

Published in *Journal of Institutional and Theoretical Economics* **151** (1995), 326–347. Copyright © 1995 by J. C. B. Mohr (Paul Siebeck) Tübingen.

Crime, Punishment, and Social Expenditure*

Jean-Pierre Benoît

Martin J. Osborne

Department of Economics
and School of Law
New York University[†]

Department of Economics
McMaster University[‡]

Revised October 1994

Abstract

Criminal activity can be controlled by punishment, and by social expenditure both on enforcement and redistributive transfers which increase the opportunity cost of imprisonment. Individuals may differ in the combinations of these policies that they prefer. We study the dependence of these preferences on the individuals' characteristics and the nature of the crime under consideration. A political mechanism determines the policy adopted by society. Differences in the policies adopted across societies are explained by the nature of the political mechanism and the initial distribution and level of incomes.

JEL classifications: K42, H53, D72, D78.

*Aloysius Siow inspired our interest in this topic. We are very grateful to him for many discussions. We are also grateful for comments from Orley Ashenfelter, Ted Bergstrom, Dan Bernhardt, Martin Browning, Curtis Eaton, William Furlong, David Greenberg, Lewis Kornhauser, Peter Kuhn, Peter McCabe, Claude Montmarquette, Nancy Morawetz, Burt Neuborne, Carolyn Pitchik, Richard Revesz, Diana Simon, Dan Usher, Mike Veall, Junsen Zhang, and two anonymous referees. Benoît's work was supported in part by grants from the School of Business at Columbia University and the C. V. Starr Center for Applied Economics at New York University. Osborne's work was supported in part by grants from the National Science Foundation Grant (SES-8510800), the Social Sciences and Humanities Research Council of Canada, and the Natural Sciences and Engineering Research Council of Canada. This paper was circulated previously under the title "Crime, Punishment, and the Redistribution of Wealth".

[†]Address: benoitj@fasecon.econ.nyu.edu, or Department of Economics, New York University, 269 Mercer Street, New York, NY 10003, USA.

[‡]Address: osborne@mcmaster.ca, or Department of Economics, McMaster University, Hamilton, Canada L8S 4M4

1. Introduction

“On 2 March 1757 Damiens the regicide was condemned . . . to be ‘taken [to] . . . a scaffold . . . , [where] the flesh will be torn from his breasts, arms, thighs and calves with red hot pincers, his right hand, holding the knife with which he committed the said parricide, burnt with sulphur, and, on those places where the flesh will be torn away, poured molten lead, boiling oil, burning resin, wax and sulphur melted together and then his body drawn and quartered by four horses and his limbs and body consumed by fire, reduced to ashes and his ashes thrown to the winds’” (FOUCAULT [1979, 1]; quotation from original document).

The severity of penalties imposed for crimes has shown substantial variation over time. Offenses that in eighteenth century Europe were punished by torture and death now carry only prison sentences. Similarly, the penalties imposed for crimes vary significantly between countries. What factors account for these differences?

BECKER [1968] analyzes a model in which an individual decides whether or not to engage in criminal activity by comparing the risks and rewards of criminal and non-criminal behavior. We use Becker’s framework as a starting point to analyze society’s choice of a crime–control policy.

One way of making crime less attractive is to punish criminal actions, say through incarceration. While it is costly to maintain prisons it is not clear that more severe punishment is more costly. Indeed, more Spartan jails are *less* costly to maintain. Since there is no necessary connection between the harshness of a punishment and its direct cost, we assume that the latter does not increase with the severity of punishment.¹ Other methods of making crime less attractive necessarily entail a direct resource cost. Expenditure on enforcement can be increased, for example, or re-training programs can be introduced. Less obviously, programs that involve redistributing income from rich to poor (at a direct resource cost to the rich) may reduce crime: those who commit crimes in Becker’s model are, *ceteris paribus*, those with the lowest market incomes, and increasing their return from legal activities may be an effective crime–control policy.

We study a model in which society has two instruments to control crime: the severity of punishment, changes in which entail no direct resource cost, and

¹In the short run, institutional constraints within a society may be such that punishment can be made more severe only by increasing the length of sentences and hence increasing expenditures on jails. In the long run and across societies, this is not true. Indeed, the harshness of punishment in some countries stems precisely from the miserable jail conditions.

social expenditure, that is, any crime-reducing policy with a direct resource cost, including expenditures on police and redistributive transfers.

If changes in the level of punishment were to entail no change in cost whatsoever, then society would choose punishments as severely as possible.² There is a significant indirect variable cost, however: since the legal system cannot be perfect, with positive probability innocent people will be punished. Taking into account this probability of mistaken punishment gives us an important explanatory variable.

In our model each person in society favors a different punishment-expenditure scheme. Denote the level of punishment utility by v , and index possible expenditure levels by the parameter α .

Each individual has some preferred *policy* (α, v) . Our theory attributes differences in individuals' desired policies to differences in their market incomes, the extent to which they are protected from crime, their probability of being mistakenly punished, and the nature of the admissible tax schemes.

Variations between *societies* are attributed to differences in the level and distribution of wealth, differences in the technologies for perpetrating crime and for apprehending criminals, and differences in the natures of the political systems used to aggregate individual preferences into a policy for society. For example, we show that as the general level of income rises the punishment adopted becomes less severe, whereas if the political mechanism gives disproportionate weight to a wealthy elite, punishment is relatively severe. (A verbal summary of some of our results is given in Table 1 at the end of the paper.)

The fact that there are two instruments available for controlling crime yields insights that cannot be obtained in a standard "deterrence-type" model, which considers only punishment. For instance, with only sanctions available, a desire to reduce the crime rate entails a harshening of punishments. With other policies available, this need not be true. In fact, we find that depending on how crime affects different income groups, a reduction in criminal activity may actually be obtained through less severe punishments and greater expenditure. The intuition for this result is conveyed by considering the example of robbery on the subway. Increased expenditure may be a desirable means by which to reduce crime in this case since the reduction in crime that it engenders makes people less concerned about not being able to afford other means of transport. If this effect is strong enough then it may be desirable to save on the (indirect) cost of punishment by making punishment more lenient. In

²We focus on crimes for which the net benefit to society is never positive. If in some cases a "crime" yields a positive net benefit then there may be a reason to limit the penalty, in order that the "optimal" amount of crime be committed. See POLINSKY and SHAVELL [1979, 1984] for analyses of this case.

a similar vein, while those with a lower probability of being falsely convicted of a crime favor more severe punishments, we find that their aim may not be a lower overall crime rate.

