DOES IT PAY TO PLEAD GUILTY? DIFFERENTIAL SENTENCING AND THE FUNCTIONING OF CRIMINAL COURTS

DAVID BRERETON*
JONATHAN D. CASPER

Do defendants who plead guilty receive lighter sentences than those with similar charges and attributes who exercise their right to trial? The assertion that they do has long been at the heart of the literature describing and explaining the plea-bargaining process, though it has been questioned in some important work published recently. The existence of sentence differentials is particularly hard to document statistically, because a successfully operating policy of punishing those who go to trial will in fact minimize the number of cases in which the sanction for trial has to be imposed. Examination of data from three California counties, as well as consideration of various theoretical concerns, leads us to argue that sentence differentials *are* likely to characterize jurisdictions whose disposition patterns are based on inducing most defendants to plead guilty.

I. INTRODUCTION

If there is any proposition at the heart of the common wisdom about criminal courts, it is the assertion that defendants who plead guilty are likely to receive less harsh sentences than defendants with similar characteristics and charges who are convicted after trial. It is this sentence differential (whether conceived of as a reward to guilty pleaders or as a punishment of those who waste the court's time by "needless" trials) which has traditionally been seen as the engine driving the plea-bargaining assembly line.

This is a time, however, when much of the received wisdom about criminal courts is being called into question (see, generally, Church, 1979; Casper, 1979). The commonplace notion that plea bargaining is the product of caseload pressure has been criticized both on historical and theoretical grounds

* We wish to thank Milton Heumann and James Eisenstein for their useful comments on a previous draft. This research was supported, in part, by Grant 79-NI-AX-0042 from the National Institute of Justice. All responsibility for the arguments and conclusions rests with the authors.

LAW & SOCIETY REVIEW, Volume 16, Number 1 (1981-82)

(see, e.g., Heumann, 1975; Feeley, 1978; Friedman, 1979). The notion that bargains themselves are largely the product of the necessity to get cases over with has, likewise, been attacked by scholars who assert that concerns for substantive justice and appropriate sentencing are central to the plea-negotiation process (Heumann, 1978; Feeley, 1979b; Utz, 1978; Mather, 1979).

Like many other evident "truths" about the operation of criminal courts, the notion that guilty-pleaders receive lighter sentences than similarly situated defendants who exercise their right to trial has also been viewed with increased skepticism in recent research. In his introduction to this *Review's* special issue on plea bargaining (Vol. 13, no. 2), Malcolm Feeley suggested that we simply did not have enough data to support the argument that it pays to plead guilty. According to his analysis, "(t)here are far too few trials in relation to the number of guilty pleas to allow much confidence in a comparison of the sentences that follow each process. Even if cases were randomly assigned this ratio of 1 to 9 would make inferences extremely hazardous. And actually the situation is far more complex because of the host of factors that channel cases into one or the other alternative" (Feeley, 1979b: 202). In addition, some quite careful recent empirical work has argued that sentence differentials do not exist, at least in the jurisdictions studied. For instance, in an important study of criminal courts in three cities, Eisenstein and Jacob concluded that, once other factors were taken into account, ". . . the effect of dispositional mode [was] insignificant in accounting for the variance in sentence length" (Eisenstein and Jacob, 1977: 270). The INSLAW study of plea bargaining in the District of Columbia reached a similar conclusion: "Contrary to expectations, sentence concessions were not routinely awarded to suspects entering guilty pleas . . . . The plea process reduces criminal behavior, largely by increasing the number of convictions without offsetting losses resulting from more lenient plea bargain sentences" (Rhodes, 1978).

From our reading of the literature, the spirit of skepticism and revisionism in recent writing on criminal courts has by no means relegated the notion of a sentence differential to the dustbin of discarded ideas. A recent article in this *Review*, for example, pointed to the existence of substantial sentence differentials (Uhlman and Walker, 1980; see also Uhlman and Walker, 1979). Indeed, we believe that most participants in and analysts of criminal court processes probably accept the old common wisdom that guilty-pleaders are rewarded.

The data and theoretical arguments presented here support this common wisdom. In presenting "just another" case study confirming the existence of sentence differentials, we hope to accomplish two tasks. The first is to present evidence for the existence of such differentials in the three court systems we have been studying and to discuss issues of generalizability from these to other courts. The second is to caution against the recent skepticism about the existence of such differentials. We offer theoretical arguments as to why we ought to expect sentence differentials in most court systems. We believe that the recent studies suggesting that differentials may not exist should be viewed with great care.

II. THE EVIDENCE TO DATE

Little research has been conducted on sentence differentials, perhaps because, until recently, everyone was so sure that they existed. However, what has been done can be quite neatly divided into two categories: research which has relied primarily on the perceptions of participants; and research which has focused on statistical analysis of various "objective" measures of the determinants of sentences. Before moving to an analysis of our own data, we shall briefly consider the significance of the findings generated by each of these approaches.

Perceptions of Participants

One way in which to determine whether or not a differential exists is simply to ask various actors in the system. Judges, for instance, can be asked if they tend to treat guilty pleaders more lightly than trial cases. Prosecutors can be questioned about their bargaining procedures, and defendants can be asked if they were told they would be better off if they did not go to trial. Given the difficulty of working with more "objective" measures, it is not surprising that the bulk of the evidence relating to differentials has taken this form. Thus, on judges we have the *Yale Law Journal*'s (1956) survey of Federal District Court judges, and more recently the extensive anecdotal evidence collected by Albert Alschuler (1976). On defendants there is Newman's early study (1956; cf. also Newman, 1966) and Casper's (1972) analysis of defendants' perspectives on the criminal justice system. And on prosecutors and attorneys we again have contributions from Alschuler (1968; 1975), plus Vetri's (1964) survey of chief prosecuting attorneys and Mather's (1973) ethnographic study

of public defenders in Los Angeles (cf. also, generally, Heumann, 1978).

