JURY TRIAL VS. GUILTY PLEA:

A PROSECUTOR'S COST-BENEFIT COMPARISON

By

Eli M. Noam

Research Working Paper No. 318A

April 1980
JURY TRIAL VS. GUILTY PLEA: 
A PROSECUTOR'S COST-BENEFIT COMPARISON

Eli M. Noam

This paper compares the costs and the benefits of the two major forms of case disposition, jury trial and guilty plea. In most instances, a prosecutor can make a tactical choice between seeking to convict a defendant through a trial or attempting to obtain a guilty plea. Normally it is easier - and hence cheaper - to obtain a plea. But such a disposition may be less effective, because it is usually accompanied by a reduced sentence. On the other hand, guilty pleas lead to an assured conviction, whereas some trials end in acquittal. Furthermore, trials have a hidden cost: they tend to take up much of a court's time, and the resultant court congestion may have undesirable side effects on the treatment of other cases. All these factors must be taken into account, no analysis exists that deals with them in an integrated fashion.

A number of caveats should be stressed at the outset. The definition of rights that is used in this paper is a restricted one, limited to the value of reduced crime. Clearly there are important considerations beyond the cost of crime in determining criminal justice policy, and the greater cost-effectiveness of one form of disposition does not imply its superiority, in general. Furthermore, the discussion in this paper is based on a theoretical model of the court system; such models are, by necessity, highly simplified reflections of reality, and the results should be viewed with this in mind. And finally,

---

1 Assistant Professor of Business and Lecturer in Law, Columbia University. This paper was supported by the Graduate School of Business and by the Center for Law and Economic Studies, Columbia University.
the empirical investigation of the paper relies primarily on data from one jurisdiction; general conclusions may therefore be inappropriate.

The Model

It is assumed that the crime rate is influenced by the probability and by the severity of sanction. Of course, these are not the only factors that determine crime, since many social, cultural, and economic conditions contribute to it. But these are factors over which the prosecutor is assumed to have, while he can influence to some extent the probability and severity of punishment. Whether those deterrence factors exist is in dispute; however, while their existence is perhaps less clear in the case of "crimes of passion", it seems not unreasonable to expect them for "crimes of property". A professional burgular, on the margin, may well reduce his activity if its "price" increases.\(^2\) This can be expressed as the relation

\[
\frac{C}{N_p} = \alpha W^\beta \pi^\delta
\]  

(1)

Where \(C/N_p\) is per capita property crime, \(W\) is the average sentence, \(\pi\) is the probability of conviction, \(\beta\) and \(\delta\) are the elasticities with respect to severity and probability, and \(\alpha\) is a coefficient that accounts for the other factors contributing to crime.

As a first step of the analysis, we assume that a trial or plea disposition affects only the probability of conviction \(\pi\). Let \(\pi\) be defined as the percentage of convictions \(K\) relative to crimes \(C\) and let the sum of pleas \(P\) and trials \(T\) equal to the caseload \(\lambda\), and \(\phi\) be the probability of a trial resulting in a conviction. Then

\[
\pi = \frac{K}{C} = \frac{\phi T + P}{C} = \phi T + (\lambda - T)
\]  

(2)
\[ C = N \alpha W^\beta \pi^\delta \alpha \alpha W^\beta (\phi T + P)^\delta S = N \alpha W^\beta (\phi T + P)^\delta \frac{1}{1+\delta} \]  
(3)

Hence the marginal impact of a trial on crime through changes in probability is

\[ \frac{dC}{dT} = \frac{1}{1+\delta} N \alpha W^\beta \frac{1}{1+\delta} (\phi T + P)^\delta \frac{1+\delta}{1+\delta} - 1 \phi \]  
(4)

For a guilty plea, it is

\[ \frac{dC}{dP} = \frac{1}{1+\delta} N \alpha W^\beta S (\phi T + P)^\delta - 1 \]  
(5)

Let the average expenditure of a trial to prosecutor be \( E_T \) and that of a plea disposition \( E_P \), and let the social harm due to a crime be \( K \). The marginal effect of a budget dollar spent on a trial \( (\beta_T) \) on the total loss due to crime \( (7) \) is therefore,

\[ \frac{dZ}{d\beta_T} = \frac{dT}{d\beta_T} \cdot \frac{dC}{dT} \cdot \frac{dZ}{dC} = \frac{1}{E_T} \cdot \frac{\phi}{1+\delta} \frac{1}{(N \alpha W^\beta (1+\delta))} (\phi T + P)^{-1} - 1 \]  
(6)