Relation to the Literature

BECKER [1968] limits penalties by assuming that there may be a social gain from the commission of an offense. STIGLER [1970] criticizes this assumption of unexplained social gain, and proposes that enforcement is costly because innocent persons are sometimes convicted. EHRLICH [1973] expands on Becker's model; he notes that income inequality is an explanatory variable in that framework. In a paper more directly related to ours, HARRIS [1970] incorporates a positive probability of wrongful conviction into Becker's model, and addresses some of the same issues as we do. However, his focus is different; in his model there is no room for differences of opinion on the best policy, for example, so that the nature of the political mechanism is not an explanatory variable.

These authors follow Becker in working with a "reduced form" model, in which the starting point is a given function that measures the social loss to crime; society minimizes this loss by choosing penalties and enforcement levels. This model is not well-suited to analyze the influence of the original distribution of income or the nature of the political mechanism on the policy adopted.

Our work also draws less directly on the work of AUMANN and KURZ [1977] and BECKER [1983], who analyze positive models of wealth redistribution. Only EATON and WHITE [1991] and SCHOTTER [1985, 96–98], as far as we are aware, directly consider the policy of redistributing wealth to reduce crime. Schotter informally discusses the issue from the viewpoint of social justice. Eaton and White study some of the issues that we do. In their model, which is quite different from ours, there are two individuals, and the possibility of theft tomorrow reduces the incentive for time-consuming productive activities today.

2. The Model

Society consists of a large number of individuals, each of whom is concerned about her own after-tax income y and the crime rate c . Each individual's preferences are the same, and are represented by the von Neumann–Morgenstern utility function $u(y, c)$, which is increasing and concave in y and decreasing in

c. The concavity of u in y means that individuals are risk-averse when comparing lotteries with the same crime rate but different (after-tax) incomes.

In Section 6 (which some readers may prefer to read before proceeding) we derive u in a model in which the primitive is a utility function w defined over income alone, and crime causes each individual to lose a certain amount of income. In this model the standard assumption of concavity of w is not sufficient to guarantee that u is concave in y . However, the additional conditions that are needed do not appear to be unreasonable.

In order to control the crime rate, the members of the society collectively maintain a mechanism to catch and punish offenders. We assume that punishment takes the form of a jail sentence, and index it by the utility v that each individual experiences in jail. (We do not address the reason why punishment takes this form.)³

Jails are costly to maintain, but, as discussed in the introduction, the relation between the cost and the severity of the punishment is unclear. Thus we assume that changes in the level punishment entail no direct cost. (Note however that we *do* allow for some fixed cost associated with operating the punishment system).

The society has another instrument at its disposal: the level of expenditure on policies, like enforcement, that have a direct resource cost. The redistribution of income may also be such a policy, as the following argument shows. Given any punishment, increasing an individual's disposable income reduces the difference between the return to market activity and the expected return to crime (since the latter includes a positive probability of being subjected to the punishment). Thus, holding other characteristics fixed, low-income individuals are more likely to be criminals, so that increasing the disposable income of the poorer members of the society decreases the crime rate. However, to the extent that such individuals vary in their degree of risk-aversion, talent for crime, and ethics (as reflected in the way their utility function evaluates illicit gains), some of the money that is redistributed to the poor will be "wasted" on people who would not have been criminals anyhow. This wastage can be minimized by identifying non-income characteristics that are correlated with a propensity for crime. One such characteristic might be a past history of legal difficulties, whether this be actual time spent in jail or simply "trouble with the law". Special transfers can be directed to people with these histories. Of course, a danger in such transfers is that they provide an incentive for

³For simplicity, we assume that an individual receives either her punishment utility or her market income, but not both. This is consistent with an atemporal model, a model in which punished individuals have restricted incomes following their release from prison, and one in which the severity of a prison term, but not its length, is varied.

people to engage in the activity they are meant to discourage. Nonetheless, one would expect a non-zero optimal amount of these transfers. We observe such transfers in the form of job training programs for convicts and out-reach programs for “troubled” youths—programs that are often explicitly motivated by a desire to induce non-criminal behavior.⁴

Thus we assume that the society has two independent instruments at its disposal to control the crime rate: the severity of punishment, and the level of expenditure. The fact that the former involves no direct cost confronts us with the question of why the society does not impose punishments as high as possible. We argue that at least part of the answer is that in any legal system, mistakes are inevitable: with positive probability, innocent persons are punished. Thus punishment involves an *indirect* cost. We assume that the probability q_i that individual i is punished even though she is innocent may vary between individuals—perhaps richer individuals have the resources to better establish their innocence.⁵

We assume that the (direct) cost of crime-control is shared between the members of the society. First consider how the money to be spent is raised. Admissible tax schemes are indexed by a parameter $\alpha \in (0, 1)$ in such a way that schemes that raise more revenue have higher values of α . In the tax schedule indexed by α an individual with market income x pays the tax $T(x, \alpha)$. For simplicity we assume that $T(x, \alpha) = \alpha\tau(x)$: all admissible tax schedules are “scaled up” versions of some basic schedule τ . We do not assume a specific form for τ ; it could, for example, be linear, or quadratic. We assume only that the marginal tax rate is always between 0 and 1:

$$0 \leq \alpha\tau'(x) \leq 1.$$

When the money that is raised by a tax scheme is spent, the resulting crime rate is $C(\alpha, v)$. (In Section 7, which may be read independently of the rest of the paper, we specify a model in which individuals rationally decide whether or not to commit crime, given the relative risks and rewards to these activities, and derive $C(\alpha, v)$ in the case that expenditure is made on redistribution.) We assume that an increase in expenditure and an increase in the severity of

⁴We thank Richard Revesz for bringing this point to our attention.

⁵We treat these probabilities as exogenous. In a more elaborate model the overall level of q_i 's would be endogenous. Thus, for instance, more stringent evidentiary requirements could be adopted, reducing all q_i 's and increasing the crime rate (due to the increased difficulty of convicting a criminal). We stay with the simpler assumption of fixed q_i 's as this already allows for the consideration of more than one decision variable and enables us to focus on the issue of expenditure versus punishment.