Some of this research is only peripherally concerned with differentials, but collectively these findings strongly suggest that most principal court actors share the view that it pays to plead guilty. There are, however, some reasons why we should view these participant accounts with a degree of caution. Participants do not necessarily possess the best understanding of a process just because they happen to be closest to it. They may be heavily influenced by each other, and what the interviewer is getting may be simply an account of how actors in that process have reconstructed reality, rather than an account of the reality itself. Even in the absence of social pressures of this kind, participants may make systematic errors in drawing inferences from what they observe. This possibility was raised by Eisenstein and Jacob when they suggested that, although empirically false, the impression that a differential existed ". . . was nurtured by the occasional heavy sentence after a jury trial or an exceptionally light sentence after a guilty plea" (1977: 271). Significantly, this argument also dovetails neatly with some recent findings in the field of cognitive psychology. Nisbett and Ross (1980) have reported a variety of evidence indicating that individuals show a good deal of insensitivity to the possibility of sample bias (see also Eisenstein and Jacob, 1977: 297). In this respect it could be argued that court participants tend to exaggerate the costs of going to trial, because they fail to take sufficient notice of the fact that the more serious crimes, and defendants with worse records, are over-represented in the trial category. Similarly, Nisbett and Ross also argue that people are inclined to exaggerate the importance of well-elaborated, graphic, and concrete data about specific instances, vis-a-vis statistically based generalizations (what they call the "vividness" criterion). On this point, our observations indicate that court personnel are indeed much more likely to "prove" the existence of a sentence differential to an outside observer by recounting detailed horror stories of "the guy who refused an offer" than by referring to statistical findings. Assuming that such horror stories do occasionally occur, and assuming further that there is an asymmetry of sorts (i.e., instances of uncharacteristically harsh treatment still outnumber those of uncharacteristically gentle treatment) it is possible to understand the tendency to significantly inflate the size of the observed differential.

On the other hand, there are also good reasons for *not* dismissing the perceptions of participants lightly. These individuals may not think like trained social scientists are supposed to think, but they do have access to data which are in many ways superior to that available to most researchers. (At least this is likely to be true for judges, prosecutors, and defense lawyers, whose contact with the system is continuous.) Thus they tend to know more than outsiders do about special circumstances of the case, characteristics of the judge involved, "what wasn't on the official record but was really important," and so on. Likewise, if the respondents are prosecutors or judges, they are presumably in the best position to know whether their own sentencing and/or charging patterns are influenced by the mode of disposition chosen by the defendant. Perhaps most importantly, such participants are also able to track cases in a way which provides a potentially more rigorous test of the differential sentencing hypothesis than does a correlational approach. If a case goes to trial they often know what a defendant was offered to settle, if an offer was made, and they can compare this with the actual outcome instead of just imputing a differential (as we must do) on the grounds that the defendant was otherwise the same as those who pleaded guilty and received lighter sentences. Court participants might still exaggerate or otherwise misread what they observe, but the errors they make are unlikely to be simplistic ones. Because they work with quite elaborate conceptions of "going rates" for particular crimes and particular classes of defendants, they are likely to exercise intuitively a fair degree of control for other possible determinants of sentence outcomes (cf., e.g., Mather, 1973; 1979).

On balance then, we would argue that although the perceptions of court participants cannot be regarded as sufficient proof that differentials exist, they are well enough grounded to require that we take them seriously. This means, in turn, that a fairly convincing counter case has to be made before we should conclude that differentials are illusory.

Quantitative Approaches

Prima facie, the most appropriate way to check the accuracy of these perceptions is to collect as much "hard" data as possible on the various determinants of sentence and then use this to see if the relationship between sentence and mode of disposition is real or spurious. However, there are both

conceptual and data-related problems in employing such an approach. A large number of factors can affect the sentence a defendant receives, but reliable measures of several of them (e.g., "strength of evidence") can be hard to come by. Court records are often incomplete, and first-hand observation is time-consuming, costly, and subject to difficulties in ensuring coder reliability. Moreover, even if these obstacles can be overcome, the researcher still has to accumulate enough cases to make worthwhile statistical analysis possible—no small problem when, with a handful of exceptions, only between 5 and 20 percent of cases heard by courts go to trial. Finally, even if it is possible to accumulate enough data about enough cases, the data will probably still come from only one or perhaps two jurisdictions, and will thus raise doubts as to how applicable these results are to other parts of the criminal justice system.

The major *conceptual difficulty* entailed in using statistical approaches to determine sentence differentials can best be stated as follows: *when differentials are most effective, they are least observable*. As Heumann (1979) argues, some cases which go to trial are regarded as doing so justifiably because they raise important legal questions, because the defendant's guilt is genuinely in question, and so on. In these cases, it is unlikely that the defendants who choose trials will suffer a penalty, for they are not considered to have wasted the court's time. By contrast, defendants whose cases are regarded as "dead-bang" are much more likely to receive a more severe than usual sentence if they go to trial, either because the court wants to discourage others from doing the same or because, as some judges and prosecutors claim, such "frivolity" deserves to be punished for its own sake. Provided the penalties were large enough and certain enough, we would expect the defendants in most of these "inappropriate" cases to be convinced of the wisdom of pleading guilty, thereby leaving the "appropriate" cases heavily over-represented in the trial category. In the same way that deterrence theory suggests that the best evidence for an effective use of nuclear weapons is the absence of their actual employment, one might expect that an effectively managed sentence differential policy would minimize the number of cases in which defendants would, in fact, actually have to be punished for exercising their right to trial. The somewhat perverse extension of this argument, of course, is that the strongest evidence for differentials may be their statistical nonexistence! Naturally, we do not want to

carry the argument to this extreme, but in a general way the point holds: we need to be very cautious about drawing conclusions about the role of differentials on the basis of tests of statistical strength alone. This does not mean that such approaches are useless—indeed we shall shortly turn to presentation of some data of our own—but it does mean that the issue needs to be carefully framed.

Two recent and influential studies assert that differentials did *not* exist in their jurisdictions. The Eisenstein and Jacob study is carefully designed and appears to attempt to overcome some of the difficulties we have cited above. The study is based on three quite different jurisdictions rather than just one, thus reducing (though not eliminating) doubts about generality. The authors accumulated a comparatively complete body of data for a large number of cases, including information not only on the more obvious predictors of sentence outcome, such as charge and record, but also the defendant's social class and race, the judge or judges who heard the case, and, to some extent, the strength of the evidence presented against the defendant. Although this still leaves out some potentially relevant variables, and although there are measurement problems with some of those which are included, this is probably about as complete a body of data as we can hope to collect, barring some revolution in social science data gathering (and financing) procedures.

