

Summer 1977

A Survey of Judges' Responses to Simulated Legal Cases: Research Note on Sentencing Disparity

William Austin

Thomas A. III Williams

Follow this and additional works at: <https://scholarlycommons.law.northwestern.edu/jclc>

 Part of the [Criminal Law Commons](#), [Criminology Commons](#), and the [Criminology and Criminal Justice Commons](#)

Recommended Citation

William Austin, Thomas A. III Williams, A Survey of Judges' Responses to Simulated Legal Cases: Research Note on Sentencing Disparity, 68 J. Crim. L. & Criminology 306 (1977)

This Criminology is brought to you for free and open access by Northwestern University School of Law Scholarly Commons. It has been accepted for inclusion in Journal of Criminal Law and Criminology by an authorized editor of Northwestern University School of Law Scholarly Commons.

A SURVEY OF JUDGES' RESPONSES TO SIMULATED LEGAL CASES: RESEARCH NOTE ON SENTENCING DISPARITY

WILLIAM AUSTIN* AND THOMAS A. WILLIAMS, III**

I. INTRODUCTION

A common assumption made by observers and researchers of the American judicial system is that sentencing disparity (discrepant sentences assigned for similar offenses and similar offenders) is all pervasive. For example, this is the inference one may draw from the report on the courts by The President's Commission on Law Enforcement.¹ Moreover, the belief that disparate sentences have produced gross injustices for convicted offenders and have undermined the deterrent effect of the law has prompted several proposals for changing federal sentencing procedures.² However, Hogarth prefaces his exhaustive work on judicial sentencing with a cautionary note regarding sentencing disparity:

The extent of disparity in sentencing may be more apparent than real. Within each offense category there may be many different combinations of facts related to the offender, the offense, and the surrounding circumstances which could properly affect the selection of a penalty. Without adequate statistical control over the types of cases appearing before the courts, it would be wrong to assume that there is genuine lack of uniformity in sentencing.³

This statement is not meant to imply that sentencing disparity does not exist. Rather, it is more accurately interpreted to mean that

*William Austin, Assistant Professor of Psychology, University of Virginia; Ph.D., University of Wisconsin-Madison, 1974 (Sociology and Social Psychology); B. A., Western Michigan University, 1969.

**Thomas A. Williams, III., Assistant to Adult and Juvenile Corrections Specialist, Commonwealth of Virginia Division of Justice and Crime Prevention; B. A., University of Virginia, 1976.

¹ THE PRESIDENT'S COMMISSION ON LAW ENFORCEMENT AND ADMINISTRATION OF JUSTICE, TASK FORCE REPORT: THE COURTS 23 (1967).

² See Advisory Council of Judges of the National Council on Crime and Delinquency, *Model Sentencing Act*, 9 CRIME & DELINQUENCY 339 (1963); Address by U.S. Attorney General Edward Levi, N.Y. Times, Feb. 3, 1976, at 13, col. 1; S. 1, 94th Cong., 1st Sess. (1975).

³ J. HOGARTH, SENTENCING AS A HUMAN PROCESS 7 (1971).

due to the existence of uncontrolled variance within offense categories, researchers have not yet been able to determine the *precise* degree of sentencing disparity, even when many factors are statistically controlled.⁴ In addition, most research efforts in this area have been relatively uninformative as to what extent disparities indicate differences in judicial attitudes (prejudices, views on the goals of legal sanction, methods of categorizing offenders) and to what extent they reflect "legitimate" differences in the nature of the evidence (mitigating and incriminating factors).

This criticism seems particularly true of several "classic studies" which have probably heavily influenced public views that widespread disparity exists.⁵ All of these researchers have concentrated on the effect of defendant characteristics and some⁶ have drawn inferences about the role of judicial attitudes in sentencing. Although these studies have contributed to our awareness of the potential impact of demographic and personal variables in sentencing, they uniformly contain methodological weaknesses due either to lack of statistical controls⁷ or to failure to collect data on judicial attitudes.⁸ Two reviews of the impact of defendant characteristics⁹ make this point by sug-

⁴ For extensive discussions of previous research on sentencing disparity, see *id.* at 7-12; R. HOOD & R. SPARKS, KEY ISSUES IN CRIMINOLOGY 141-54 (1970).