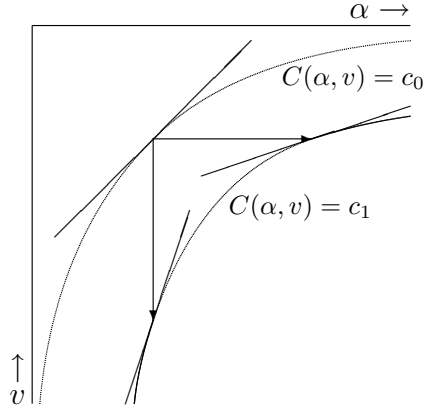


Figure 1. Contours of the function C . We assume that as α increases for fixed v the slope of the contours decreases, and as v decreases for fixed α this slope increases.

punishment (reduction in v) reduce the crime rate:⁶

$$D_\alpha C(\alpha, v) \leq 0 \quad \text{and} \quad D_v C(\alpha, v) \geq 0.$$

Further, we assume that as expenditure increases (given punishment) it becomes less effective relative to punishment in reducing crime, and as punishment becomes more severe (given expenditure) it becomes less effective relative to expenditure. These assumptions are illustrated in Figure 1: for each given value of v the slopes of the contours of C decrease as α increases, and for each given value of α these slopes increase as v decreases. Mathematically the assumptions are equivalent to

$$D_v C \cdot D_{\alpha\alpha} C - D_\alpha C \cdot D_{\alpha v} C > 0 \quad \text{and} \quad D_\alpha C \cdot D_{vv} C - D_v C \cdot D_{\alpha v} C < 0. \quad (1)$$

(These assumptions imply that C is quasi-convex, as in Figure 1.)

We are interested in the willingness of noncriminals to pay taxes to control crime and do not consider the policies favored by criminals or by those who receive subsidies. Presumably a criminal would want low penalties (except to the extent that she herself fears being the victim of crime and favors having others work), and those who receive subsidies would favor massive redistribution. If individuals' fortunes were subject to random variation then a person's preferences over policies might take into account how she would feel in the event she were a criminal; even in the absence of such variation the policy

⁶We use D for the derivative operator; for example, we write $D_\alpha C(\alpha, v)$ for $(\partial C / \partial \alpha)(\alpha, v)$ and $D_{\alpha v} C(\alpha, v)$ for $(\partial^2 C / \partial \alpha \partial v)(\alpha, v)$.

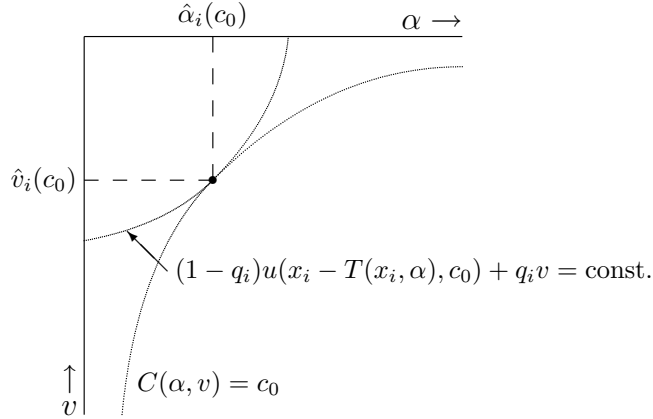


Figure 2. An illustration of the solution of (2). Note that for each value of c there is a different family of indifference curves of $(1 - q_i)u(x_i - T(x_i, \alpha), c) + q_i v$.

pursued by a society might be influenced by criminals' policy preferences. For our purposes it seems preferable to avoid these complications by simply assuming that crime policy is determined by noncriminals. If i is such an individual, then her most preferred way of achieving any given crime rate c is the scheme $(\hat{\alpha}_i(c), \hat{v}_i(c))$ that solves the problem

$$\max_{(\alpha, v)} (1 - q_i)u(x_i - T(x_i, \alpha), c) + q_i v \quad \text{subject to} \quad C(\alpha, v) = c. \quad (2)$$

This problem is illustrated in Figure 2 (note that c is a parameter of the problem). An analysis of $(\hat{\alpha}_i(c), \hat{v}_i(c))$ yields predictions about the dissension within an economy concerning the crime-control policy that should be adopted to achieve any given crime rate. (Note that $\hat{\alpha}_i$ and \hat{v}_i depend not only on c , as our notation records, but also on the parameters of the problem: q_i and x_i .)

To say something about the policy that the society adopts, we need to consider the political mechanism that aggregates the diverse preferences of the individuals. If individual i is a dictator then she adopts the policy (α_i^*, v_i^*) that solves

$$\max_{(\alpha, v)} (1 - q_i)u(x_i - T(x_i, \alpha), C(\alpha, v)) + q_i v. \quad (3)$$

The resulting crime rate is $c_i^* = C(\alpha_i^*, v_i^*)$, which can alternatively be obtained as the solution of the problem

$$\max_c (1 - q_i)u(x_i - T(x_i, \hat{\alpha}_i(c)), c) + q_i \hat{v}_i(c). \quad (4)$$

More generally, we assume that if no individual is a dictator then the policy the society adopts is a compromise among the most preferred policies of all

non-criminal individuals, as determined by a given political mechanism that gives nonnegative weight to each individual’s most preferred policy. We refer to an individual’s most preferred policy as the policy that she “proposes”.

One final note on our model. That people are sometimes mistakenly convicted of crimes is undeniable. That individuals’ fears of being unjustly imprisoned account for the limitations on punishment is less obvious. A closely related explanation would be that people see a cost to anyone being falsely convicted, whether or not it happens to be them. One might further conjecture that individuals of a given class find it more objectionable when “one of their own” is mistakenly convicted. The comparative statics we perform on changes in an individual’s chance of mistaken conviction would then be paralleled by comparative statics on the chance that someone in the individual’s class is falsely convicted. With greater generality, we could write the individual’s utility function as $u(y, v, c)$, where $D_v u > 0$ indicates some indirect cost to punishing. With appropriate assumptions many of our results would again emerge, but the explanatory power of the role of mistaken conviction would disappear.

Our analysis of the model proceeds as follows. We begin, in Section 3, by studying how an individual’s favored punishment–expenditure policy for achieving a given crime rate depends on that crime rate; we find that the favored policy depends on a characteristic of the individual’s utility function, namely the cross partial $D_{yc}u$, which we interpret. Then, in Section 4 we study how differences between individuals within a society in the probability of mistaken conviction and market incomes affect the policies that they favor. Our analysis concludes in Section 5, where we study differences between societies: we assume that the policy implemented in a society is some aggregate of the favorite policies of the members of the society, and consider how that implemented policy depends on the probability of mistaken conviction in the society, the nature of the political mechanism, the general level of income, and the distribution of income. The final two sections propose detailed models of the functions u and C .