Despite the quality of their data, however, Eisenstein and Jacob's argument seems to us less conclusive than it might first appear. Their overall conclusion is that the effects of the mode of disposition on sentence are very small, though stronger for some judges than for others. Their general rejection of the differential sentencing hypothesis is based on the results of multiple regression analysis (where length of sentence is the dependent variable) and multiple discriminant function analysis (where prison/no prison is the dependent variable), which together indicate that the amount of variance explained by mode of disposition is consistently much lower than that accounted for by the original charge or by the characteristics of the defendant. For example:

> [W]hen we take a closer look and consider not only the type of disposition but also the offense on which a person is convicted, his personal characteristics, the strength of the case against him, and the identity of the courtroom workgroup that sentenced him, the effect of dispositional mode is insignificant in accounting for the variance in sentence length. It accounts for as little as 3 percent, and at most 7 percent, of the variance that can be accounted for by all these factors (Eisenstein and Jacob, 1977: 270).

The problem is that their approach mis-states the issue, since to our knowledge no one has ever argued that mode of disposition is the *major* factor in the sentencing decision. After all, common sense alone suggests that the seriousness of the charge, under most circumstances, will be the most important determinant of the sentence received—both because the penalty structure available to the sentencers is contingent upon the conviction charge, and because the amount of harm done is itself an important consideration in sentencing decisions. Similarly, who would expect a defendant with a prior felony to do better than a first-timer charged with the same offense, just because the former pleaded guilty and the latter did not? The appropriate question to ask is not whether mode of disposition is *more or less strongly related* to sentence decisions than are other factors, because the sentence differential hypothesis does not assert this. Rather, the appropriate question to ask is *whether or not mode of disposition makes a difference ceteris paribus*; this assertion encompasses the view that guilty pleaders are punished less harshly than similarly situated defendants who exercise their right to trial. Because of the way Eisenstein and Jacob choose to analyze their data, it is difficult to evaluate this issue in their three jurisdictions. Their conclusion that a sentence differential did not exist in their three jurisdictions seems to be a mis-specification of the issue at hand.

Another recent study which asserts that sentence differentials do not exist is the INSLAW project on plea bargaining in Washington, D.C. (Rhodes, 1978). The study, which uses PROMIS data, appears to have included the relevant defendant and case-related variables (e.g., strength of evidence). Moreover, the data are analyzed in a way which attempts to assess the effects of mode of disposition directly rather than simply in comparison to the effects of other factors. Although the finding is generally characterized as being that there was no sentence differential, what Rhodes in fact reports is that for three of four classes of cases (assault, larceny, and burglary) there was no sentence differential, while for the fourth—robbery—a very marked sentence differential appeared. No satisfactory account is given as to why these offenses might show different disposition effects; moreover, the sentence differential (both as to prison/no prison and sentence length) does appear for the one offense in which a substantial number of defendants were likely to be sentenced to prison. Thus, the Rhodes study does suggest that in the District some

substantial proportion of defendants did not appear to receive the expected sentence differential, while others did. Moreover, although we have no quarrel with the study on methodological grounds, generalizability of the finding from a single jurisdiction seems problematic.

III. OUTLINE OF THE STUDY

Our analysis focuses on disposition patterns for robbery and burglary arrestees in three large California jurisdictions, for the period 1974 through 1978. Most of the data were supplied by the California Bureau of Criminal Statistics (B.C.S.), although some were collected by us in the course of a broader study of the implementation of the California Determinate Sentencing Law. In all we looked at 1759 cases from San Bernardino County, 2514 from San Francisco, and 2520 from Santa Clara County.[1] In order to maximize the number of cases in the trial category, the five-year period was treated as a single unit, although the relationships reported here appear to hold true for individual years as well. Similarly, we restricted our analysis to those whose most serious arrest charge was either burglary or robbery, so that we could focus more directly on the question of sentence differentials, with seriousness of crime controlled for at the point of data collection.

In addition to the original charge and the mode of disposition,[2] we were able to obtain measures of the final conviction charge, the type of attorney who handled the case, and the defendant's previous record, plus the standard demographic variables of age, race, and sex. This set is similar to that employed by Eisenstein and Jacob, except that we do not have a measure of the strength of evidence against the defendant—or an indicator as to which judge heard the case. The absence of these data is undoubtedly a limitation, but an unavoidable one given the character of the official records from which we worked. (We shall argue below that such omissions are unlikely to have affected our results a great deal.)

---

[1] The Santa Clara data contain no cases for 1977, as the County did not report to the Bureau of Criminal Statistics for that year.

[2] In coding mode of disposition, we have combined bench and jury trials. Several factors suggested this as a reasonable way to proceed. First, our observations and interviews in the three jurisdictions did not suggest that bench trials were a form of "slow plea." In addition, they are relatively rare (ranging from 5 to 20 percent in the three jurisdictions during this period). Finally, there was no evidence that those who had bench trials received less harsh sentences than those who had jury trials.

Our major dependent variable was treated dichotomously as state prison/no state prison. The latter category included not only defendants who received jail time, or some combination of jail time and subsequent probation, but also a few who received straight probation. Although it would have been theoretically more desirable to treat jail and probation as separate categories, there were too few cases in the trial category receiving probation for meaningful analysis to have been possible. Because of the nature of California's Indeterminate Sentence Law (which was operative until the middle of 1977), we did not examine the effects of mode of disposition on sentence length. Under the ISL, the judge in prison cases simply sentenced the defendant to the term "prescribed by law" (usually a very open-ended one-year-to-life or five-years-to-life), and the eventual length of the term was determined by the parole board. Thus, there was little opportunity for serious sentence-bargaining in prison cases. As a result, the in/out decision was by far the most important issue to be determined at the court stage of the proceeding.