For a plea bargain, it is similarly

\[ \frac{dZ}{d\beta_P} = \frac{1}{E_P} \frac{1}{1+\delta} (N \alpha W^\beta (1+\delta)) (\phi T + P)^{-1} \]  
(7)
So far, it was assumed that only the probabilities of conviction are changed by a trial or a plea. This assumption can now be relaxed and the impact of a disposition on the severity of conviction also be considered. The mechanism by which this can be explained is as follows. Most courts are characterized by acute congestion. A decision to go to trial has an effect on the degree of contestion and indirectly on the treatment of many other cases. Because it is impossible to hold trials for more than a small fraction of pending cases most others must be disposed of through the guilty plea process. In order to induce a defendant to give up his right to a jury trial and the chance for acquittal that goes with it, he normally receives a reward in the form of a sentence reduction over what he could have expected to receive after a jury trial conviction. This plea-bargaining has become essential for case disposition, and in some jurisdictions 95% of all cases are disposed of in that fashion. The sentence reduction that is given to defendants in exchange for a guilty plea is the price for a guilty plea. One can expect that the number of guilty pleas will be related to the attractiveness of the sentence concession. Therefore, in a congested court system where many cases must be plea-bargained, the sentence reduction can be expected to be higher than in a court without a heavy workload. Sentence "discount" levels at a rate that clears the docket will develop. Even with some individual variation in the granting of the discount, occurring expected discount levels are reached through the daily negotiations between prosecutors and defense attorneys. This functional relation of court dependency of sentence levels to court congestion can be expressed as

$$W = b \frac{M_i}{\lambda} V$$ (8)
where

\[ W = \text{average effective sentence} \]
\[ V = \text{average jury sentence} \]
\[ M = \text{number of judgeship equivalents} \]
\[ \mu = \text{trial capacity per judgeship} \]
\[ \lambda = \text{case load} \]
\[ b = \text{coefficient}. \]

\[ \frac{M\mu}{\lambda} \] is the ratio of trial capacity in the court to case load and a measure for court congestion. The smaller \( \frac{M\mu}{\lambda} \), the more congested the court, and the smaller is the effective sentence \( W \). The verification of this equation and an estimation of its magnitude are reported in the later part of the paper. The implications of the relationship between congestion and sentence severity are different for pleas than they are for trials. If trial capacity is fully utilized, \( M\mu = T \), and the decision for a trial is equivalent to reducing the available trial capacity for the remaining defendants by one trial unit, and similarly reducing the case load by one unit. For the fully utilized court,

\[ W = \frac{bM\mu V}{\lambda} = \frac{bM\mu V}{M\mu + P} \]

and the reduction, by one, of available trial capacity and of the remaining case load

\[ \frac{dW}{dT} = \frac{dW}{d(-M\mu)} = -bV \left( \frac{(M\mu + P) - M\mu}{(M\mu + P)^2} \right) = -\frac{bVP}{\lambda^2} \]  \( \text{(9)} \)
For a plea case, trial capacity is not affected, but remaining case load is reduced by a plea disposition. Hence, the court is less congested than before, and effective sentencing level is higher

\[
\frac{dW}{dp} = \frac{dW}{d(-\lambda)} = \frac{bVT}{\lambda^2} = \frac{bMuV}{\lambda^2} \tag{10}
\]

The effects of trials and pleas, when both the probability and the severity of conviction are considered, can now be calculated. Again, let the probability of convictions \( \pi \) be their number, \( K \) as a share of total crime \( \pi = \frac{K}{C} \), and

\[
C = Na^\delta W^\pi \delta = (Na^\delta W^\delta K) \tag{11}
\]

The effect of a trial is then

\[
\frac{dC}{dT} = \frac{1}{1+\delta} (Na^\delta W^\delta K)^{1+\delta} \frac{1}{NaS} (W^\delta W^{-1}dK + K^\delta W^\delta W^{-1}dW) \tag{11}
\]

which is

\[
\frac{dC}{dT} = C \cdot \frac{1}{1+\delta} \left( \frac{B}{T} \cdot \frac{\lambda-T}{\lambda} + \frac{\delta\phi}{K} \right) \tag{12}
\]

The marginal effect of a budget expenditure on a trial is then

\[
\frac{dZ}{dB_T} = \frac{1}{E_T} \cdot C \frac{1}{1+\delta} \left( -\frac{B}{T} \cdot \frac{\lambda-T}{\lambda} + \frac{\delta\phi}{K} \right) \tag{13}
\]
Similarly, for a guilty plea disposition it is

\[
\frac{dZ}{dP} = \frac{1}{E_P} \cdot C \frac{1}{1+\delta} (\frac{\beta}{\lambda} + \frac{\delta}{\lambda})^k
\]  

(14)

III. Empirical Results

The empirical part of this study is based on data from the District of Columbia; some of the parameters, however, must rely on nationwide investigations. The results should therefore be considered exemplary only. As crimes of property, five of the seven FBI index crimes are used. The existence and the magnitude of equation (8) \(W_i = \frac{b(Mu)_i V_i}{\lambda_1}\) was analyzed and shown to be \(W_i = (2.1557) \frac{(Mu)_i V_i}{\lambda_1}\) with an \(R^2\) of .8XXX, and significance at the .99 level.

