A COST-BENEFIT MODEL
OF CRIMINAL COURTS

By

Eli M. Noam*

Revised Edition
July 1980

*Associate Professor of Business and Lecturer in Law, Columbia University. This paper was supported by a research grant by Columbia University, Graduate School of Business. Special thanks go to James G. McCafferty of the Administrative Office of the United States Courts, to Michael Hindelang of the Criminal Justice Research Center in Albany, and to Marsha Carey of the Federal Judicial Center. All responsibility remains with the author. The comments of a referree were helpful in clarifying the model.
A Cost-Benefit Model of the Criminal Court System

This paper describes a method of measuring the impact of courts on the reduction of crime. It raises the question as to the marginal effect of resources in the criminal courts. This issue is important because there are at present many well-meant calls for additional public funds for courts. Yet other parts of the criminal justice system such as corrections or police also have serious resource problems, to say nothing of the many other worthy courses in need of support. Hence, the case for additional funds for courts cannot be based simply on the notion that they are "important" in some abstract sense, because there are many other important institutions, but it requires some evidence, among which must be a showing that additional funds are likely to make a substantial difference.

One form of such evidence is a cost-benefit analysis, in which inputs are related to outputs. The problem in applying such a method to the courts is that it is difficult to measure the effects of budget changes and hard to value the "output" of courts. This study is an attempt to overcome some of these problems, at least conceptually, in the framework of a simplified model of court operations.

Criminal courts are concerned with, among other objectives, the enforcement of law. In the aggregate, the courts' operation effects crime,
directly or indirectly. Therefore, the study will concentrate on the law enforcement aspect of court de-congestion, and will trace their benefits in terms of crime reduction.

The paper proceeds as follows. First, a model is specified. This is followed by an empirical application for the District of Columbia. The results suggest that there are substantial benefits to additional resources in the criminal court system. These benefits are in the order of $5-11 dollars for each dollar of additional court resources.

The Model

With more resources available to the court, additional judges, clerks, and other personnel could be hired, and therefore cases can be tried at a higher rate. But what is the significance of this higher rate of case processing on crime? This question can be approached by looking at the effects of heavy caseloads. The courts in which the bulk of criminal cases are processed in this country are urban courts, usually characterized by a severe congestion. It is not possible to hold trials for more than a small fraction of pending cases, and most others must be disposed of in a different fashion. This happens primarily 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 is normally promised as a reward a reduced sentence. This process is essential for case disposition; in some jurisdictions 95% of all cases are plea bargained. The sentence reduction that is granted in exchange for a plea is the
"price" for a guilty plea. It is not granted voluntarily but rather by necessity, since without it the court, inundated by a huge caseload, would not be able to function. It is elementary that the magnitude of a price affects the quantity of transactions. Thus it is reasonable to expect that the number of guilty pleas is related to the attractiveness of the sentence reduction.\(^1\) Hence in a congested court system, where the prosecutor must dispose of relatively many cases by plea-bargaining, he must offer relatively attractive terms. Thus the sentence reduction will here be higher than in a court with a low workload, and this reduction is independent of the relative merit of the cases. But a prosecutor does not offer better deals than he has to; hence, sentence reductions at a level that "clears the docket" come about and some average sentencing levels develops through the daily negotiations between prosecutors and defense attorneys. Indeed, the term "the going rate" is frequently used in plea negotiations.\(^2\) Functionally, the relationship between sentence reduction and court congestion is hypothesized, subject to later empirical verification, to be a linear relationship.

\[
W = b \frac{M \mu}{\lambda} \nu
\]

where

\[
\begin{align*}
V &= \text{average jury sentence} \\
W &= \text{the average sentence} \\
M &= \text{number of judgeships} \\
\mu &= \text{trial capacity per judgeship} \\
\lambda &= \text{case inflow} \\
b &= \text{coefficient}
\end{align*}
\]
\( \frac{M_u}{\lambda} \) is the ratio of total trial capacity in the court to case inflow, and is a measure for the congestion in the court. The smaller this ratio, the more congested the court, and the lower the effective sentence \( W \) in relation to a given \( V \).