⁵ See E. GREEN, JUDICIAL ATTITUDES IN SENTENCING 100-02 (1961); Bullock, *Significance of the Racial Factor in the Length of Prison Sentences*, 52 J. CRIM. L.C. & P.S. 411 (1961); Gaudet, *The Sentencing Behavior of the Judge*, in ENCYCLOPEDIA OF CRIMINOLOGY 449 (V. Brandom & S. Katash eds. 1949); McGuire & Holtzoff, *The Problem of Sentencing in the Criminal Law*, 20 B.U.L. REV. 413, 426-33 (1940); Sellin, *Race Prejudice in the Administration of Justice*, 41 AM. J. SOC. 212 (1935).

⁶ See E. GREEN, *supra* note 5.

⁷ See Bullock, *supra* note 5.

⁸ See E. GREEN, *supra* note 5.

⁹ See Chiricos & Waldo, *Socioeconomic Status and Criminal Sentencing: An Empirical Assessment of a Conflict Proposition*, 40 AM. SOC. REV. 753 (1975); Hagan, *Extra-Legal Attributes and Criminal Sentencing: An Assessment of a Sociological Viewpoint*, 8 L. & SOC'Y REV. 357 (1974).

gesting that the earlier studies have substantially overestimated the importance of these variables. More recent investigations¹⁰ are more sophisticated in pinpointing the relative influence of judicial attitudes and defendant characteristics. But these studies still face the methodological obstacle of *equalizing cases*; hence, estimating the degree of sentencing disparity.

The purpose of this investigation is to gain a relatively "pure" estimate of the degree of sentencing disparity among the same type of judges within the same state jurisdiction. The research strategy facilitated this task: an assembled group of state district court judges responded to the same hypothetical legal cases, thus eliminating the cumbersome problem of equating different legal cases within offense categories. Now, of course, the gain in methodological parsimony is a loss in "realism." One must proceed on the assumption that the simulated data possess a high degree of ecological validity, or "generalizability." Even though the judges were not *responsible* for actual judicial outcomes, it is reasonable to assume that the judges employed decision rules similar, if not identical, to the ones they use in actual trial proceedings.

II. PROCEDURE

A. SUBJECTS AND INSTRUCTIONS

Data were gathered from forty-seven Virginia district court judges attending a state judicial conference in October, 1975.¹¹ Since the goal of this study was to estimate the degree of sentencing disparity by eliminating variance in similar legal cases, the judges were each given an identical booklet containing descriptions of five legal cases. This booklet, distributed to the judges in a conference auditorium by the organizers of the judicial conference, bore a cover page identifying the sponsors as the "Joint Executive Committee and the Education Committee" and expressly stating that the study was a survey of comparative judgments designed to assess the rate of agreement among judges. The subjects were asked to read carefully the instructions and case descriptions, to recommend a verdict, and, if a guilty verdict was rendered, to suggest an appropriate sen-

tence based on the law of the Commonwealth of Virginia. Upon completion, the surveys were returned to the conference coordinator.

B. DESCRIPTION OF CASES¹²

Each of the case descriptions conveyed the basic evidential factors requisite for a verdict: defendant's name, the criminal charge and a synopsis of the testimony. Nonevidentiary information was also supplied. All of the cases described minor felony or misdemeanor offenses—the types of cases assigned to Virginia district courts.

Case 1 concerned an eighteen year old female defendant who was apprehended for *possession of marijuana*. She was arrested with her boyfriend and seven other acquaintances. Evidence of a substantial amount of smoked and unsmoked marijuana was found; however, no marijuana was discovered directly on the defendant's person. She had no previous criminal record, was a good student from a middle class home and was neither rebellious nor apologetic for her actions.

Case 2 involved a charge of *reckless driving* against two teenage boys who were arrested for "drag racing." After the police officer's signal to stop, one boy pulled his car over immediately while the other continued to accelerate for one-half mile. The first boy had no prior record and claimed he was not racing but was hurrying home with his father's borrowed car. The second boy had two previous traffic convictions which included speeding.

Case 3 concerned the arrest of a woman accused of *shoplifting*. An off-duty police officer witnessed the theft of a pair of gloves, a bottle of perfume and some handkerchiefs—a value of \$23.95—and arrested the defendant in the presence of a security officer. All items were new and no sales receipts were produced. The defendant had one prior conviction for shoplifting; her husband was a bank vice-president; she was a religious person and remorseful because of her predicament.

Case 4 concerned a *burglary*. Residents of a household returned home to find their television set missing, and, hearing a noise, noticed two persons carrying a large object down the street. Soon afterwards, officers arrested a forty-six year old man and his son, who were

¹⁰ See J. HOGARTH, *supra* note 3; R. HOOD, SENTENCING IN MAGISTRATES' COURTS (1962).