The procedures used to analyze these data were very straightforward, involving nothing more than multi-dimensional contingency tables. This was possible because of the trichotomous and dichotomous nature of our dependent and independent variables, and because of the relatively large numbers of trial cases which we were able to accumulate. Certainly, other more sophisticated techniques such as regression or discriminant function analysis have the advantage of neatness, but the contingency tables assist in visualizing the form a relationship takes. For instance, a straightforward comparison of coefficients for the various independent variables will indicate whether some kind of relationship exists, but particularly if the amount of variance explained is relatively small, it will tell us little about the form of this relationship; e.g., it will not be clear whether differentials are to some extent applied to all categories of defendants, or whether they are concentrated in one or two sub-groups such as robbers with previous records. More importantly, the use of interval level techniques in the analysis of nominal and ordinal data requires us to violate a number of the assumptions on which these techniques are based. It is sometimes claimed that in the social sciences such violations

are unavoidable, but there is little reason to commit them when a simpler technique seems to suffice.[3]

Table 1. Comparison of Prison Rates, Type of Charge, and Previous Record for Guilty Pleaders and Trial Defendants

		COUNTY					
		San Bernardino		San Francisco		Santa Clara	
		Plea	Trial	Plea	Trial	Plea	Trial
1.	SENTENCE						
	Imprisonment	36	66	29	71	29	65
	Jail or Probation	64	34	71	29	71	35
		100	100	100	100	100	100
2.	CHARGE						
	Burglary	70	53	60	41	76	51
	Robbery	30	47	40	59	24	49
		100	100	100	100	100	100
3.	RECORD						
	None	20	7	13	10	20	12
	Some, but no prison served	58	50	59	44	62	53
	Previous prison term served	21	42	27	46	18	35
		99	99	99	100	100	100
4.	APPROX. N	(1548)	(211)	(2242)	(272)	(2327)	(193)

Note: Percentages sum downward.
Percentages rounded to nearest whole number.

IV. RESULTS

As Table 1 indicates, a straightforward comparison of the disposition patterns in each of the three jurisdictions shows that a higher proportion of trial defendants went to prison than did those who pleaded guilty. But because both those arrested for robbery and those with more serious records were over-represented in the trial category—presumably because courts were less willing to encourage a plea and because the anticipated sentence was so great either way that defendants figured they had little to lose by going to trial—this kind of uncontrolled comparison is likely to be quite misleading. In order to correct for this, therefore, Table 2 compares mode of disposition with sentence while simultaneously controlling for arrest charge and the defendant's previous record. (It should

[3] An alternative approach would be to use techniques like categorical regression (e.g., Kritzer, 1978). Because our emphasis is upon the existence of plea/trial differentials among similarly situated defendants, the actual tables themselves and chi-square analysis suffice to establish and indicate our basic finding.

be noted here that the two groups did not differ significantly in respect to demographic characteristics, and thus additional controls proved unnecessary.)

Table 2. Prison Rates for Guilty Pleaders and Trial Defendants, Controlling for Record and Arrest Charge

Previous Record	Arrest Charge: Burglary		Arrest Charge: Robbery	
	Plea: Guilty	Plea: Not Guilty	Plea: Guilty	Plea: Not Guilty
A. SAN BERNARDINO				
None	14 (211)	9 (11)	38 (109)	75 (4)
	N.S.		N.S.	
Some, but no prison term served	22 (604)	33 (48)	54 (101)	83 (37)
	S.		S.	
Previous prison term served	60 (202)	71 (48)	83 (101)	95 (37)
	N.S.		N.S.	
B. SAN FRANCISCO				
None	16 (134)	33 (6)	42 (105)	55 (37)
	N.S.		S.	
Some, but no prison term served	15 (781)	49 (43)	32 (517)	70 (20)
	S.		S.	
Previous prison term served	41 (392)	81 (58)	53 (191)	87 (64)
	S.		S.	
C. SANTA CLARA				
None	8 (386)	7 (15)	53 (99)	75 (8)
	N.S.		N.S.	
Some, but no prison term served	19 (1025)	47 (47)	52 (336)	72 (50)
	S.		S.	
Previous prison term served	44 (323)	82 (33)	79 (102)	94 (34)
	S.		S.	

Notes: Each entry shows the percentage (rounded) of defendants in this category sent to prison. The total number of defendants falling into this category is shown in parentheses.
The significance test employed was X^2 corrected for continuity. Following convention, a relationship was declared significant if $p \leq .05$. This statistic was computed after cross-tabulating "guilty-not guilty" with "imprisoned-not imprisoned" for each of the combinations of the controlling variables in each of the counties. To improve table readability, we here present only the percentage of defendants imprisoned.

Although the small number of cases in some of the trial categories means that the relationships do not always qualify as statistically significant, the overall pattern of these results is remarkably consistent. In only two instances do prison rates for trial cases fail to exceed prison rates for guilty pleaders by

at least ten percentage points (burglary arrestees with no previous records in San Bernardino and Santa Clara), and these deviations are hardly significant given that only a handful of trial cases are involved.

There is some indication that the *size* of the differential varies both across jurisdictions and between some of the categories, but we are not in a position here to determine whether there is a theoretical reason for this, whether it is due to a nonrandom distribution of "errors" (quite possible because of the small n's in the trial cells), or whether it is because certain other unknown determinants of sentence are not evenly distributed across all of the categories. At any rate, such considerations are largely beyond the scope of the present paper. The main point to emphasize here is that, although the strength of the relationship between mode of disposition and sentence is doubtless reduced once controls are introduced, the relationship obviously cannot be explained as simply an outcome of the different characteristics of the two sets of defendants.

Is it nonetheless possible, as the earlier quote from Feeley implied, that these differences might disappear if more variables were introduced into our analysis? Consideration of three of the more plausible candidates for intervening variables leads us to suggest both *a priori* and empirical reasons why they are unlikely to undermine the clear relationship observed between mode of disposition and sentence.

1. *Nature of the Crime*. Especially as it is recorded by the California B.C.S., "arrest charge" is arguably a fairly crude indicator of the nature of the crime which was actually committed. For one thing, it does not indicate whether multiple or single counts were alleged. For another, it tells us nothing about the "gory" details of the crime—i.e., we do not know whether violence was employed in the commission of the crime, whether a weapon was used, and so on. This leaves open the possibility that the apparent differential could be a result of the fact that those who plead not guilty have committed more numerous or heinous robberies or burglaries than their relevant control group. Fortunately, we are able to employ some additional data for burglary and robbery cases in Superior Court for 1976 and 1978 as a means of partially testing this hypothesis. By using local court records it was possible to determine the total number of charges alleged against each defendant and the relative seriousness of the first three of

these charges.[4] On these measures (see Table 3) there is no indication either that more charges were initially leveled against non-guilty pleaders, or that these charges were of greater aggregate seriousness, with the single exception of robbery arrestees in San Francisco. These findings thus tend to suggest that a more refined measure of the nature of the original charge is unnecessary.

2. *Judges*. Eisenstein and Jacob found that, in the three systems which they studied, the "identity of the courtroom work-group" (which they operationalize as the judge involved) was one of the two or three most significant factors in the in/out decision. To what extent might the differential be due to the fact that lenient judges are over-represented among guilty pleaders, and harsh judges among trial cases?

