Mary Law School to read Appendices A and B, and then to consider the narrative summaries of a series of Georgia murder cases. Based on these narratives (from which all information about the verdict was deleted), the students were asked to classify the cases under the scoring rules.

The experiment took place in two parts. The first phase was a "trial run" (as it were) in which the ratings of the students and the author were contrasted, the aim being to root out cryptic features of the classification scheme. Several minor changes were made, but they all involved elaborating the rules rather than altering them.[56]

---

56. The version of the procedure in Appendix A is the final version.

In the test phase, 9 new students were given 25 Georgia cases to classify,
none of which had been used in the "clarification phase" of the experiment. [57]
For every case, the students provided ratings about deliberateness, victim-status,
vileness, and self-defense. The total number of ratings that emerged was therefore
25x4x9 = 900.

Over the 900 scores, the students and the author agreed 87% of the time.
The 13% disagreement rate could have three possible origins: (a) the scheme might
contain flaws that provoke occasional discrepancy; (b) some case narratives are
genuinely ambiguous on key points; (c) some student ratings are clearly inconsistent
with the classification rules. All of these problems probably were present to some
extent.

When obvious student errors are excluded (e.g. the case was classified
"not vile" although there were two murder victims), the agreement rate rises
above 90%. And when we set aside rating "disputes" tied to subtle semantic
distinctions (e.g. does an "ambush" in a robbery killing imply prior intent
to kill, as opposed to prior intent to surprise and rob?), the rate of agreement
reaches 93%.

Of course the scheme is not perfectly objective: could there be any
universal definition of (say) psychological torture? But the degree of consensus
that did arise in the experiment strikes the author as surprising and gratifying.

Simplifying Assumptions

Other tests that were performed pertained to the internal structure of
the classification procedure. Unlike Gross and Mauro, for example, we assumed
that the sex of the victim was not important per se. How can we defend such a

---

57. The cases were randomly chosen as follows: 100 cases had been selected at
    random by the National for State Courts to illustrate the use of a certain
    computer program. The Center had numbered the cases, and we simply chose every
    fourth one for our sample.

hypothesis, and various others like it that are implicit in the scoring rules?

The hypotheses in question are of the general form: characteristic X is irrelevant to the sentencing decision. To test any one of them, we might divide the cases in a given class into those having characteristic X and those that do not. If the death rates in these two subdivisions do not differ significantly, the axiom about X is sustained (at least in that class). The word "sustained" does not imply that the hypothesis has been decisively verified; rather we have avoided the self-contradictory situation in which, having proceeded on the assumption that X is not important, we wind up producing clear evidence that it is.

We performed several such hypothesis tests, three of which are presented below. The results were consistent with the assumption under scrutiny although, - as the numbers below will suggest, we sometimes had to proceed with rather small data samples.

> Hypothesis 1: In death-sentencing decisions, the
> age and sex of the victim do not matter.

This contention goes against the notion that juries are especially harsh on those who kill women, the very old, or the very young. Killing such "helpless victims", it is sometimes suggested, is more abhorrent than slaying those who might somehow have "fought back".

Certainly the victim's age and sex do not matter when $s \leq 2$, for virtually no kilings in that range lead to death sentences. After dichotomizing age into the two ranges 15-60 and "0-14 plus 61 and up", we reach the following contingency tables for Category 3 and class (1,1,2):

DEATH SENTENCING RATES BY AGE AND SEX OF VICTIM

	CATEGORY 3 KILLINGS		CLASS (1,1,2)	
	MALE	FEMALE	MALE	FEMALE
Age 15-60:	.22(64)	.17(23)	.89(18)	.79(14)
Other Age:	.38(8)	.33(9)	.89(9)	.82(11)

Neither of these tables depicts statistically significant variation [58] around the relevant average death risk (23% in Category 3; 85% in Class (1,1,2)).