A verification of the existence of this relationship and an estimation of the magnitude of the parameter \( b \) are given in the empirical part of the paper.

A second equation of the model relates crime and criminal sentences. An existing crime rate is the product of many cultural, social and economic conditions; these factors are not, however, changed by adding resources to the criminal courts, and for the purposes of this analysis they are assumed to be given. What is variable by courts is the "price" which a defendant must potentially face for a crime, i.e., probability and magnitude of punishment for an offense. Such expected punishment, one would assume, must have some effect on crime rates. While there is disagreement about this point, particularly on the effect of punishment on so-called crimes of passion or violence it is reasonable to believe that some rational decision-making occurs in the commission of "crimes of property". A professional burgler may well reduce this activity if its "price" increases. Not every burgler will be prevented, of course, but some marginal ones may not be committed. The existence of deterrence in such circumstances, while not conclusively settled,\(^3\) has received growing empirical support.\(^4\) This effect on probability and severity of conviction on the crime rate can be expressed by the equation

\[
\frac{C}{P} = a W^b \pi \gamma
\] (2)
where \( \frac{C}{P} \) is per capita crime, \( W \) is the average sentence, \( \pi \) is the probability of a case resulting in conviction, and \( \alpha, \beta, \) and \( \gamma \) are coefficients.\(^5\)

For the present it is assumed that the case flow into the court from the police is the stable rate \( \lambda \) (though this assumption will be related later); that the court tries as many cases as its limited capacity \( M \) permits; and that the rest of the cases are plea-bargained. We also assume a probability of conviction, given trial, which is a constant \( \rho \); this is not unreasonable, considering that juries in their verdict are not likely to be influenced by court congestion. For guilty plea-cases the probability of conviction is certain by definition. Hence the average probability of conviction for a case that enters the court system is the rate of convictions divided by the case flow,

\[
\pi = \frac{\rho M \mu + (\lambda - M \mu)}{\lambda} = 1 - \frac{M \mu \eta}{\lambda}
\]

where \( \eta = 1 - \rho \) is the probability of acquittal in a trial.

The changes in \( \pi \) with respect to an increase in judgeships is therefore

\[
\frac{\delta \pi}{\delta M} = - \frac{\mu \eta}{\lambda}
\]

There are two further equations in the model. First, the relation of budget to court capacity is given by a simple budget function

\[
M = \frac{B - F}{E}
\]

with \( B \) the total court budget, \( F \) the fixed overhead, \( M \) the number of judgeships, and \( E \) the average cost of each. Secondly, the social cost of crime is defined as
with $C$ the crime rate, $r$ an adjustment factor for unreported crime, and $k$ the average direct social loss due to a crime. Equation (5) becomes, after substitutions

$$Z = Crk = Prw^\beta \pi^r k = \alpha\left(\frac{bMwV}{\lambda}\right)^\beta (1 - \frac{Mun}{\lambda})^\gamma r k$$

(6)

The marginal effect of a court budget allocation on social losses due to crime is then the derivative

$$\frac{dZ}{dM} = rk\alpha\left(\frac{bMwV}{\lambda}\right)^\beta [\beta M^{\beta-1} (1 - \frac{Mun}{\lambda})^\gamma + M^\beta \gamma(1 - \frac{Mun}{\lambda})^{-1}\frac{\gamma}{\lambda}]$$

(6)

$$= rk C \left[\frac{\beta}{M} + \gamma\left(\frac{-Mun}{\lambda}\right)/(1 - \frac{Mun}{\lambda})\right]$$

(6)