In order to substantiate this particular hypothesis, three conditions would need to be met. First, obviously enough, it would have to be shown that individual judges in a jurisdiction varied significantly from each other along a leniency-harshness dimension. Second, it would need to be established that these same judges did not engage in differential sentencing themselves. Otherwise, if the contrary were true, our results could understate rather than exaggerate the significance of differentials, due to individual-level differences cancelling each other out (cf. Gibson, 1978, on this "levels of analysis" problem in relation to sentencing patterns). And third, it would have to be shown that relatively harsh judges handled a disproportionate number of the trial cases and vice versa.

We lack adequate evidence in respect to the first two conditions, but it is possible to present some data from 1976 and 1978 with respect to the third. Specifically, our findings indicate that in each of the three jurisdictions studied, a large proportion of the judges heard at least some trials and no one judge heard more than 20 percent. This makes it quite difficult to argue that any one group of judges dominated the trial cases. On the other hand, it is true that responsibility for sentencing of guilty pleaders was concentrated in the hands of one or a

[4] The California Bureau of Criminal Statistics uses a so-called "hierarchy" index to measure the seriousness of crimes. Because the range of numbers is large (running from around 100 to more than 100,000), and because their scale has lower numbers for more serious crimes, we have here made use of a transformation suggested by Charles Hubay in an unpublished manuscript (1978). This involves computing the reciprocal of the hierarchy number for each charge, multiplied by 100,000. The transformed score for each arrest charge is then summed for each defendant in order to give an overall measure of the charge seriousness.

Table 3. Number and Aggregate Seriousness of Charges for Guilty Pleaders and Trial Defendants, Controlling for Most Serious Arrest Charge

		COUNTY											
		San Bernardino				San Francisco				Santa Clara			
		Arrest Charge											
		Burglary		Robbery		Burglary		Robbery		Burglary		Robbery	
		Guilty	Not Guilty	Guilty	Not Guilty	Guilty	Not Guilty	Guilty	Not Guilty	Guilty	Not Guilty	Guilty	Not Guilty
$\bar{X}$	Number of Charges	2.1	1.8	2.6	2.1	1.9	1.8	2.7	4.3	2.5	2.3	2.7	2.4
$\bar{X}$	Seriousness	30.6	26.5	35.9	35.8	27.0	27.1	36.8	44.6	32.4	29.9	37.5	39.4
N		(424)	(46)	(180)	(41)	(497)	(23)	(438)	(54)	(629)	(33)	(416)	(63)

Note: Only data for 1976 and 1978 were collected.

few individuals, due to the fact that all three jurisdictions used a master sentencing calendar system.[5] But for two reasons, this is unlikely to explain the differential. The first is that a significant proportion of defendants in each jurisdiction pleaded guilty before a variety of different judges and then had their cases returned to the master calendar for sentencing by the master calendar judge. In such cases, the deal worked out among prosecutor, defense attorney, and the judge taking the plea almost invariably settled the issue of prison/no prison and hence bound the hands of the eventual sentencer. For cases in which the defendant both pleaded guilty before and was sentenced by the master calendar judge, the argument that lenient judges are more often involved in pleas and harsher judges more often in trials requires the assumption that the master calendar position attracts or is given to more lenient judges, presumably because they will facilitate more pleas. But in each of the jurisdictions the position is rotated regularly, usually once per year, among a relatively large pool of judges. Thus, over time there is no strong reason to expect one or the other type to dominate this position.

3. *Strength of Evidence*. As stated earlier, we do not have either direct or indirect measures of the strength of the case against the defendant. However, it seems unlikely that this variable would have much impact on the observed relationship. On the one hand, there is the possibility that where the evidence is overwhelmingly strong, defendants may be more inclined to go to trial, *viz*: if prosecutors are less willing to bargain in these types of cases (since a conviction is so certain), defendants may be more likely to decide that they have nothing to lose by pleading not guilty. But on the other hand, to the extent that the trial docket is comprised of cases in which there is some real issue about guilt or innocence (cf. our discussion above), the evidence against trial cases should tend to be weaker rather than stronger. Putting these two arguments together, the best conclusion appears to be that there is no good *a priori* reason to expect the *average* strength

[5] Under the master calendar system prevailing in the three counties, a single judge was administrative head of the Superior Court system. His functions included presiding over a substantial proportion of pretrial (plea-bargaining) conferences, as well as doing the majority of sentencing of defendants who plead guilty. All three systems used a version of the master sentence calendar, by which most cases which resulted in a guilty plea were returned to a central sentencing court, regardless of the judge who took the plea. In San Bernardino and San Francisco, the master calendar judge did virtually all the master sentence calendar, while in Santa Clara the master sentence calendar was spread among several judges.

of evidence to be much different for the trial category than for the guilty pleaders category.

V. IMPLICATIONS

We have argued that the evidence presented here suggests quite clearly that a substantial sentence differential existed in each of the three jurisdictions studied. Although there is some possibility that variables not included in our analysis may have an impact on the relationship between mode of disposition and sentence, we have sought to show above that such an impact, if it occurs at all, should be minor. Moreover, as noted earlier, statistical approaches to the sentence differential issue should be treated with great care because a "successful" differential sentencing policy may reduce the differentials actually observed. Thus, there seems as much reason to believe that our findings underestimate the true role that differentials play in these systems as to believe that the observed differentials are spurious.

The more challenging questions concerning this study relate much less to validity and more to generalizability. Specifically, what does this tell us about other sentencing decisions in these courts? And, most importantly, to what extent are these findings likely to hold true for other court systems? We shall attempt to answer briefly each of these questions in turn.

Intra-Jurisdiction Generalizability

The courts studied here made decisions not only on whether defendants should be sent to prison, but also on whether they should be sent to jail (and if so for how long) or given probation. Moreover, from mid-1977, following the introduction of the Determinate Sentencing Law, they began making decisions about the length of time to be served in prison cases as well. Because of limitations in our data set, primarily relating to the small n's in the trial category, we were unable to examine these decisions directly. But there seems no reason to assume that the courts, which were apparently so willing to use prison sentences to reward or punish, should have suddenly become coy when it came to questions of jail and probation.

