BLINDFOLDED JUSTICE
LED BY AN
INVISIBLE HAND

Criminal Justice as a Variable
of Exchange Transactions

ELI M. NOAM
Columbia University

The article argues that instead of viewing the criminal justice process as
determined by judicial or prosecutorial discretion, it should be seen as the
outcome of exchange transactions that are designed to adjust resource
constraints between case inflows and outflows. These exchanges, over
which judges and prosecutors have only little control in the aggregate,
generate changes in the criminal justice system in an unplanned or
"invisible hand" fashion. They alter the sentence level, affect the standard
of proof and the rules of evidence, and transform the function of trials.
Several of the analytical conclusions are confirmed by an empirical
investigation.

The traditional view of criminal justice accords to the judge a
preeminent role in the shaping of the criminal process. It is
the judge who applies the law, evaluates the admissibility of
evidence, metes out punishment, and often determines guilt or
innocence. This traditional model of criminal justice may be
termed "judicial," as distinguished from the "prosecutorial" model,
which has largely superseded it in the view of observers
holds that most decision-making has shifted from the courtroom
to the prosecutor's office. A major manifestation of this move is
basic point, however, is that reapportionment’s effects need to be examined in conjunction with districting changes.

Future research is needed to clarify the theoretical conditions under which empirical tests are conducted. We applaud Hardy and Newcomer’s effort to test hypotheses more rigorously. Yet, we believe, additional effects of changes in legislative structure (e.g., redistricting) need to be taken into account in order to provide a more complete explanation of the variables that may influence policy.

NOTE

1. In addition to our study, Hardy and Newcomer question the results of Douglas Feig’s (1978) research.

REFERENCES


ROGER A. HANSON is Senior Researcher at the Institute for Court Management. He is currently directing experimental and quasi-experimental studies in the civil and criminal justice field, including the use of telephone hearings to conduct court business, expedited appellate procedures, mediation of prisoner disputes, and discovery rule limitations.

ROBERT E. CREW, Jr., is Principal Research Scientist at the American Institutes for Research. His major interests revolve around the problems involved in the evaluation of legal and social programs. He holds a Ph.D. in political science from the University of North Carolina and has been Assistant Dean at the Woodrow Wilson School of Public and International Affairs and a member of the political science faculty at the University of Minnesota.
the prevalence of guilty pleas,¹ a form of case disposition in which the prosecutor, after some bargaining, recommends a sentence that is nearly invariably accepted by the court, thus reducing the judiciary's function in most cases to mere ratification (Alschuler, 1968; Newman, 1966). In New York City, for example, the percentage of criminal dispositions by trial, rounded off to the nearest integer, is actually zero.

In contrast to both the "judicial" and the "prosecutorial" approaches, this article, argues for an "exchange" model of criminal justice. This means that many basic institutional aspects of the criminal process are determined not by legislative, judicial, or executive discretion, but are instead impersonally generated by a multitude of exchange transactions between the parties in criminal cases, the purpose of which is to reconcile the caseload of the court with its capacity to process these cases.

The prevalent belief in the discretionary powers of prosecution is natural, given its daily exercise as witnessed by criminal lawyers. But to conclude that the prosecution is therefore the locus of decision-making power is either a fallacy of composition or a failure to distinguish between individual decisions and their aggregate effect—between what economists label micro and macro analyses.

To use an analogy: A car dealer has considerable discretion in the sense that he or she may adjust a car's price, assess the value for a trade-in, extend different warranties, and accept payment by various methods. Yet in all likelihood the dealer's decisions represent merely fluctuations around some prices and industry practices over which he or she has no control. The dealer's discretionary powers are limited if he or she wants to stay in business. Of course the situation of prosecutors is different from that of persons engaged in commercial activities; but the example shows that discretion can be more apparent than real.

The lack of actual discretion has several implications for the workings of the criminal justice system. To demonstrate them, this article discusses the changes in sentencing as well as in convictions, standard of proof, and admissibility of evidence that are caused by the inexorable need for negotiated transactions.
The results of three empirical tests are also briefly reported in order to buttress the arguments.