\(k\), the social loss due to an average index crime, is estimated to be \(k=593\), using the Wolfgang-Sellin index of crime severity for each crime and converting severity into a money "equivalent" by the average number of trials per judge per year was, in the Federal Court System in 1972, 47 trials. Since this includes also civil cases, a consideration of the respective time averages results in an estimated criminal trial capacity \(\mu=59\) per judgeship per year. The number of judgeships \(M\) is, for the District of Columbia, calculated as \(M=15.7\) full-time equivalents of criminal judgeships, pro-rated from the 45 total judgeships (handling crime, family, traffic, etc.) on the basis of case frequencies and average time per case. \(\beta\) and \(\delta\), the elasticities of crime with respect to severity and probability, are derived from an estimation by I. Ehrlich; weighted averaged, for property crimes, are \(\beta = .6206\) and \(\delta = .7105\). In that study several additional demographical variables explain the crime rate, such as median income, percentage of poor, percentage of non-whites, etc. Values for these variables, taken from the 1970 census figures for
Washington, D.C.\textsuperscript{11} were substituted, and result in a parameter of $\alpha = 0.0294$. $\lambda$, the felony case inflow, is $\lambda = 3737$ for 1976.\textsuperscript{12} $N$, the population for Washington, D.C. is $N = 741,582$ in 1970.\textsuperscript{13} $V$, the average trial sentence severity, is $V = 17.45$ for the five index crimes.\textsuperscript{14} The probability of conviction by trial is found to be $\phi = 0.74$,\textsuperscript{15} the number of jury and bench trials $T = 920$ and of guilty plea dispositions $P = 2817$.\textsuperscript{16} The average cost per disposition to the prosecutor is estimated with the help of a time study for a different jurisdiction, which determined the time requirement for different prosecutorial activities.\textsuperscript{17} It is possible to calculate the average cost per time unit by dividing the prosecutor's total budget by the number of attorneys and their average workloads, and arrive at a total cost for a trial, including its preliminary states, of $E_T = 2538$, and for a guilty plea disposition, also including preliminary states, of $E_P = 284$.\textsuperscript{18}

These parameters are substituted into the equations of marginal benefits. The first results are those of equations (6-9) where only the effects of changes in probability are expressed. These results are

$$\frac{dC}{dT} = 8.74 \quad \frac{dC}{dP} = 11.80 \quad (,)$$

$$\frac{dZ}{dE_T} = 2.03 \quad \frac{dZ}{dE_P} = 24.57 \quad (,)$$

It can be seen that the impact of a guilty plea on crime is somewhat larger than that of a trial; when the relative costs of the two dispositions is taken into account, a plea is by far the more advantageous form of disposition. When both probability and severity effects are considered, as in equation (12) and (14), the results are
\[
\frac{dC}{dT} = -5.62 \quad \frac{dC}{dP} = -12.40
\]
\[
\frac{dZ}{dE_T} = 1.30 \quad \frac{dZ}{dE_p} = 25.80
\]

Again, a plea bargain disposition is more effective than a trial; the severity impact of a trial partly offsets that of probability while for a plea disposition both severity and probability effects are negative. When the relative cost of the dispositions is considered, the high cost-effectiveness of a plea is striking in comparison to that of a trial. Thus, the results show that the impact of a dollar spent on a guilty plea seem to be far higher than for a trial. Given limited resources, and a congested court, a plea bargain is here the obvious strategy when the reduction of crime is the prosecutor's primary goal. Of course, it has been long clear that a plea is the cheaper method of disposition. But given the more complex interactions of probabilities and severities for these dispositions, their relative effectiveness was not obvious. The results of this paper suggest that in Washington, D.C. a market cost-benefit advantage for plea bargains exist, confirming the rationality of its prosecutor's reliance on them. Therefore, when the availability of plea bargaining is reduced, substantial additional resources need to be provided in order to maintain the previous level of deterrence.

It should be stressed again, however, that these results, based as they are on data from Washington, D.C., need not necessarily apply to other jurisdictions as well. To generalize the conclusions this would be the next step of research.

Columbia University