3. The Nature of the Crime: A Preliminary Result

What is the nature of the crime? More specifically, how does it affect people of different income levels? It turns out that whether rich people or poor people are more sensitive to changes in the crime rate is crucial to understanding the diversity of preferred policies. Mathematically, we are asking about the sign of $D_{yc}u$. If $D_{yc}u$ is positive, then an increase in income reduces an individual’s sensitivity to changes in the crime rate. Think of crime on the subway—the

richer you are, the less frequently you take the subway, so that this crime affects you less the more money you have. If $D_{yc}u$ is negative, then an increase in income increases an individual's sensitivity to crime. Take the theft of luxuries from homes—the richer you are the more luxuries you have to lose, so that rich people may be more sensitive to such a crime. Thus, $D_{yc}u > 0$ for crimes that affect the poor more than the rich, while $D_{yc}u < 0$ for crimes that affect the rich more than the poor.

Now consider how individual i 's proposed policy $(\hat{\alpha}_i(c), \hat{v}_i(c))$ for achieving a *given* crime rate c depends on c . The crime rate can be decreased by increasing expenditure (increasing α) and/or making punishment more severe (reducing v). One might expect that the best way of reducing crime would be to do both. Indeed, if $D_{yc}u = 0$ this is true. However, if $D_{yc}u \neq 0$ the situation is more involved.

Suppose that $D_{yc}u$ is positive. Then if the crime rate falls, the marginal utility of a dollar decreases, so that direct expenditure is less expensive in utility terms. If $D_{yc}u$ is large enough, this may induce the individual to favor an increase in expenditure so much that she prefers to achieve a reduction in crime by *lessening* penalties. On the other hand if $D_{yc}u$ is negative enough, then an individual may want to accomplish a decrease in crime by increasing penalties and *reducing* expenditure.⁷ These results are summarized in Figure 3 (see the Appendix for details).

$$\begin{array}{ccc} \hat{v}'_i(c) > 0, \hat{\alpha}'_i(c) > 0 & | & \hat{v}'_i(c) > 0, \hat{\alpha}'_i(c) < 0 & | & \hat{v}'_i(c) < 0, \hat{\alpha}'_i(c) < 0 \\ \hline & & 0 & & D_{yc}u \rightarrow \end{array}$$

Figure 3. The dependence of the signs of $\hat{v}'_i(c)$ and $\hat{\alpha}'_i(c)$ on the value of $D_{yc}u$. Depending on the shape of the functions C and u the middle region may not exist.

To better understand the intuition behind these results consider the following. It may be desirable to reduce a crime like subway robbery by increasing expenditure and decreasing punishment. The reason that accepting a high tax rate, and hence a diminution of disposable income, is attractive, is that when a reduction in crime is accomplished, the individual is less concerned about not being able to afford other means of transportation. If this effect is strong enough (i.e. if $D_{yc}u = D_c(D_y u)$ is positive and large enough) then the individual will want to save on the “cost” of punishment at the same time. Conversely, an increase in the crime rate makes riding the subway all the more

⁷Notice the analogy to the possibility of inferior inputs in a production function. We can think of C as a production function and α and v as the inputs.

dangerous. The individual will want to have enough money to avoid the subway, and hence may want to accomplish this increased crime rate by a drastic cutback in expenditure accompanied by an increase in punishment.

On the other hand, consider the theft of luxuries. Reducing this theft through social expenditure is unattractive, since the less frequently your luxury goods are stolen, the more you value your money in order to purchase them (i.e. $D_{yc}u$ is negative). If this effect is strong enough then you may want to accomplish a lower crime rate by punishing more severely, but reducing expenditure (hence saving your money).

A word of caution is in order when interpreting these results. In our model there is only one crime. Thus, our results compare policies for controlling crime in *two* societies that differ as to the nature of their “primary” crime. They are not about how two different crimes should be controlled within the *same* society.

4. Differences Among the Policies Proposed by Individuals within an Economy

Individuals differ in their probabilities q_i of misapprehension and in their market incomes x_i . We now investigate how these differences affect the policies they propose.

4.1 Differences in Probabilities of Mistaken Conviction

Consider two individuals A and B who differ only in that A has a lower probability of mistaken conviction than has B . A has a lower marginal cost of punishment, and hence favors achieving any given crime rate with a harsher punishment and smaller expenditure than does B . At the same time A has a lower overall cost to reducing crime and so one might think that A favors a lower crime rate. This is not necessarily so, however.

Consider again the crime of robbery on the subway (a crime for which $D_{yc}u > 0$). Achieving a crime rate with relatively little expenditure, as A favors, leaves her relatively wealthy, and relatively unconcerned about this crime. Hence she may favor a higher crime rate than B does. For the theft of luxuries, on the other hand, A 's lower cost of crime deterrence and the fact that this crime affects wealthier people more reinforce each other, so that A favors a lower crime rate. We now make these arguments precise.

For individuals A and B with $q_i^A < q_i^B$, the slopes of A 's indifference curves in α - v space are higher than the slope of B 's indifference curves. Hence A prefers to achieve any given crime rate with more severe punishment and

lower expenditure. Regarding \hat{v}_i and $\hat{\alpha}_i$ as functions of q_i , we can express these comparative static results as $\partial\hat{v}_i/\partial q_i > 0$ and $\partial\hat{\alpha}_i/\partial q_i > 0$. (We use the partial derivative notation involving ∂ to emphasize that these are comparative statics; \hat{v}_i and $\hat{\alpha}_i$ as we have defined them are direct functions only of c .)

A comparative static calculation on the solution of (4) yields

$$\partial c_i^*/\partial q_i = -\hat{v}'_i(c_i^*)/(1 - q_i)\Delta,$$

where Δ , the second derivative of the objective function, is negative for a nondegenerate maximum. Thus A favors a lower crime rate unless $D_{yc}u$ is positive and sufficiently large (see Figures 3 and 4).

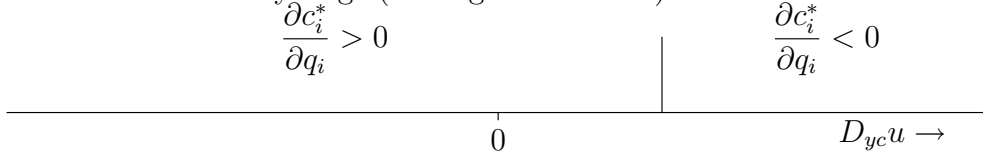


Figure 4. The dependence of the sign of $\partial c_i^*/\partial q_i$ on the value of $D_{yc}u$.