**SENTENCES**

Under a "judicial model," a jury weighs the evidence for conviction, and the judge, after considering all circumstances, determines the sentence. By contrast, in a "prosecutorial" model these decisions are made by the prosecutor. Yet in contrast to a judge or a jury, a prosecutor is dependent on an agreement with the defendant. Without such agreement, there is no guilty plea; thus, prosecutorial discretion is constrained by the need to keep cases moving through court. Furthermore, the bargaining process over such an agreement does not take place in a vacuum but is affected by general conditions within the court. In a congested court system, only a very limited number of cases can be tried, while the majority of cases must be disposed of by inducing the defendant to plead guilty. Since a guilty plea is normally extended by a defendant in return for a reduced sentence over the one that could be expected after trial; this "sentence discount" is, in effect, the "price" the prosecutor pays to obtain a guilty plea. In a heavily congested system the prosecutor must induce a relatively large number of defendants to plead guilty, or the processing of the case flow would come to a near stand-still. Thus, the higher the required percentage of guilty pleas, the better, on average, must be the terms offered to defendants. In other words, the "discount" in sentences is related to the number of guilty pleas that is required in order to clear the case docket, just as prices in general are associated with some quantities of goods or services. The extent of the plea-bargain discount is therefore only seemingly in the hands of the prosecution. Sentence discounts are not so much "set" by prosecutors as arrived at by the multitude of bargains between prosecutors and defense attorneys on many unrelated cases; a change in sentence discounts need not be based on a conscious policy decision by a prosecutor's office (Jacoby, 1980), but is rather the gradual result of near-imperceptible responses to work-
load pressures and requirements for guilty pleas in order to clear dockets. With prosecutors and defense attorneys engaged in numerous daily plea transactions, certain levels of expected discounts are established and an equilibrium is reached where the dockets are cleared. While some discretion in granting discounts is possible, there is a limit to it if the prosecutor deals with defense attorneys who know the “going rate,” and insist on getting it for their clients.\(^2\)

An increase in the sentence discount can occur in different ways: by a reduction in the plea sentence, by an increase in the trial sentence, or by a combination of both. Graphically, this can be shown by the two curves of Figure 1, in which sentence severities change with court congestion, such that guilty-plea sentences decrease as the courts get overloaded, and trial sentences increase at the same time.\(^3\)

The substantiation of these functional relations is a question for empirical research. Such a study has been undertaken,\(^4\) and it confirms the existence of curves of the kind drawn in Figure 1.\(^5\)
Briefly, the results show that, with a statistical significance at the 95% confidence level, each 10% increase in court congestion is associated with an approximately 2% reduction in the sentences that follow a guilty plea, and a 1% increase in sentences that follow a trial conviction.⁶

The conclusion that resource pressures in the courts affect sentencing seems straightforward enough, though Heuman (1975, 1977) seems to come to a different conclusion; however, his results can be reconciled, since he only looks at quantities of plea-bargaining (absolutes or rates), which is not conclusive without an investigation of changes in the price of the plea, that is, in the discount in sentence severity. Quite possibly, the stable number of trials that was observed for Connecticut had been "bought" by more lenient plea sentences; that is, the sentence level had been changed in order to keep the number of trials constant.

What this means is that in jurisdictions where congestion is small, plea sentences are higher than in courts that are congested. The sentence disparities that exist across jurisdictions are therefore, at least partly, not based on different attitudes of judges that can be dealt with by fiat, persuasion, or other well-meant efforts to "unify" standards. Rather, part of sentencing disparity is created by the difference in the courts' caseloads relative to their resources.

A second observation is that the effect of congestion relegates to impotence the statutory sentences that are set by a legislature. In theory, the range of criminal sentences is given by statute; it is the expression of the societal evaluation of the seriousness of the offense and of the appropriate sanction for it. In the judicial model of the criminal justice process, the judge selects the appropriate measure of punishment within that statutory range. In the prosecutorial model, the same discretion is exercised by the prosecutor, who must take into account the remorse of a defendant as evidenced by a plea of guilty. But when the bulk of dispositions is by guilty pleas, as in most urban jurisdictions, the plea-bargain sentence level is for all practical purposes the effective sentence, while the statutory sentence has become merely a reference point. The plea-bargaining sentence may have
an only incidental relation to the legislature's sentence range. Because the discrepancies between actual and statutory sentences become a public embarrassment, we witness the vogue of the "mandated" sentence, whose purpose is to eliminate the plea-bargain discount entirely. As long as courts are as congested as they are, these attempts are doomed to fail. Most probably, a prosecutor will reduce the charges, thereby keeping his or her bargaining option intact. Alternatively, the prosecutor must hold a trial, since the defendant has no incentive to plead guilty; but the diminished resources for the trying of other cases will lead to their receiving lower sentences. Thus, what seems to be a new and "tough" policy is in reality a tradeoff in severity between different classes of offenses.