There is no evidence that the killers of women are more "death-prone" than the killers of men.[59] Those who murdered the elderly and young did fare worse than other Category-3 killers; note, however, the small sample size for "outer-age" victims. That this age-related discrepancy might be a fluctuation is suggested by the absence of a similar pattern in Class (1,1,2).[60]

> Hypothesis 2: The death risk associated with killing an employee in his official capacity is the same regardless of his occupation.

This hypothesis asserts that, despite statutory distinctions, killing a policeman is not more likely to elicit a death sentence than slaying a cab driver, gas station attendant, or liquor store cashier. Perhaps the "contemporaneous offense" aspect of the latter murders (which usually involve robbery) is given about equal weight to the attack on society itself when a law enforcement officer is slain.

---

58. Even when one considers the prior record of the defendant and the setting of the trial (urban/rural), this outcome is unchanged. (By "not statistically significant," we mean that usual sampling error could account for the observed differences.)

59. The victim's sex could appear important because of a spurious correlation: sexual abuse -- which evokes revulsion whether the victim is male or female -- is disproportionally frequent in female-victim cases.

60. It is also possible that age is a factor only in the "borderline" Category 3 cases. But the data, though consistent with this possibility, hardly confirm it.

An appropriate place to test Hypothesis 2 is the (1,1,1) class, which includes most employee killings that are not vile. Within that class, there were seven killings of peace officers (policemen, state troopers, detectives, security guards); two of them (29%) led to death sentences. There were 22 killings of other kinds of workers; these brought forth six death verdicts for a rate of 27%. And, interestingly, the death-sentencing rate for the other 19 (1,1,1) killings is 26%. One can hardly question the homogeneity of the grouping on the basis of these numbers.

> Hypothesis 3: In killings clearly plotted in advance, the death risk is the same for the instigator and the executioner.

The classes (2,0,1) and (2,0,2) lend themselves to assessing this hypothesis. In (2,0,1), there were two cases in which the defendant instigated the murder but did not take part in it; neither ended in a death verdict. In (2,0,2), three cases involved instigators with no physical connection to the slaying; one of these defendants was sentenced to death.

Given the aggregate death rates in (2,0,1) and (2,0,2), one would have "expected" 1.9 death sentences among the five instigators studied. They collectively received only one but, given the sparsity of data, to attribute significance to this deficit might be farfetched. A more complex model that distinguishes instigators, executioners, and those active in both phases of the killing might be useful, but the data do not establish the need for it.

In all, the data did not force us to abandon any of the simplifying assumptions used in constructing the classification scheme.

Predictive Power

After devising the classification rules from study of more than 400 cases, we considered a "holdout sample" of 100 or more. These new cases were classified under the scoring rules (which they played no role in developing), and their dispositions were noted.

The sentencing patterns in the new data were similar to those in the original set. Of the nine additional (1,1,2) cases, eight led to death verdicts, a fraction that closely follows the proportion in the earlier cases. Of the 62 new cases with s-values between 0 and 2, only one yielded a death sentence. And, at a rate nearly indistinguishable from the earlier figure, 22% of the Category 3 killings brought death sentences to their perpetrators.

Being so similar to the original set, the 100 new cases were combined with it for the Table 1 calculations. But the test involving them helped confirm that the classification rules reflect general sentencing practices, and not just the idiosyncracies of a limited number of cases.

A Final Note

The 513 cases studied comprise about 7/8 of those in the post-Furman Baldus File. We excluded all cases with handwritten narrative summaries because a substantial proportion of them (written by the same person) were not fully legible. We also discarded some cases whose summaries, although legible, were cryptic or otherwise disheartening (e.g. a narrative that began "Facts very confusing".). A few other records were set aside because of inconsistencies between the computer coded variables and the narratives (e.g. on whether the case was pre- or post-Furman).

We have no reason to believe that these exclusions biased the analyses that were performed. They seem quite inconsequential compared to the circumstance that, because of the generosity of Professor Baldus, we were able to work with a huge and excellent data set.