The marginal effect of a budget allocation in terms of crime reduction $\frac{dZ}{dB}$ is then by the chain rule

$$\frac{dZ}{dB} = \frac{dM}{dB} \frac{dZ}{dM}$$

With $\frac{dM}{dB} = \frac{1}{E}$, we obtain as a result for the marginal effect of the court budget

$$\frac{dZ}{dB} = \frac{1}{E} rk C \left[\frac{\beta}{M} + \gamma\left(\frac{-Mun}{\lambda}\right)/(1 - \frac{Mun}{\lambda})\right]$$

(7)

The model can now be carried a step further. So far, the case load $\lambda$ has been assumed as given, but this assumption is now relaxed, with $\lambda$ a variable and a function of the crime rate
\[ \lambda = f(C) = t \cdot C \] (8)

The introduction of this relation changes the model into a system of simultaneous equations since crime is also a function of case load, \( C = h(\lambda) \)

Equation (2) for crime becomes, after substitutions of (1), (3) and (8).

\[ C = P^x \left( \frac{bMuV}{tC} \right)^\beta \left( 1 - \frac{Mun}{tC} \right) \] (9)

which results in

\[ \frac{1}{(Pa)^{\frac{1}{Y}}} \left( \frac{bMuV}{t} \right)^{\frac{B}{Y}} = D = \frac{M^{\frac{B}{Y}}}{B+1} - \frac{(\frac{\mu n}{t})^{\frac{B+Y}{Y}}}{C^{\frac{B+Y}{Y}}} \] (10)

where \( D \) is a short band parameter for the left hand side.

Taking the total derivative we have

\[ dB = \frac{\partial D}{\partial x} dx + \frac{\partial D}{\partial y} dy = 0 \] (11)

so that

\[ \frac{dC}{dM} = \frac{\partial D}{\partial x} / \frac{\partial D}{\partial C} \]

The partial derivatives are

\[ \frac{\partial D}{\partial M} = \frac{B-Y}{YM} - \frac{\mu n(\beta+\gamma)}{t \gamma C} \] (12)

\[ \frac{\partial D}{\partial C} = -\frac{(B+1)M^{\beta/Y}}{B+Y+1} + \frac{\mu n(\beta+\gamma+1)M^{\gamma}}{t \gamma C} \] (13)
and we can solve after some algebra

\[
d\mathbf{C} = \frac{\mathbf{B} \mathbf{t} \mathbf{C}^2 - \mathbf{M} \mathbf{u} \mathbf{n}(\mathbf{B}+\mathbf{Y}) \mathbf{C}}{\mathbf{t} \mathbf{C}(\mathbf{B}+1) \mathbf{M} - \mathbf{u} \mathbf{n}(\mathbf{B}+\mathbf{Y}+1) \mathbf{M}^2}
\]

so that

\[
d\mathbf{Z} = \mathbf{r} \mathbf{K} \mathbf{B} \mathbf{t} \mathbf{C}^2 - \mathbf{M} \mathbf{u} \mathbf{n}(\mathbf{B}+\mathbf{Y}) \mathbf{C} \cdot \mathbf{E} \cdot \mathbf{t} \mathbf{C}(\mathbf{B}+1) \mathbf{M} - \mathbf{u} \mathbf{n}(\mathbf{B}+\mathbf{Y}+1) \mathbf{M}^2
\]

This equation is the expression, for the simultaneous equation model, of the marginal effects of resources in the criminal courts.

**Empirical Analysis and Data**

The empirical investigation is based on data from the District of Columbia, where available; the crimes considered are the four FBI "index" property crimes. As a first step of the empirical estimation the existence of the relationship between sentence reduction and court congestion is established by estimating equation (1).

\[
\mathbf{W}_i = \frac{\mathbf{bM}_i \mathbf{uV}_i}{\lambda_i}
\]