A more difficult issue relating to intra-jurisdiction generalizability is raised by the fact that we have examined sentence patterns only for burglary and robbery arrestees. Just because differentials are found in respect to these relatively

serious crimes, it does not necessarily follow that they will be maintained for lesser crimes as well. Indeed, the INSLAW study found differential sentences only for the most serious offense examined (Rhodes, 1978). Provided it can be shown that the main reason for differentials is to induce guilty pleas, it can also be argued that the size of the inducement will decrease as charges become less serious, because other costs attached to going to trial will assume proportionately greater significance. A defendant who is arrested for small-time drug possession may forgo a trial even in the absence of an anticipated reward, because the likely sentence is so small and because pleading guilty minimizes time, expense, and uncertainty (assuming, of course, that the defendant considers conviction to be likely). A robber with a previous record, on the other hand, might still choose a trial, even if conviction is probable, because the anticipated sentence is so large that the defendant is prepared to expend considerable resources to gain a small chance of avoiding it. In the case of the latter group, there is clearly less of a natural incentive for defendants to plead. Therefore, the inducement to plead must be larger than that which would have to be offered to someone from the former group (cf. Nardulli, 1978 for an extension of this argument).

It may be logical that differentials should vary in size under these circumstances, but it is less clear that this is what actually happens. As Table 2 indicated, within our sample the size of the differential did not seem to be systematically affected by the defendant's previous record or, except in the case of San Bernardino, by the nature of the crime charged. This provides some empirical basis for arguing that a similar differential might likewise be maintained for other categories of defendants not included in this study. There are several theoretical reasons why this might be so.

First, punishments, and to a lesser extent rewards, are likely to be valued for their own sake rather than just because they can be used to obtain pleas. For instance, if participants believe that it is wrong for defendants to go to trial on "frivolous" grounds, they will be inclined to punish a defendant in a "petty" case, even though the resultant sentence may be much greater than would be necessary to discourage defendants in other petty cases from going to trial. Indeed, precisely because the case is regarded as trivial, the waste is likely to be perceived as greater than where serious charges are involved.

Second, court participants are concerned with legitimizing their behavior not only to the outside world, but also to each other. It would seem more difficult to offer a rationale for differentials for some crimes but not others, than to justify their use in general. The former requires an admission that sentences are being manipulated deliberately in order to obtain pleas, whereas the latter can be defended on broader grounds such as "frivolity should be discouraged," "cooperation and contriteness should be rewarded," and so forth. This consideration is likely to be quite important, given that court personnel usually attach considerable value to the consistent application of rules, whether these rules are formal or informal.

In sum, then, there are theoretical reasons to expect differentials to characterize not only the serious cases examined but other cases as well. On the other hand, the data and arguments presented by Rhodes and Nardulli suggest that sentence differentials may vary with type of charge. Given these uncertainties, the safe conclusion would seem to be that the issue needs further study.

Cross-Jurisdiction Generalizability

A more fundamental issue of generalizability is whether the data obtained for these three systems apply to other court systems as well. The best response to this is that the jurisdictions reported on here *did* differ from each other along a variety of dimensions, such as general level of severity, nature of the local political climate, and manner in which the plea-bargaining process was structured. The fact that differentials were found in all three, despite these variations, is a fairly strong indication that their presence is not restricted to any particular type of system. (Admittedly, these jurisdictions also had an important characteristic in common—each functioned within the context of California sentencing laws—yet there seems no reason why this law should account for the presence of differentials.) Moreover, as we have seen, such a claim of generalizability is indirectly supported by the reported perceptions of participants themselves and by the fact that, with the notable exceptions of Eisenstein and Jacob's book and the INSLAW study (which dealt with a single jurisdiction), little or no empirical work has claimed to have identified systems where significant differentials are absent. This is not to say that such systems cannot be found, but there is as yet little reason to expect that they will be other than deviations from the central tendency.

VI. SENTENCE DIFFERENTIALS: SOME MORE GENERAL CONSIDERATIONS

If we accept the argument that sentence differentials characterize most criminal courts, we still must examine the question of why the practice is so widespread. The standard explanation, so far as we can tell, is one of functional necessity: sentence differentials are common because they are considered by courtroom participants as the only way to induce sufficient numbers of defendants to plead guilty. Large numbers of guilty pleas are valued because of caseload considerations or simply because most trials are seen as a waste of time (Heumann, 1978; Feeley, 1979b).

We have little to quarrel with in the essence of this argument. Although there is still much to learn about the process by which defendants arrive at the decision to plead guilty, available evidence suggests that in addition to cost and time considerations and the desire to minimize uncertainty, many defendants would not plead guilty absent a belief that they stood a good chance of receiving a lighter sentence. For instance, in a survey conducted by Casper (1978) in 1976 of defendants in Baltimore, Detroit, and Phoenix, a very large proportion of defendants in each system answered in the affirmative when asked if their final plea had come about *as a result* of a plea bargain they had discussed with their lawyer (68 percent in Baltimore, 78 percent in Detroit, 87 percent in Phoenix). Whether these same defendants would have eventually pleaded guilty even in the absence of a bargain is harder to determine, but it seems extremely unlikely that enough of them would have done so to produce the 85 percent or higher plea rates found in most jurisdictions. Even more direct support for the functional necessity argument is provided by two recent studies of the effects of "no deal" policies on plea rates. In both Church's analysis (1976) of the impact of a law-and-order, "no bargains" chief prosecuting attorney on disposition patterns in a midwestern county, and Heumann and Loftin's recent (1979) Wayne County study of the impact of a no-bargaining policy vis-a-vis the Michigan Gun Law, the authors argue that trial rates were kept low following the implementation of these policies *only* because each system developed new ways of offering concessions to potential guilty pleaders.[6]

[6] Some ostensibly contrary evidence might appear in the Alaska experience, where the plea rate has remained high despite the "abolition" of plea bargaining. But on closer reading, there is no evidence that the end of

Despite the fact that the functional necessity argument appears to be a strong explanation for the prevalence of sentence differentials, a number of challenges to its sufficiency have been mounted. These include Eisenstein and Jacob's assertion that it is the *illusion* of a differential, not its substance, that is necessary, as well as some suggestions that factors in addition to its utility in inducing guilty pleas may be responsible for the existence of a differential.