In order to dispose of cases, the sentence discount may have to be large; and it can be made large by raising the reference "price" of jury and bench trials. The function of trial sentences, precisely because they are rare, then becomes one of price-setting. Given that nearly all sentences are plea-bargained, the notion that these sentences are reductions is true in a technical sense only. A better description is that there is a surcharge for a trial, that is, a price for the exercise of the constitutional right of trial, and that the magnitude of the price is based on the pressures within the system. What makes this so troublesome is that the price—the higher sentence—exists not in order to deter crime, but to deter trials. A trial sentence is, at least in part, an administrative accommodation and an element to keep the court system in equilibrium, rather than a "just" punishment. The consequences of this can be an extreme burden that is borne by those who still go to trial and are found guilty (Lachman and McLaughlan, 1977). Their punishment may have to be draconian in order to keep the guilty plea sentences relatively high, if that is the strategy, and will reduce the strategy to absurdity. An example is found in the case of Bordenkircher v. Haynes (1978), in which a defendant who was charged with passing a forged check for $88.30 and who refused to plead guilty was sentenced to life imprisonment, with the Supreme Court upholding.
CONVICTIONS, STANDARD OF PROOF, AND EVIDENCE

With court congestion transforming the criminal process increasingly into a system of exchange, the probability that a case will end in a conviction is also lowered. Intuitively, this is not obvious: On one hand, a guilty plea eliminates the probability of a case resulting in an acquittal at the end of a trial, and this tends to increase the probability of conviction; on the other hand, cases may be less vigorously pursued in a resource-constrained criminal justice system. These two effects operate in different directions, and the overall effect is ultimately an empirical question. (The overall effect that was found is a decrease of about 1.3% in the probability of conviction with each change of 10% in congestion, with a very high degree of statistical significance.) Thus, the more congested the court, the lower is the probability that a case will end in a conviction. The probability of conviction therefore does not depend solely on the merits of a case, but is also a function of court resources.

The decreased conviction rate suggests that proportionately more innocent people are tried. What may be termed a “crowding out” of guilty defendants by innocent ones occurs in a congested court. As courts are more heavily burdened and as the sentences drop for plea bargains and increase for trials, the incentive for guilty defendants to plead guilty is stronger than before. This incentive does not exist to the same extent for defendants who believe that they are innocent and who hence rate their chances of acquittal as fairly high (unless they are as a group very risk-averse). Most innocent defendants are likely to keep demanding a trial, and their relative number among those tried increases, since their likelihood to plead guilty is less sensitive to the expected sentence than that of guilty defendants. Because trials are much more expensive than guilty-plea negotiations, the implication is that the prosecutor’s office must spend a large share of its budget trying innocent defendants rather than guilty ones, relative to the allocation in an uncongested court. This seems an undesirable
result, but since it is generated by the realities of resource pressures within a system, it cannot be corrected by a simple policy decision.

STANDARD OF PROOF AND EVIDENCE

Another element of the criminal process that is transformed by exchange is the standard of proof for conviction. Under the traditional model, a defendant is convicted if his or her guilt is established beyond a reasonable doubt. But in reality, when nearly all cases are disposed of by guilty plea, this standard need not be met by a prosecutor, because in plea-bargaining one can trade doubt for sentence reductions. The higher the doubt and the weaker the prosecutor's bargaining position, the greater is the sentence reduction offered to a defendant (Kalven and Zeisel, 1966). Thus, the traditional two-stage procedure of a trial, where first guilt or innocence is determined and then the sentence is set, is integrated into a single stage. Instead of the doubt determining the verdict, it now determines the sentence. This need not be viewed as undesirable; under the judicial model, a small difference in the evidence can make a large difference in the outcome because the verdict is a dichotomous variable. Under the exchange model, however, these discontinuities are smoothed out through variations in the sentence. The consequence for such a one-step system is that the standard of reasonable doubt has become irrelevant. Its effect is restricted to the limited number of trials; for the majority of cases, the measure of proof becomes an indeterminate concept.

Exchange transactions induced by resource constraints also affect the rules of evidence. Because so few cases end up in court, rules as to what evidence will be admissible are less important than under the traditional model. Rather than affect the question of conviction, they primarily affect the sentence. For example, a class of evidence that has become inadmissible because of a Supreme Court ruling does not cease to be important. Instead, its trading value declines. This decline may perhaps be considerable
but is certainly not total, since the evidence, even if inadmissible, will affect the assessment of the prosecutor and defense of the case and of potential admissible evidence that may yet be developed. There is therefore no reason to believe that the prosecutor's or the police's acquisition of excluded evidence will end because it has become inadmissible in court. Since it is still of value, albeit a reduced one, its acquisition may still be a rational strategy if the cost of acquiring alternative and admissible evidence is high in relation to its trading value. An exclusionary rule is therefore no barrier to the collection of inadmissible evidence, but acts more like an excise tax, reducing the value of the evidence and hence its collection, but not eliminating it.