The existence and the magnitude of this relationship was analyzed by estimating a regression across the federal district courts with heavy criminal workloads. Data for all federal criminal cases for the year 1973 was used which contains the sentences for each case, reported according to a sentence severity index used by the Administrative Office of United State Courts. Thus the average sentences \( V_i \) (for jury dispositions for the four property crimes) and \( W_i \) (for overall average dispositions for the four property crimes) was calculated for each district from extensive data which includes every disposition of a criminal case in the federal system.
the ratio of trials to the total case-inflow, is available from court administration statistics. Population figures $P_i$ for the judicial districts are from congressional hearing reports and from census figures for the countries that comprise a judicial district. With this data, equation (1) was estimated, using ordinary least squares, and found to be

$$W_i = 2.11 \left(\frac{(M_i V_i)}{\lambda_i}\right)$$

The $t$-value for the coefficient $b$ is significant at the .99 level. The $R^2$ is a high .9032. It is possible to raise the $R^2$ slightly further by the use of additional demographic variables, but the explanatory power of congestion alone is very high already. Further variables, besides not being part of the model, would add complexity for its own sake. In sum, strong evidence exists for the hypothesis that sentences in a court are directly related to the degree of congestion in that court. Of the other parameters, $k$, the average loss due to a property crime, is of particular interest. One possible definition for $k$ is the direct economic loss due to a crime. Average figures for that loss were found by a Presidential Commission; their weighted average, for the four property crimes, is $k_1 = 402$ in 1980 dollars. It should be noted that the commission did not assign any positive value to the welfare that accrues to the offender from his crime. To consider such benefits is to treat a property crime as a form of transfer payment; crime could then even have a positive net social benefit, if the fruits of burglary or robbery are more valuable to the criminal than to the victim. Because of the
conclusions that would flow from such observations, the benefit to criminals is not counted by the commission, and the present study, being in agreement with this assumption, follows this procedure, through it refrains from using the term "social" loss and speaks instead of "loss to victims." In addition, it also uses an alternative measure for the damage of crime which includes more than the material losses to victims. The disutility to the victim of a crime goes usually far beyond the economic cost of the lost object. But how can such disutility be estimated? It is a subjective concept and not easily measure, but in the case of crime, previous research can provide some help. In a classic study population samples were surveyed to find the public perception of the "seriousness" of different crimes. The results are the well-known Wolfgang-Sellin index of crime severity. In the same study, the seriousness that is attaches to a non-violent, non-disruptive, "taking" of different sums of money is also surveyed, and a "Power Function of Money" is estimated from the results. This function describes perceptions of seriousness according to the amounts of money that is lost: Severity perception = 16.93 (money-loss). If this function is now inverted one can express, for each offense for which a severity index value is known, a money "equivalent", so that

\[ \text{Money loss equivalent} = \left( \frac{\text{offense severity index number}}{16.93} \right) \]

By substitution the seriousness index figures for the four major property crimes the associated dollar amounts can be found; their weighted average is \( k_2 = 641 \), in 1980 dollars.
The remaining parameters are obtained as follows: $r$, the adjustment factor for unreported crime, was taken from a recent victimization survey in Washington, D.C.\textsuperscript{15} in which thousands of businesses and households were questioned. The weighted average is $r = 1.87$. $E$, the cost of an additional judgeship, was calculated by the Administrative Office of U.S. Courts, and is in 1980 dollars $E = $220,000\textsuperscript{16}. $\beta$ and $\delta$, the elasticities of crime with respect to severity and probability, are derived from an estimation of 1. Ehrlich.\textsuperscript{17} Weighted averages for property crimes are $\beta = -.6206$ and $\gamma = -.7105$. In that study, several demographic variables explain the crime rate, such as median income, percentage of poor, or percentage of non-whites. These variables, taken from 1970 census figures for Washington, D.C.\textsuperscript{18} were substituted and result in parameters of $\alpha = .25$. $t$, the ratio of cases $\lambda$ to the crime rate $C$ has been on the average during the past five years $t = .67$.\textsuperscript{19} $\lambda$, the felony case inflow is $\lambda = 3762$ for 1976.\textsuperscript{20} The population for Washington, D.C. is $P = 741.582$\textsuperscript{21} from the 1970 census. $\eta$, the probability of acquittal given a trial, is for 1970-1975 $\eta = .14$.\textsuperscript{22} The average number of trials per judge per year was, in the Federal Court System, $\mu = 47$.\textsuperscript{23} Since this also includes civil cases, a consideration of the respective time averages results in an estimated criminal trial capacity $\mu = 59$ per judgeship per year. Finally, $M$ is the number of judgeships in the District of Columbia. In 1976 a total of 44 trial level judgeships existed in Washington\textsuperscript{25}, handling criminal as well as civil and other cases, and 40% of all cases in the court system were criminal.\textsuperscript{26} Assuming that trials are allocated in a similar proportion and adjusting for the fact that civil cases take on the average longer,\textsuperscript{27} criminal judgeships comprise roughly the equivalent of $M = 15.6$ full-time judgeships.
Results