Even though A may prefer a higher or lower crime rate depending on the nature of the crime, she always prefers harsher punishment. This follows from the fact that $v_i^* = \hat{v}_i(c_i^*)$: since c_i^* depends on q_i and v_i^* also depends on q_i (independently of the dependence via c_i^*), we have

$$\begin{aligned} \partial v_i^*/\partial q_i &= \hat{v}'_i(c_i^*)(\partial c_i^*/\partial q_i) + \partial\hat{v}_i/\partial q_i \\ &= -(\hat{v}'_i(c_i^*))^2/(1 - q_i)\Delta + \partial\hat{v}_i/\partial q_i > 0 \end{aligned}$$

(where \hat{v}'_i is, as before, the derivative of \hat{v}_i holding q_i fixed).

We observe in passing the advantage of considering at least two variables in a deterrence model. If only punishment were available, then a desire for harsher punishment would go hand in hand with a desire for a lower crime rate. We have seen, however, that such a conclusion is unwarranted: one must consider the nature of the crime.

4.2 Differences in Market Incomes

We first note the following three points.

- Whether the potential benefit from a reduction in the crime rate is greater for a richer individual than for a poorer one⁸ depends on whether

⁸In speaking of “poorer” people, we mean those who are nonetheless rich enough to be paying taxes, rather than receiving subsidies (see the discussion between equations (1) and (2)).

richer individuals are more or less sensitive to changes in the crime rate than poorer ones (i.e. on the sign of $D_{yc}u$).

- The fact that individuals are risk-averse in income ($D_{yy}u \leq 0$) means that an extra dollar of expenditure becomes less costly in utility terms as income increases.
- The fact that the marginal tax rate is positive means that an increase in expenditure becomes more costly (in dollar terms) as income increases.

The crime rates that individuals would like to achieve vary with their income. Suppose that sensitivity to crime increases with income ($D_{yc}u < 0$; cf. the example of the theft of luxuries discussed above). This tends to make richer people favor a lower crime rate. At the same time richer people are paying more in taxes, but each dollar is worth less to them. When these last two factors just balance each other, the impact on utility of an increase in the scale α of the tax scheme is independent of income. In a sense, the tax system is then *utility neutral*. In this case, only the increased sensitivity to crime matters, so that, when $D_{yc}u < 0$, a richer person wants a lower crime rate. Mathematically we have the following:

$$\frac{\partial c_i^*}{\partial x_i} = \left(-\frac{1 - q_i}{\Delta} \right) [-\hat{\alpha}'_i(c)\{D_{yy}u \cdot (1 - \hat{\alpha}_i\tau') \cdot \tau + D_yu \cdot \tau'\} + D_{yc}u \cdot (1 - \hat{\alpha}_i\tau')],$$

where Δ is, as before, negative. The term in braces is the derivative with respect to x of

$$D_yu(x - T(x, \alpha), c) \cdot D_\alpha T(x, \alpha) = -\frac{\partial}{\partial \alpha} u(x - T(x, \alpha), c). \quad (5)$$

Thus, this term gives the change with income of the utility loss from an increase in the scale α of the tax system. For a utility neutral tax system this loss is independent of income, so that the term is zero. When the utility loss from higher α is greater for the rich, the term is positive—the tax scheme is “utility progressive”; when the utility loss is greater for the poor, taxation is “utility regressive”. To the extent that tax schemes are designed with a positive (and increasing) marginal tax rate in order to account for the fact that a dollar is worth less to a richer individual, we might expect tax schemes to be close to utility neutral.

When the tax scheme is neutral, the sign of $\partial c_i^*/\partial x_i$ is determined by $D_{yc}u$. If richer people are more sensitive to crime ($D_{yc}u < 0$) then they favor a lower crime rate, while if they are less sensitive to crime ($D_{yc}u > 0$) they favor a higher crime rate. The intuition for these results can again be illustrated by

the two examples we discussed in Section 3: for a crime like robbery on the subway, you prefer a higher crime rate if you are richer (since you can then afford to avoid the subway if necessary); for the crime of the theft of luxuries you prefer a lower crime rate if you are richer.

Now consider how an individual's preferred policy (α_i^*, v_i^*) varies with income. We just saw that the variation in the preferred crime rate depends on the nature of the crime. So the only remaining question is how people want to achieve different crime rates. Section 3 answered this. Thus, if $D_{yc}u$ is negative, but not too large in absolute value, richer people favor more expenditure and harsher punishment in order to achieve a lower crime rate. If $D_{yc}u$ is positive, but not too large, then it is poorer people who favor more expenditure and harsher punishment. All the possibilities are summarized in Figure 5 (compare this to Figure 3).

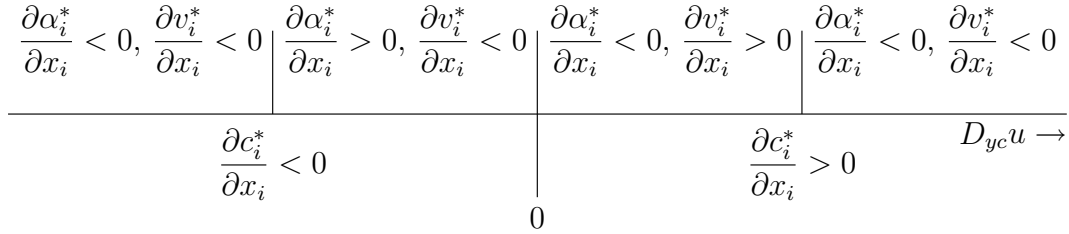


Figure 5. The dependence of the signs of $\partial\alpha_i^*/\partial x_i$, $\partial v_i^*/\partial x_i$, and $\partial c_i^*/\partial x_i$ on the value of $D_{yc}u$ when the tax system is utility neutral.

Now consider the effect of a non-neutral tax scheme. If the tax scheme is utility progressive, the increased burden on the rich works towards making them favor less expenditure, whereas if the tax system is regressive the opposite is true. (This can be seen in (9) in the Appendix, where the coefficient on the term in braces—the degree to which the tax system is utility progressive—is negative.)

5. Differences in Policies between Economies

The policy chosen by a society is determined in our model by a political mechanism that aggregates the preferred policies of all the members of the society. Differences in this mechanism thus explain variations in the observed policy. Other explanatory variables in our theory are the level and distribution of market incomes.

5.1 *Dependence on the Probability of Mistaken Conviction*

The fact that people may be mistakenly convicted of crimes is an important explanatory variable in our model. Consider two societies, A and B , which differ only in that the probability of a false conviction is lower in A than in B (perhaps there are many judicial safeguards in A). The results of Section 4.1 imply that every person in society B has a counterpart in society A who prefers harsher punishments. Thus, assuming that the political mechanism reflects this overall desire for increased punishment, society A will have more severe punishment than society B . The overall crime rate may be higher or lower in A , depending on the nature of the crime. If, in A , richer people are more sensitive to crime, the crime rate will be lower in A than in B ; if, in A , poorer people are sensitive enough to the crime rate, then the crime rate will be higher in A than in B . (All our results for this case are summarized in Table 1.)