¹¹ Virginia Judicial Conference, American Academy of Judicial Education (Oct. 1975).

¹² A more complete description of the cases is available from Professor William Austin, Department of Psychology, University of Virginia, Charlottesville, Va. 22901.

pawning the TV set, which was identified by its serial numbers. The father, whose fingerprints matched those found in the house, had four previous convictions for petty larceny, one for reckless driving and one for assaulting a police officer. The son had only one traffic conviction and claimed his father had told him a friend gave him the television set.

Case 5 involved an arrest for *drunken driving*. A male defendant was arrested after demonstrating blatant signs of erratic driving. A "balloon test" revealed the driver was intoxicated. The defense attorney pleaded that the defendant had a serious drinking problem.

III. RESULTS

The overall pattern of judges' recommendations for the five cases was a generally high rate of agreement on the verdict, but substantial variance in the choice of sentencing mode and the magnitude of penalty within some modes.

Case 1

The results for the case of marijuana possession were quite dramatic. Eighteen judges (38.3%) voted "guilty," while twenty-nine (61.7%) "not guilty" verdicts were recorded. This was the lowest rate of verdict agreement of the five cases. This may be the result of the lack of consensus on marijuana laws or it may reflect the strength of the evidence—marijuana was not found directly in the possession of the defendant.

Among the judges who voted "guilty," sentencing responses were quite varied. Eight judges (44.4%) recommended probation; four (22.2%) would impose a fine; three (16.7%) would issue probation *and* a fine; and three judges (16.7%) were in favor of a jail term.

Case 2

This case required judges to consider two defendants on the charge of reckless driving. Judges showed a high rate of agreement on the verdict for both boys: all forty-seven judges voted guilty for the second defendant, and thirty-three (70.2%) voted guilty for the first defendant.

The sentencing data paralleled the verdicts. For defendant two, judges clearly favored suspension of driver's license *and* a fine (thirty-six judges or 76.6%); six (12.8%) opted for a fine

alone; three (6.4%) for suspension of license; and two (4.2%) for fine and a jail term. However, judges were in less agreement for defendant one. Sixteen judges (48.4%) favored a fine; eleven (33.3%) a fine and suspension of license; four (12.1%) a fine and traffic school; and two (6.1%) suspension of license. There were also large variations in the *magnitude* of sentences within sentencing modes. The range in fines recommended was \$25 to \$500 for defendant two, and \$10 to \$500 for defendant one.

Since this is not an experimental study, it is impossible to determine *why* more consensus was registered for defendant two, but the data strongly suggest that the strength of the evidence affects disparity in verdicts and sentences. For the two defendants, more consensus on a verdict resulted in less sentencing disparity. This interpretation is supported by the responses of ten judges who wrote on their questionnaires that defendant one was guilty of the lesser charge of "speeding" or "improper driving," rather than "reckless driving."

Case 3

In this shoplifting case with eyewitness testimony, once again agreement on the verdict was high: forty-one judges (87.2%) voted guilty. However, even though the defendant had a previous conviction, there was great disparity on the choice of sentence. Nine different "types" of sentences were recommended. Twelve judges (29.3%) favored a fine and a suspended jail term; seven (17.1%) favored suspended jail term alone; seven (17.1%) favored only jail but suspended part of the term—ranging from ten days to ninety days actual term served; five (12.2%) favored a fine plus a jail term—ranging from \$25/10 days to \$150/60 days; and the remaining eight judges voted for five other modes of sentences.

Case 4

This case involved a burglary charge against a father and his son. The final disposition of this type of case does not lie with Virginia district courts; rather, district court judges hear preliminary evidence and decide whether to bind the case over to a grand jury. Thus, many judges, true to their role, made only this type of recommendation when they thought the defendants were guilty and did not recommend a sentence.

The judges recorded a high agreement for the verdict and sentence for the father. The evidence was stronger against the father and he had a prior criminal record. All of the judges voted guilty (31) or to certify to the grand jury (16). Of the thirty-one judges who determined a sentence, twenty-six (83.9%) recommended either a jail term (13) or a jail term and a fine (13). However, these sentences ranged from thirty days and \$100 to five years at a state prison farm. The remaining five judges (16.1%) recommended jail or jail and fine with part of the jail term suspended.

More disparity was recorded for the son who had pleaded that he was unaware that he was participating in a crime. Twenty-seven judges (57.4%) voted guilty and twenty (42.6%) not guilty. Of the judges who recommended a guilty decision, eight voted to certify to the grand jury and eight felt the defendant was guilty of the lesser charge of "larceny." The sentencing for the son was also more disparate. Five (27.7%) of eighteen judges recommended a jail term; seven (38.9%) favored jail with part of the term suspended; three (16.6%) voted to suspend the entire jail term; two favored jail and a fine; and one judge decreed a fine and jail with part of the term suspended.

