The Optimal Use of Nonmonetary Sanctions as a Deterrent

By Steven Shavell*

A theoretical model of deterrence is studied in which the imposition of nonmonetary (as opposed to monetary) sanctions is socially costly. It is therefore desirable that the system of sanctions be designed so that sanctions are imposed infrequently. If courts possess perfect information, the optimal system is such that sanctions are never imposed — all who can be deterred will be — but, realistically, courts' information will be imperfect and sanctions will be imposed.

This article studies the optimal use of nonmonetary sanctions (such as probationary restraints on conduct, imprisonment, the death penalty) in a simple model in which the threat of sanctions may deter parties from committing contemplated acts. An important assumption of the model is that the imposition of nonmonetary sanctions is socially costly. The motivation for this assumption is that the imposition of nonmonetary sanctions is often associated with direct claims on goods and services (as with the operation of the prison system) and, in contrast to the case with monetary sanctions, results in disutility to punished parties that is not balanced in any automatic way by additions to the utility of other parties.

Social welfare is defined in the model to be the benefits parties obtain from committing acts less the harm done, the costs of apprehending parties, and the costs of imposing sanctions.[1] The social problem is to choose a set of sanctions and the probability of apprehension in order to maximize social welfare. As will be emphasized, the solution to this problem will involve design of a system of sanctions under which sanctions are actually imposed infrequently—since their imposition is assumed to be socially costly.

The social problem is first examined under the hypothetical assumption that *courts*[2] *can obtain perfect information about parties who have been apprehended*. With perfect information, courts will be able to recognize the two situations in which it will not be optimal to impose sanctions. One is where a party's act, such as speeding in an emergency, was socially desirable. A policy of imposing sanctions here will not be optimal because it might discourage socially desirable acts and, even if not, will result in pointless imposition of costly sanctions. The other situation where imposing sanctions will not be optimal is where, although a party's act was undesirable, it could not possibly have been deterred given the probability of apprehension. (Deterrence may have been impossible given the probability of apprehension because the disutility of sanctions is bounded.) Consider, for instance, the act of killing in the heat of passion or stealing a large amount where the chances of discovery are low. If deterrence of an act was impossible, clearly it will not be optimal to impose a sanction.

The remaining situation that courts will be able to recognize, assuming that they can obtain perfect information, is where a party's act was undesirable and could have been deterred by a sufficiently high sanction given the probability of apprehension. In this type of situation, it will be best for courts to adopt the policy of imposing sanctions large enough to deter, for then—since deterrence

*Professor of Law and Economics, Harvard Law School, Cambridge, MA 02138, and Research Associate, National Bureau of Economic Research. I thank L. Bebchuk and A. M. Polinsky for their comments on this paper and the National Science Foundation (grant no. SES-8420226) for financial support.

[1] Here and below, the word "sanctions" will mean nonmonetary sanctions.

[2] The term "courts" refers to the social authority responsible for deciding upon sanctions.

will be accomplished—there will never be an occasion in which a sanction is applied. In other words, any act that can be deterred given the probability of apprehension will be deterred without the actual imposition of costly sanctions. Suppose, for instance, that a party would obtain a benefit of 50 from committing an undesirable act, that the probability of his being apprehended is 10 percent, and that the disutility of possible sanctions can range as high as 1,000 (corresponding, say, to the death penalty). Then the party could and would be deterred under the optimal sanctioning policy; a sanction exceeding 500 would deter him. Moreover, note here, and generally, that the size of the optimal sanction is not uniquely determined. The sanction could as well be extremely high (1,000) or barely sufficient to deter (just over 500), since an extremely high sanction is no more costly than a lower one if neither will be imposed. In addition, observe that the size of the optimal sanction does not depend on the harmfulness of the act (whether the act does great or only moderate harm, the optimal sanction will still be any sanction exceeding 500), presuming that the harmfulness of the act is substantial enough to make it undesirable. The optimal sanction will, however, be affected by the probability of apprehension and by the benefits a party would derive from an act, because the expected sanction must be high enough to offset the benefits. Finally, the optimal probability of apprehension will reflect a balancing of two factors: the greater is the probability, the more parties who can possibly be deterred, but the larger are the expenses of policing behavior.

The social problem is then reconsidered under the realistic assumption that *courts cannot obtain perfect information about parties*. Thus courts will not be able to employ sanctions as just described; they will make various "errors" relative to the situation under perfect information, resulting in the discouragement of some socially desirable acts and in the actual imposition of sanctions. Courts may turn out to impose sanctions for three reasons: they may fail to recognize that an act was really socially desirable; they may fail to recognize that an undesirable act really could not have been deterred; they may fail to recognize how high a sanction was really necessary to deter an undesirable act that could have been deterred. Furthermore, unlike in the case where courts' information is perfect, the optimal magnitude of sanctions is uniquely determined[3]—it represents an appropriate compromise between achieving greater deterrence and the social costs borne due to the actual imposition of sanctions—and it does depend on the magnitude of harm done. The optimal probability of apprehension will reflect factors similar to those in the case where courts' information is perfect.[4]

I. The Model

Parties decide whether to commit harmful acts from which they would derive benefits.