Results can be found by substituting the parameters into equations (7) and (15). When case flow $\lambda$ is assumed to be given, the marginal effect of added resources in the courts is found to be

$$\left( \frac{dZ}{db} \right)_1 = \$9.33$$ (7)

When the case flow is a function of the crime rate, the marginal effect of courts is found, from the simultaneous model (15)

$$\left( \frac{dZ}{db} \right)_2 = \$11.33$$ (15)

When material losses, as defined by the President's Commission, are substituted for our disutility values, the results are

$$\left( \frac{dZ}{db} \right)_3 = \$5.85 \quad \text{and} \quad \left( \frac{dZ}{db} \right)_4 = \$7.10$$

These are all high benefit-cost ratios, indicating that additional resources for criminal courts can have a significant impact in reducing the losses due to crime. Comparing these figures with the much lower ratios found for many other government expenditures, a strong case can be made for increased budget allocations to the courts.

Conclusion

The paper has defined a model to determine the benefits— in terms of value of reduced crime to victims— of additional budget allocations for the criminal court system. The model required an empirical analysis of the
impact of court congestion on sentencing, and an estimation of that effect. It was then possible to calculate the cost-benefit ratio of marginal court budgets. The results show very high returns of 5-11 dollars of crime reduction per dollar of additional budget allocation.

It must be stressed that the court benefits that were calculated are only those of crime reduction of four major property crimes. Additional court resources may also improve the quality of justice and the speed of its disposition. Thus the values that were found are only a lower boundary of overall benefits. On the other hand, the results assume a congested criminal court. Where no congestion exists, the cost-benefit ratios are different and smaller. Clearly, there are considerations beyond cost effectiveness where courts are concerned, and many other arguments in favor of court expansion can be made. The results of this study suggest, however, that even in terms of crime reduction, additional budget allocations for the criminal courts are an effective and sensible social policy.
FOOTNOTES


2. See, for example, the dialogue in James Mills, One Just Man, 1975.


5. α includes the impact of other factors which contribute to crime, such as poverty, or which deter from it, such as police, and which are assumed to be given.

6. Defined as those Courts with 300 or more criminal trials per year; the Middle District of Florida was omitted because of problems with its data.


8. Administrative Office of U.S. Courts, James A. McCafferty, Chief, Statistical Analysis and Reports Branch. Part of the computation of the data was done at the Criminal Justice Research Center in Albany.


11. Milton Heumann's claim that Connecticut guilty pleas are unaffected by case pressures ("A Note on Plea Bargaining And Case Pressure", Law and Society Review, Vol. 9, 1975, p. 515) is based on relative frequency and overlooks changing prices (sentence severities); lacking a theoretical model, his findings can be reconciled with several contradictory hypotheses, among them our equation (1).

12. President's Commission of Law Enforcement and the Administration of Justice, Assessment Task Force Report, p.42. Values for Aggravated Assault were extrapolated using the Wolfgang-Sellin index.


14. Ibid. p. 289. Average characteristics for larceny and robbery were chosen.


21. 1976 Federal Census, Population Figure Tables.


26. Ibid., p. 45.

27. Federal Judicial Center, op. cit., p. 45.