5.2 *Dependence on the Political Mechanism*

Now suppose that the only difference between societies A and B is that in society A the political mechanism gives more weight to a relatively wealthy élite whose members are not likely to be falsely imprisoned (i.e. for whom q_i is low). Combining the results of Sections 4.1 and 4.2 we see that for a tax scheme that is close to utility neutral, members of the élite favor harsh punishment and a low crime rate when D_{ycu} is negative, and a relatively high crime rate and more severe punishment when D_{ycu} is positive enough. Thus we predict that a society ruled by a wealthy élite will have relatively harsh punishment and a low crime rate when the wealthy are more affected by crime than the poor, and a relatively high crime rate and relatively harsh punishment when wealthy people are much less affected by crime. (Again, see Table 1.)

5.3 *Dependence on the General Level of Income*

Consider two economies, one of which is wealthier than the other. Will the wealthier society choose a higher level of social expenditure to control crime? Will it pass its increased wealth onto criminals in the form of lessened penalties, or on the contrary punish more severely?

In order to gain some insight into these questions we proceed as follows. Suppose that economy A is obtained from economy B by doubling everybody's income. To abstract away from questions of changing marginal utility, assume that utility is linear in income. Redefine the parameter α so that the tax paid by a person in A is twice as high as that paid by the same person in B . Any

given person's income in A is twice what it is in B , so this means that

$$T_A(2x, \alpha) = 2T_B(x, \alpha), \quad (6)$$

and that the scheme indexed by α raises twice as much money in A as in B .

For the moment suppose the crime is theft and that in A the amount a thief can expect to steal is twice as much as it is in B . In a very simple economic model of the decision to steal, a person with potential market income z in economy B is just indifferent between stealing an amount s and working if $ps + (1 - p)v = z + a$, where p is the probability of being caught, v is the punishment utility, and a is the income transfer to the individual. The "same" individual in the richer economy, with potential market income $2z$, is just indifferent between stealing an amount $2s$ and working if the transfer a doubles and the punishment utility doubles as well. The two economies then have equal levels of crime:

$$C_A(\alpha, 2v) = C_B(\alpha, v). \quad (7)$$

We use this simple example to motivate (7), which is the assumption we now make. (This assumption remains valid in more involved models of the decision to commit crime, though clearly not in all such models.) Given (6) and (7), an individual with market income $2x$ in society A has the following utility when the tax scheme chosen is α and the punishment utility is $2v$.

$$(1 - q)u(2x - T_A(2x, \alpha), C_A(\alpha, 2v)) + q(2v) = \\ 2[(1 - q)u(x - T_B(x, \alpha), C_B(\alpha, v)) + qv].$$

(We have used the linearity of u in income.)

Thus the solution of problem (3) for an individual with income $2x$ in A is $(\alpha_i^*, 2v_i^*)$ if the solution to the same problem for an individual with income x in B is (α_i^*, v_i^*) . If we take account of our reindexation of the tax schemes, we conclude that each individual in B has a counterpart in A who favors twice as much expenditure and a punishment utility twice as high.

The intuition for this result rests on the observation that what matters about the punishment is its severity *relative to* the general level of income. If all incomes and the amount that may be obtained in theft double while the punishment utility remains constant, then the punishment becomes relatively harsher. Since the probability of false conviction is positive, this increase in severity comes at a cost, and if an individual was not willing to pay this cost in the original economy then she is not willing to pay it after incomes double.

In summary, our model identifies a tendency for a wealthier society to spend more on enforcement and/or redistribution and impose less harsh punishment. While this result is immediate once the problem has been properly framed, it contrasts with the results that would be obtained in other analyses of punishment. For instance, under a “moral” theory of punishment, in which penalties are set at the “just” levels, there is no reason to lessen the severity of punishment simply because society has become richer—at least no immediate reason, since the crime remains as “immoral” as before.

5.4 *Dependence on the Distribution of Income*

The distribution of income affects the cost of controlling crime. If the variance in income in society is large, then the punishment must be very severe and/or social expenditure high in order to deter the poorest individuals from criminal activity. For this reason one would expect the chosen crime rate to be higher in an economy with a very unequal distribution of income.

The interrelationships in our model are complex. However, we can obtain some insight into the forces at work if we hold the level of expenditure fixed and focus on the severity of punishment. In this case let $V(c)$ be the punishment necessary to induce a crime rate of c , and assume (as before) that $D_c V > 0$. Then the crime rate c_i^* that individual i proposes solves

$$\max_c (1 - q)u(y, c) + qV(c), \quad (8)$$

so that

$$(1 - q)D_c u(y, c_i^*) + qD_c V(c_i^*) = 0.$$

Let γ be a parameter that indexes the shift in V as the distribution of income changes. We choose γ so that a larger value represents a distribution of income with a longer left tail, and accordingly assume that

$$D_\gamma V < 0 \quad \text{and} \quad D_{c\gamma} V > 0.$$

That is, as the distribution of income becomes less equal, the punishment necessary to achieve any given crime rate increases in severity and the marginal change in the severity of punishment required to reduce the crime rate also increases.

Performing a comparative static calculation we obtain

$$\frac{\partial c_i^*}{\partial \gamma} = \left(-\frac{q}{\Lambda} \right) D_{c\gamma} V > 0$$

(since Λ , the second derivative of the objective function, must be negative for a nondegenerate maximum). Thus the model suggests that the more income inequality there is at low income levels, the more crime the policy chosen will induce. The effect on the severity of punishment itself is uncertain: on the one hand higher punishment is needed to attain any give crime rate, but on the other hand marginal changes in punishment are less effective.

In order to analyze the effect of an increase in inequality at high incomes as well as at low incomes, an additional factor must be considered. An increase in inequality at high incomes increases the income of the average taxpayer, so that if $D_{yc}u < 0$ the tendency for richer individuals to want a lower crime rate may reverse the previously found desire for a high crime rate; if $D_{yc}u > 0$ this desire is reinforced. That is, a society with a very unequal distribution of income may have a low crime rate for a crime like the theft of luxuries (since that is a crime to which rich people are sensitive) while having a high crime rate for robbery on the subway.

6. A Model of the Utility Function

Crime affects people in many ways. Theft reduces disposable income with positive probability, and induces expenditures on insurance and protection that reduce income with certainty. Shop-lifting and white-collar crime indirectly reduce income by raising prices. Violent crime induces expenditures on protection, and with positive probability impairs the ability to enjoy income.