There is also a second effect of an exclusionary rule of evidence. The inadmissibility of certain previously admissible evidence makes a prosecutor's collection of evidence more costly in many cases. It also tends to provide more potential grounds for appeal. If the prosecutor's resources are unchanged while the cost of marshaling evidence has increased due to an exclusionary rule, the result is an added need to dispose of cases by guilty plea instead of by trial. With the demand for pleas increased, the sentence discount must be increased, too. This, as has been shown, will tend to raise jury trial sentences and reduce plea sentences. For those defendants who opt for a trial, the potential sentence thus becomes more severe. Defendants who exercise their right to trial and are convicted, in effect, pay for the new rule through increased sentences. Some of the risk-adverse defendants will therefore plead guilty instead of going to trial as before. For those who plead guilty, on the other hand, sentences become lighter. They profit from new rulings that may have no relation to their own cases.

CONCLUSION

We have seen that important elements of the criminal justice process are affected by the need for exchange transactions between prosecutors and defendants that are caused by caseload
pressures. The analysis indicated that this pressure causes guilty-plea sentences to be lighter and trial sentences to be more severe. This has implications for the relative uniformity of sentences among courts, the standard of proof, the allocation of resources on cases, and the effect of an exclusionary rule. These changes are not planned, but are instead impersonally generated via exchange transactions. In that sense, important parts of the criminal justice system are not consciously directed, but can instead be described as led by an invisible hand.

The absence of direct control by public policy makers is a troubling conclusion, given the importance of the crime problem in American society. Yet it is a problem susceptible to improvement. One remedy is to reduce the congestion that causes the exchange mechanism to take hold to such an extent. This policy, useful for other reasons as well, is not unrealistic as long as some congestion is acceptable. It does not imply the elimination of all crowding. Instead, it would be optimal to expand a court up to the point where the social costs of enlargement are equal to their incremental social benefits. Relative to the cost of additional police deployment that would make a difference on crime, an expansion of judges, prosecutors, and their staff is an excellent social expenditure. Such enlarged court and prosecution resources would permit society to regain control over key policy issues in its criminal justice system.

NOTES

1. The literature on plea-bargaining is summarized in a special issue of Law and Society Review: 23, 2 (1979). See also the annotated bibliographies in National Institute of Justice (1980) and Markowitz (1978).

2. Similar to those transactions between defendants and prosecution, exchanges also exist between entire institutions of the criminal justice system, such as police and the prosecutor; they are described in Cole (1970).

3. The nonlinearity is similar to that of Kort (1968, 1977) and Nagel (1977). Court congestion includes all pending cases, not only those that are tried. Several jurisdictions, including the state of Alaska, have imposed partial bans on plea-bargaining. In that case, the curves in Figure 1 continue to exist, but they may be transposed vertically.
4. For those federal courts with more than 300 criminal trials per year, sentencing data were obtained from the Administrative Office of U.S. Courts (1973). Average sentences were regressed over measures of court congestion, defined as the ratio of inflow of cases to the number of judges. Control variables for district income, employment, and geographical region were included.

5. Instead of using a cross-sectional analysis for different jurisdictions, it is possible to use time-series analysis for the same court. The PROMIS data system can lend itself to this task once it has been in operation for a sufficient time period (Forst and Brosi, 1977), or, when it is used in more jurisdictions, for a cross-sectional analysis.

6. The data for the state of Alaska show an increase of sentences for minor offenses only after plea-bargaining was banned. However, Alaska's court system is not heavily congested.

7. See note 3.

8. This conclusion may be further tested as PROMIS data become available for more jurisdictions.

9. Unless the defense knows with certainty that admissible evidence could not be developed by the prosecution.

10. For one jurisdiction where this question was investigated empirically, the results, allowing for changes in the crime rate, show that a tripling of the courts is optimal (Noam, 1981).

REFERENCES


BORDENKIRCHER v. HAYNES (1978) 434 U.S. 357


ELI M. NOAM is Associate Professor of Business and Lecturer in Law at the Graduate School of Business of Columbia University.
INDEX

to

LAW & POLICY QUARTERLY

Volume 3

Number 1 (January 1981) pp. 1-120
Number 2 (April 1981) pp. 121-256
Number 3 (July 1981) pp. 257-376
Number 4 (October 1981) pp. 377-504

Authors:

CREW, ROBERT E., Jr., see Hanson, R. A.
HANNEKE, CHRISTINE R., see Shields, N. M.