In attempting to account for their surprising finding that differentials did not exist in three jurisdictions they studied, Eisenstein and Jacob suggest that what is functionally necessary is not an actual differential but a widespread *belief* on the part of defendants that trials are punished by harsher sentences. They maintain, moreover, that this belief might be sustained even in the absence of actual differentials. It must be acknowledged that defendants are probably more susceptible to embracing such a belief, even if it is ungrounded, than are other court participants, for the simple reason that their contact with the court system is much more sporadic and thus affords less opportunity for "testing out" these beliefs. In addition to whatever past experience they may have had, defendants have two primary sources of information about the existence of differentials: the jailhouse culture in which they may be awaiting disposition, and the advice given them by their attorneys. The jailhouse culture provides a wider range of past experience, but presumably matters less than the specific advice given by a defendant's attorney about what happens in general in the court or what is likely to happen in the defendant's particular case. Although defendants may mistrust their attorneys on many dimensions, they are likely to believe a gloomy prediction about the consequences of a trial, if for no other reason, because of the possibility of a self-fulfilling prophecy.

Thus, we would argue that most defendants would be likely to believe incorrectly that a differential exists only if their attorneys were engaging in deliberate deception or were

---

plea bargaining also signaled the end of perceived sentence differentials. In fact, to the contrary, Rubinstein and White (1979: 380) point out in a footnote that defense attorneys sometimes said that ". . . they were reluctant to go through a full trial when there was no 'triable' question of fact; they feared that the judge would disapprove of the expenditures of state money and time on such cases. Under these circumstances, there was often a suspicion lurking in the lawyer's mind that his convicted client would have to pay a bill in the end, perhaps in the form of a longer sentence." On this evidence, therefore, there seems little reason to reject the assumption that many defendants plead guilty only because they expect a lighter sentence in return.

themselves misled. We doubt that systematic deception of defendants by their attorneys is the norm in many jurisdictions. Although attorneys may wish to avoid trials, in most instances internalized and group-enforced norms would seem sufficient to prevent consistent and widespread lying.

This brings us, though, to a second possibility: that the people doing the advising are themselves misled. In addition to the previously advanced arguments about the ability of participants to test out beliefs and control for effects of mode of disposition, two other arguments seem to militate against the notion that lawyers typically share a belief that differentials exist when they in fact do not.

Often the object of a belief is independent of the belief itself, in the sense that it will continue to exist whether or not it is believed to exist. In a court system, however, this is much less likely to be the case. Rather, if participants believe something to be true they will tend to act in a way that makes it true. (In this sense questions as to whether initial perceptions are accurate or not are to some extent irrelevant.) Specifically, if incoming judges and prosecutors believe that incumbent judges and prosecutors are favorably disposed to guilty pleaders when it comes to sentencing or charging, they will themselves come to act in this fashion (cf. Heumann, 1978, generally on this point). Similarly, if defense attorneys believe that guilty pleas get rewarded, they will attempt to ensure that such rewards are also forthcoming to their clients. In this way, a belief that sentence differentials exist may be a self-fulfilling prophecy.

Another quite different argument as to why participants are unlikely over time to perceive differentials where there are none is that court systems periodically generate their own verifiers. For instance, as pointed out earlier, many trials occur after plea negotiations have taken place but then have been broken off when no final deal can be agreed upon. Those close to such cases know what a defendant could have obtained had he or she pleaded guilty and are able to compare this with the sentence actually imposed after trial. Presumably, if such defendants did no worse over time, the belief that differentials existed would begin to decay. Likewise, occasional tests are provided by new participants entering the system. New public defenders, for example, are both less likely to accept the existence of differentials and less likely to be attuned to informal definitions of what kinds of cases should and should not be tried. What happens when they attempt to take cases to

trial will thus serve as an indicator not only to them but to other participants of the extent to which differentials really exist.

This does not mean necessarily that all systems with high plea rates will always be characterized by differentials. For instance, there may be particular small jurisdictions where more or less systematic deception of defendants takes place. It is also possible to argue that although beliefs about differentials are ultimately dependent on their existence, a degree of "loose-coupling" is involved—i.e., once a belief gets established, it takes some time to decay, even though empirical support for it no longer exists. (We would assume that this decoupling is likely to be greatest in highly integrated systems where internal socialization mechanisms are at their most effective.) However, we believe that in most systems, for most of the time, the evidence overwhelmingly indicates that sentence differentials will prevail.

Although we have sought to establish that sentence differentials are, in a broad sense, functionally necessary for the maintenance of high guilty plea rates, that may not be the only reason *why* they exist. Two other arguments about sentence differentials have often been discussed: the propensity of judges to reward guilty pleaders independent of caseload considerations, and structural properties of the trial and plea settings which may militate towards greater leniency to guilty-pleaders.

Interviews with judges (e.g., the *Yale Law Journal* survey) reveal many who argue that defendants who plead guilty *deserve* more lenient sentences. Several reasons are typically offered, including the suggestion that defendants who plead guilty have taken the first step towards rehabilitation. In addition, they have not angered judges by taking the stand and perjuring themselves by denying crimes the judge "knows" they have committed. Finally, one sometimes encounters an "equity" argument to the effect that, because the defendant who goes to trial retains the chance of acquittal, it would be unfair to those who pleaded guilty if this trial defendant were also rewarded with a light sentence if convicted. In the case of all of these assertions, the fact that the effect is to discourage defendants from going to trial is said to be incidental to the "real" purpose of the differential—the implication being that even if guilty pleas in such jurisdictions began to decline, sentencing patterns would not be modified.

It is, to be sure, difficult to know how much weight to give to these asserted justifications. No doubt such arguments are often used simply as a means of rationalizing behavior to outsiders and to each other. Even when the belief is sincerely held, this may only be because participants have lost track of the underlying purpose of the differential. As, for example, the Church and Heumann and Loftin studies suggest, they might be much more aware of the time-saving virtues of differentials if they were suddenly unable to employ them. But there seems no doubt that such considerations do motivate some judges. Just how many and how much remains a question for future research.

Another alternative to the functional necessity hypothesis focuses upon the assertion that differentials are a product of varying structural properties of the plea and trial settings. From this perspective, defendants who go to trial might continue to do worse even if there were no conscious effort on the part of other court participants to use sentences to reward or punish. A trial conviction may not allow as much flexibility in sentencing decisions as a plea bargain does. Bargaining permits prosecutors and defense attorneys to select a conviction charge which may not fit the evidence but which permits a lesser sentence; in a trial they lose control over the conviction charge and the constraints it places on ultimate sentence. In addition, the trial is likely to produce more publicity and attendant public scrutiny, as well as providing the sentencer with much more detail about the nature of the harm done by the defendant. Both of these may militate against the leniency that often attends the privacy and flexibility of the plea bargain.