We have taken as a primitive $u(y, c)$, which gives each individual's utility as a function of her income and the crime rate. One way of obtaining this function starts from the general formulation

$$u(y, c) = (1 - \pi(y, c))w(y - \ell_0(y, c)) + \pi(y, c)w(y - \ell_0(y, c) - \ell_1(y, c)),$$

where w is the basic utility function over disposable income, $\pi(y, c)$ is the probability of being the victim of direct criminal activity for an individual with income y when the crime rate is c , $\ell_1(y, c)$ is the amount lost in this case, and $\ell_0(y, c)$ is the amount paid for insurance and protection.

We are interested in the signs of $D_{yy}u$ and $D_{yc}u$. Consider two simple cases. First suppose that the direct losses to crime are zero (individuals are fully insured, for example). Then

$$u(y, c) = w(y - \ell_0(y, c)),$$

so that

$$D_{yy}u = w'' \cdot (1 - D_y\ell_0)^2 - w' \cdot D_{yy}\ell_0$$

and

$$D_{yc}u = -w'' \cdot (1 - D_y\ell_0) \cdot D_c\ell_0 - w' \cdot D_{yc}\ell_0.$$

If $w'' \leq 0$ and $D_{yy}\ell_0$ is small enough then $D_{yy}u \leq 0$. If the rate of increase with respect to the crime rate of the amount (indirectly) lost to crime increases fast enough as income increases ($D_{yc}\ell_0$ is large enough) then $D_{yc}u < 0$, while if this rate of increase is small enough then $D_{yc}u > 0$.

Second suppose that the crime is murder, so that the direct loss ℓ_1 is equal to the entire income, and the indirect loss is zero. Suppose that the probability of being murdered is independent of income, and normalize $w(0) = 0$. Then

$$u(y, c) = (1 - \pi(c))w(y),$$

so that $D_{yy}u \leq 0$ if $w'' \leq 0$, and $D_{yc}u \leq 0$.

7. A General Equilibrium Model of the Crime Rate $C(\alpha, v)$

Our previous analysis does not depend on the origin of C , which gives the crime rate as a function of the amount of direct expenditure (indexed by α) and the severity of punishment v . Many models would generate such a function. Here we briefly describe two models, which we deliberately keep simple.

Each individual i can either work or commit a crime. If she works, then she obtains the market income x_i adjusted by a tax/subsidy t_i , and is (mistakenly) punished (as above) with probability q_i . If the crime rate is c then her utility in this case is

$$(1 - q_i)u(x_i - t_i, c) + q_iv.$$

If she commits the crime then she gains g_i from it, adjusted by the tax/subsidy t'_i . The probability of her being punished is p_i , so that her utility if the crime rate is c is

$$(1 - p_i)u(g_i - t'_i, c) + p_iv.$$

Individual i decides whether or not to commit the crime by comparing these two utilities.

First suppose that the money raised by taxes is used exclusively to finance a police force. Then an increase in α means more police hours and, presumably, an increase in the probability p_i that a criminal goes to jail. Clearly, then, both an increase in α and a decrease in v result in a lower crime rate.

Now suppose that the money raised by taxes is used for redistributive transfers and that the tax scheme $T(x, \alpha)$ requires those individuals with incomes in excess of \bar{x} (and no others) to pay a positive tax. Suppose also that the tax paid (or subsidy received) by a criminal is a given function of the tax

the individual would pay (or be paid) were her income legal. (Maybe criminals pay no tax and receive no subsidies; or maybe they pay the same taxes and receive the same subsidies as workers.) Then there is a subsidy scheme $S(x)$ that distributes the revenue from the tax T among the individuals with incomes below \bar{x} in such a way that the crime rate is minimized. The resulting crime rate is $C(\alpha, v)$.

To illustrate in more detail how C could be constructed, assume that criminals pay no taxes and receive no subsidies, and that any individual who, under some policy, would decide to be a criminal, is insensitive to the crime rate. (Since the criminals in this model are those individuals with the lowest market incomes, this amounts to assuming that $D_c u(y, c) = 0$ if y is small.) Write $\bar{u}(y) = u(y, c)$ for these individuals. Then individual i is a criminal if

$$\left(\frac{1-p_i}{1-q_i}\right)\bar{u}(g_i) + \left(\frac{p_i-q_i}{1-q_i}\right)v > \bar{u}(x_i - t_i).$$

Given the distributions of p_i , q_i , and g_i in the population, as well as the distribution of income, we can then compute the crime rate for any subsidy scheme that costs α , and thus find the crime-minimizing scheme and hence the crime rate $C(\alpha, v)$.

8. Conclusion

The penalties imposed for crimes, the level of enforcement, and the extent of redistributive taxation show substantial variation across societies. These phenomena are related; both direct expenditures and punishment are instruments that can be used to control crime. Our goal has been to identify the factors that affect these instruments. We have shown the following to be important: how crime affects different members of society, the probabilities of mistaken conviction, the nature of the political mechanism, and the distribution and level of wealth. Especially interesting is the role played by the sensitivity of people to crime as their income changes—the unexpected importance of this factor emerged only upon analysis of the model. By explicitly allowing for more than one instrument in the control of crime, we have tried to expand the explanatory power of a “deterrence” model, while pointing out some pitfalls inherent in a pure punishment model.

Appendix: Details of the Comparative Statics

$\hat{\alpha}'_i(c)$ and $\hat{v}'_i(c)$

A comparative static calculation on the solution of (2) yields

$$\hat{\alpha}'_i(c) = \frac{1}{\Omega} \{ -(1 - q_i) D_{yc} u \cdot D_\alpha T \cdot (D_v C)^2 + \lambda (D_v C \cdot D_{\alpha v} C - D_\alpha C \cdot D_{vv} C) \}$$

and

$$\hat{v}'_i(c) = \frac{1}{\Omega} \{ (1 - q_i) D_v C \cdot D_\alpha T \cdot (D_{yc} u \cdot D_\alpha C - D_{yy} u) - \lambda (D_v C \cdot D_{\alpha\alpha} C - D_\alpha C \cdot D_{\alpha v} C) \}$$

where Ω is the determinant of the bordered Hessian and λ is the multiplier on the constraint. At a nondegenerate maximum we have $\Omega > 0$ and $\lambda = -q_i/D_v C < 0$. The signs indicated in Figure 3 follow from our assumptions on the signs of the partials of u , C , and T , and (1), under which the signs of the coefficients of λ in both cases are positive. If C is locally linear (so that the coefficient of λ in both cases is zero) and u is locally linear in y then the signs of $\hat{\alpha}'_i$ and \hat{v}'_i depend only on the sign of $D_{yc} u$, so that the middle region in Figure 3 does not exist.