Thus, once again, sentencing disparity seems closely related to the persuasiveness of the evidence, with judges differing in their threshold of "how much evidence is enough." Judges were almost unanimous in the disposition of the father, but even when agreeing on the *type* of sentence, they differed greatly on the *magnitude* of punishment.

Case 5

This was a drunken driving case with eyewitness testimony and positive results from a "balloon test." Agreement on a guilty verdict was nearly unanimous—forty-five judges (95.7%)—but choice of sentence revealed a complex pattern. There was clearly a dominant choice with twenty-five judges (55.6%) choosing a fine and suspension of operator's license. However, seven other types of sentences were specified, the most popular choices being fine accompanied by suspension of license and suspended jail term (eight judges or 17.7%), and assignment to an alcoholism program (five judges or 11.1%).

The results in this case were quite different from the others in that there was little variation in the magnitude of sentence within sentencing mode. For example, within the dominant choice (fine, suspended license) almost all of the twenty-five judges recommended \$200 or \$250 as the appropriate fine. This indicates that for certain types of offenses there are relatively standard penalties that work to reduce sentencing disparity. Yet, even in this case where a dominant sentencing mode and magnitude of penalty was manifest, there was disparity in choice of sentencing mode (twenty of forty-five judges opted for a different choice).

IV. SUMMARY

From the degree of disparity in the *verdict*, in choice of *sentencing mode*, and in *magnitude of penalty* within sentencing mode among the judges' responses to five legal cases, a number of conclusions can be drawn.

(1) At the most general level, it is clear that when legal cases are equalized within offense categories, judges still show substantial disparity on all three criteria.

(2) The strength of the evidence against a defendant is related to sentencing disparity. In those cases where the evidence was weakest (*i.e.*, Case 1; Case 2-defendant one; and Case 4-defendant two), disparity in mode and magnitude of sentence was great.

(3) However, strong evidence does not appear to guarantee a high agreement on the appropriate sentence. For Case 2-defendant two and Case 4-defendant one, strong evidence (high verdict agreement) "produced" high agreement on mode of sentence, but disparity in magnitude; in Case 2 strong evidence was accompanied by disparity in mode *and* magnitude; and in Case 5, strong evidence was correlated with a dominant mode and low variance in magnitude within this mode, but almost half of the judges chose another type of sentence. *Thus, the five cases produced a variety of patterns of disparity, but some form of disparity was always present.*

(4) The type of offense *appears* to be related to the degree of disparity. The most disparity was found in the drug case and the least in the drunken-driving case. It seems plausible that consensus existed among the district judges on the appropriate response to drunken driving,

but not on marijuana possession and shoplifting.

(5) It is impossible to give an exact quantitative index of the degree of sentencing disparity because of the qualitative nature of sentencing. Disparity is reflected not merely in the magnitude of the punishment, but in the type of sentence as well. However, in our data the rate of agreement on verdicts varied from 38.3% guilty verdicts to 100%; preference for a single mode of sentencing varied from 29.3% to 83.9%; and the magnitude of sentences within modes (of those that could be quantified) was absolutely large for the minor offenses studied. These figures, though roughly calibrated, are evidence of "substantial" disparity.

(6) Differences in the type and magnitude of sentences *suggest* that judges operate with a variety of "theories of legal sanction" or "decision-making guides." For example, judges favoring probation or assignment to a counseling program probably place more value on the goal of rehabilitation; those opting for heavy

finer or jail terms are more influenced by considerations of deterrence or the moral function of legal sanction (*e.g.*, establishing justice, gaining retribution, upholding community values).

Our data provide a clearer picture of the degree of disparity *for minor offenses* and lead us to speculate as to some of the causes. The results of this study should be integrated with other research efforts in this area,¹³ which, when taken together, should begin to present coherent explanations for the existence of disparity in sentencing. Once researchers go beyond the present level of knowledge by utilizing a combination of research strategies, the necessary data will be available to support, or detract from, the plausibility of numerous legislative proposals¹⁴ designed to reduce sentencing disparity.

¹³ See W. GAYLIN, *PARTIAL JUSTICE: A STUDY OF BIAS IN SENTENCING* (1974); J. HOGARTH, *supra* note 3; R. HOOD, *supra* note 10.

¹⁴ See note 2 *supra*.