Interestingly, similar arguments have provided the basis for recent claims that plea bargains may be more suited to dispensing "substantive justice" than are adversarial approaches, precisely because the greater flexibility of the former makes it possible to tailor sentences to the situation of the individual defendant (cf. Mather, 1979; Utz, 1979; Heumann, 1978; Casper, 1979). Although little research has so far been directed toward the role of these structural properties, we consider it quite unlikely that they alone will be able to explain why differentials are so widespread. Nonetheless, it does appear that their impact may be sufficiently great to again make us wary of embracing functional explanations to the extent of implying that the presence of differentials is *always*

and only the result of a conscious desire to maximize the number of guilty pleas.

VII. CONCLUSION

The point of the preceding discussion is simple enough: when guilty plea rates are high, expect to find differential sentencing. We believe that recent arguments to the effect that differentials are largely illusory do not withstand serious scrutiny, even though this revisionist challenge has been valuable in forcing us to examine more closely what is too often taken to be self-evidently true. In support of our case, we have provided evidence from a study of three large California jurisdictions, along with some more theoretical arguments as to why defendants would be a good deal less willing to plead guilty in the absence of a sentence-related inducement. The paper thus reaffirms what has been, for a long time, one of the fundamental tenets of our understanding of how courts function. Ours is not the final word on the subject, but we would claim that a good deal more evidence and argument is required before this assumption can be called into serious question.

REFERENCES

ALSCHULER, Albert W. (1968) "The Prosecutor's Role in Plea Bargaining," 36 *University of Chicago Law Review* 50.

——— (1975) "The Defense Attorney's Role in Plea Bargaining," 84 *Yale Law Journal* 1179.

——— (1976) "The Trial Judge's Role in Plea Bargaining, Part 1," 76 *Columbia Law Review* 1059.

CASPER, Jonathan D. (1972) *American Criminal Justice: The Defendant's Perspective*. Englewood Cliffs, N.J.: Prentice Hall.

——— (1978) *Criminal Courts: The Defendant's Perspective*. Washington, D.C.: Government Printing Office.

——— (1979) "Reformers v. Abolitionists: Some Notes for Further Research on Plea Bargaining," 13 *Law & Society Review* 567.

CHURCH, Thomas W., Jr. (1976) "Plea Bargains, Concessions and the Courts: Analysis of a Quasi-Experiment," 10 *Law & Society Review* 377.

——— (1979) "In Defense of 'Bargain Justice,'" 13 *Law & Society Review* 509.

EISENSTEIN, James and Herbert JACOB (1977) *Felony Justice: An Organizational Analysis of Criminal Courts*. Boston: Little Brown.

FEELEY, Malcolm M. (1978) "The Effects of Heavy Caseloads," in S. Goldman and A. Sarat (eds.) *American Court Systems: Readings in Judicial Process and Behavior*. San Francisco: W. H. Freeman.

——— (1979a) *The Process is the Punishment*. New York: Russell Sage Foundation.

——— (1979b) "Perspectives on Plea Bargaining," 13 *Law & Society Review* 199.

FRIEDMAN, Lawrence M. (1979) "Plea Bargaining in Historical Perspective," 13 *Law & Society Review* 247.

GIBSON, James L. (1978) "Race as a Determinant of Criminal Sentences: A Methodological Critique and a Case Study," 12 *Law & Society Review* 455.

HEUMANN, Milton (1975) "A Note on Plea Bargaining and Case Pressure," 9 *Law & Society Review* 515.

——— (1978) *Plea Bargaining: The Experiences of Prosecutors, Judges, and Defense Attorneys*. Chicago: University of Chicago Press.

HEUMANN, Milton and Colin LOFTIN (1979) "Mandatory Sentencing and the Abolition of Plea Bargaining: The Michigan Felony Firearm Statute," 13 *Law & Society Review* 393.

HUBAY, Charles (1978) "California Determinant Sentences Impacts." Santa Monica, California: Rand Corporation, unpublished paper.

KRITZER, Herbert M. (1978) "An Introduction to Multivariate Contingency Table Analysis," 22 *American Journal of Political Science* 187.

MATHER, Lynn M. (1973) "Some Determinants of the Method of Case Dispositions: Decision-Making by Public Defenders in Los Angeles," 8 *Law & Society Review* 187.

——— (1979) *Plea Bargaining or Trial? The Process of Criminal Case Disposition*. Lexington, Mass.: Lexington Books.

NARDULLI, Peter F. (1978) *The Courtroom Elite: An Organizational Perspective on Criminal Justice*. Cambridge, Mass.: Ballinger.

NEWMAN, Donald J. (1956) "Pleading Guilty for Considerations: A Study of Bargain Justice," 46 *Journal of Criminal Law, Criminology and Police Science* 780.

——— (1966) *Conviction: The Determination of Guilt or Innocence Without Trial*. Boston: Little Brown.

NISBETT, Richard and Lee ROSS (1980) *Human Inference: Strategies and Shortcomings in Social Judgment*. Englewood Cliffs, N.J.: Prentice Hall.

RHODES, William M. (1978) *Plea Bargaining: Who Gains? Who Loses?* Washington, D.C.: Institute for Law and Social Research.

RUBINSTEIN, Michael L. and Teresa J. WHITE (1979) "Alaska's Ban on Plea Bargaining," 13 *Law & Society Review* 367.

UHLMAN, Thomas M. and N. Darlene WALKER (1979) "A Plea is No Bargain: The Impact of Case Disposition on Sentencing," 60 *Social Science Quarterly* 218.

——— (1980) " 'He Takes Some of My Time, I Take Some of His': An Analysis of Sentencing Patterns in Jury Cases," 14 *Law & Society Review* 323.

UTZ, Pamela J. (1978) *Settling the Facts: Discretion and Negotiation in Criminal Court*. Lexington, Mass.: Lexington Books.

VETRI, Dominick R. (1964) "Guilty Plea Bargaining: Compromises by Prosecutors to Secure Guilty Pleas," 112 *University of Pennsylvania Law Review* 865.

YALE LAW JOURNAL (1956) "Comment: The Influence of the Defendant's Plea on Judicial Determination of Sentence," 66 *Yale Law Journal* 204.