$\partial c_i^*/\partial q_i$, and $\partial \alpha_i^*/\partial x_i$

The expression for $\partial c_i^*/\partial q_i$ follows from a comparative static calculation on the solution of (4). The expression for $\partial \alpha_i^*/\partial q_i$ is analogous to the one for $\partial v_i^*/\partial q_i$ obtained in the text:

$$\begin{aligned} \partial \alpha_i^*/\partial q_i &= \hat{\alpha}'_i(c_i^*) (\partial c_i^*/\partial q_i) + \partial \hat{\alpha}_i/\partial q_i \\ &= -\hat{\alpha}'_i(c_i^*) \hat{v}'_i(c_i^*) / (1 - q_i) \Delta + \partial \hat{\alpha}_i/\partial q_i. \end{aligned}$$

Thus if $D_{yc} u$ is large in absolute value and either positive or negative we have $\partial \alpha_i^*/\partial q_i > 0$, since in these cases the signs of $\hat{\alpha}'_i(c_i^*)$ and $\hat{v}'_i(c_i^*)$ are the same.

$\partial c_i^*/\partial x_i$, $\partial \alpha_i^*/\partial x_i$, and $\partial v_i^*/\partial x_i$

The expression for $\partial c_i^*/\partial x_i$ follows from a straightforward comparative static calculation on the solution of (4), as does the following:

$$\frac{\partial \hat{\alpha}_i}{\partial x_i} = \left(-\frac{1 - q_i}{\Omega} \right) \{ D_{yy} u \cdot (1 - \hat{\alpha}_i \tau') \cdot \tau + D_y u \cdot \tau' \} (D_v C)^2.$$

Since $\partial\alpha_i^*/\partial x_i = \hat{\alpha}'_i(c) \cdot (\partial c_i^*/\partial x_i) + \partial\hat{\alpha}_i/\partial x_i$ we have

$$\begin{aligned} \frac{\partial\alpha_i^*}{\partial x_i} = & (1 - q_i)\{D_{yy}u \cdot (1 - \hat{\alpha}_i\tau') \cdot \tau + D_yu \cdot \tau'\} \left[\frac{(\hat{\alpha}'_i)^2}{\Delta} - \frac{(D_vC)^2}{\Omega} \right] - \\ & \left(\frac{1 - q_i}{\Delta} \right) \hat{\alpha}'_i \cdot D_{yc}u \cdot (1 - \hat{\alpha}_i\tau'), \end{aligned} \quad (9)$$

(where $\Delta < 0$). For a utility neutral tax scheme the first term in braces is zero, so that the sign of $\partial\alpha_i^*/\partial x_i$ is the same as the sign of $\hat{\alpha}'_i \cdot D_{yc}u$.

Similarly,

$$\frac{\partial\hat{v}_i}{\partial x_i} = \left(\frac{1 - q_i}{\Omega} \right) \{D_{yy}u \cdot (1 - \hat{\alpha}_i\tau') \cdot \tau + D_yu \cdot \tau'\} \cdot D_vC \cdot D_\alpha C$$

and

$$\begin{aligned} \frac{\partial v_i^*}{\partial x_i} = & (1 - q_i)\{D_{yy}u \cdot (1 - \hat{\alpha}_i\tau') \cdot \tau + D_yu \cdot \tau'\} \left[\frac{\hat{\alpha}'_i \cdot \hat{v}'_i}{\Delta} + \frac{D_\alpha C \cdot D_vC}{\Omega} \right] - \\ & \left(\frac{1 - q_i}{\Delta} \right) \hat{v}'_i \cdot D_{yc}u \cdot (1 - \hat{\alpha}_i\tau'). \end{aligned}$$

	Rich much more sensitive to crime	Rich more sensitive to crime	Poor more sensitive to crime	Poor much more sensitive to crime
How one person wants to reduce crime	Harsher punishment			Lighter punishment
	Less spending	More spending		
Policy preferred by person with a lower probability of mistaken conviction	Harsher punishment			
	Less spending			Less spending
	Lower crime rate			Higher crime rate
Policy preferred by person with a higher income (under utility-neutral taxes)	Harsher punishment		Lighter punishment	Harsher punishment
	Less spending	More spending	Less spending	
	Lower crime rate		Higher crime rate	
Policy in a society with a lower probability of mistaken conviction	Harsher punishment			
	Less spending			Less spending
	Lower crime rate			Higher crime rate
Policy in society ruled by a wealthy élite	Harsher punishment			Harsher punishment
	Less spending			Less spending
	Lower crime rate			Higher crime rate

Table 1. A summary of some of the comparative static results. (Some of the results given here are not described in the text.) “Spending” refers to direct expenditure that reduces crime, including expenditure on enforcement and on redistributing income. Boxes that are empty correspond to cases in which the sign of the change depends on factors other than the nature of the crime.

References

- AUMANN, R. J., and M. KURZ [1977], "Power and Taxes," *Econometrica*, 45, 1137–1161.
- BECKER, G. S. [1968], "Crime and Punishment: An Economic Approach," *Journal of Political Economy*, 76, 169–217.
- BECKER, G. S. [1983], "A Theory of Competition Among Pressure Groups for Political Influence," *Quarterly Journal of Economics*, 98, 371–400.
- EATON, B. C., and W. D. WHITE [1991], "The Distribution of Wealth and the Efficiency of Institutions," *Economic Inquiry*, 29, 336–350.
- EHRlich, I. [1973], "Participation in Illegitimate Activities: A Theoretical and Empirical Investigation," *Journal of Political Economy*, 81, 521–565.
- FOUCAULT, M. [1979], *Discipline and Punish*, Vintage Books: New York. [Translation of French edition published in 1975.]
- HARRIS, J. R. [1970], "On the Economics of Law and Order," *Journal of Political Economy*, 78, 165–174.
- POLINSKY, A. M., and S. SHAVELL [1979], "The Optimal Tradeoff Between the Probability and Magnitude of Fines," *American Economic Review*, 69, 880–891.
- POLINSKY, A. M., and S. SHAVELL [1984], "The Optimal Use of Fines and Imprisonment," *Journal of Public Economics*, 24, 89–99.
- SCHOTTER, A. [1985], *Free Market Economics*, St. Martin's Press: New York.
- STIGLER, G. J. [1970], "The Optimum Enforcement of Laws," *Journal of Political Economy*, 78, 526–536.