[3] To illustrate, suppose the numerical example is modified as follows. There is another party who would obtain benefits of 150 from committing the undesirable act and thus who could not be deterred even by the maximum sanction of 1,000 (since the maximum expected sanction is 100). Suppose also that this second party cannot be distinguished from the first by courts—the sense in which courts' information is imperfect. Hence, courts must apply the same sanction to each party. It follows that a sanction of 500, but not one above 500, will be optimal: a sanction of just 500 will deter the first party; raising the sanction above 500 would not deter the second party but would lower social welfare because it would increase the expected social cost due to imposition of sanctions on the second party. (While the probability of apprehension was assumed to be fixed at 10 percent in this example, the probability would, of course, be optimally determined in the solution to a completely specified example; see Sections II and III.)

[4] The results described in this and the previous paragraph have not been discussed in the literature on deterrence to my knowledge; see Gary Becker (1968), R. Carr-Hill and Nicholas Stern (1979), A. Mitchell Polinsky and myself (1984), and references cited therein. In that literature the analysis has sometimes been carried out mainly at the level of the aggregate number of offenses (as in Becker); and where the analysis has been conducted at the level of the individual (as in Polinsky's and my article), it has not focused on the information the courts are able to obtain about an apprehended individual's benefits and the harmfulness of his act, and has usually been concerned with monetary sanctions and has implicitly assumed that such sanctions are socially costless to impose.

Let[5] b = benefit to a party from committing a harmful act; $0 \leqq b \leqq \bar{b}$; and h = harm resulting from an act; $0 \leqq h \leqq \bar{h}$.

A particular party is identified by b and h, and the distribution of parties by type is described by $f(b, h)$ = probability density of b and h; the function f is assumed to be known by courts. If a party commits an act, he might be apprehended and suffer a sanction. Let p = probability that a party who commits an act is apprehended and suffers a sanction; and s = sanction; $0 \leqq s \leqq \bar{s}$.

The level of s is determined by a sanctioning function, that is, a function of the variables that courts can observe (as discussed below). The utility of a party who commits an act will be b if he is not sanctioned and $b-s$ if he is; if he does not commit an act, his utility will be 0. Therefore, he will commit an act if[6]

$$(1) \qquad b > ps.$$

Because s is bounded by $\bar{s}$, it may be impossible to deter a party from committing an act. This will be true if

$$(2) \qquad b > p\bar{s},$$

and such a party will be referred to as undeterrable given p.

Social welfare is defined to be the benefits parties obtain from their acts,[7] less the harm done, the social costs due to the imposition of sanctions, and the costs of apprehending parties. In particular, suppose that σ = weight for calculating social costs from imposition of sanctions; σs is the social cost if the sanction imposed is s; $\sigma > 0$;[8] and $c(p)$ = social cost of maintaining the probability of apprehension at p; $c(p) > 0$; $c'(p) > 0$; $c''(p) > 0$.

Social welfare thus equals

$$(3) \qquad \int\int (b - h - p\sigma s) f(b, h)\, db\, dh - c(p),$$

where the integration is performed over the set of parties who commit the act as determined by (1).

The social problem is to choose the sanctioning function and the probability p to maximize social welfare; p will be assumed to be the same for all parties.[9] This choice will be called the optimal system of deterrence and will be denoted s^* and p^*. It will be assumed that $p^* > 0$; otherwise the social problem is without interest.

The first-best behavior of parties—that which would maximize social welfare if parties' behavior could be commanded—is for

[5]The benefits and (see below) the sanction are assumed to be bounded because, as is well known, the usual axioms of expected utility theory imply the boundedness of utility; see, for instance, Kenneth Arrow (1971).

[6]If $b = ps$, the party will be indifferent between committing the act and not; but in order to avoid having to make tedious qualifications, I adopt the convention that the individual will not commit the act in this case, and I adopt similar conventions below without further comment.

[7]In my paper (1985b, a previous version of this paper), parties' benefits were multiplied by a weight β that could be less than 1, so as to allow for society to discount parties' benefits. (Some might find the assumption of such discounting appealing where the source of a party's utility is the disutility experienced by a victim.) But the qualitative nature of the conclusions that were obtained there did not depend on whether β was less than 1, and were the same as the conclusions that will be reached below.

[8]The weight σ will exceed 1 if it measures both the resource costs of imposing sanctions and the disutility suffered by sanctioned parties.

[9]In fact, of course, p can be varied to some degree according to the characteristics b and h of parties; p will be constant only over subpopulations. Thus, if the reader wants, he may reinterpret the assumption that p is the same for all parties: he may say that for each subpopulation of parties (described by a density f) among whom p must be uniform, the social problem will be as posed here. In any case, the reader should note that there are two reasons why p will not vary in a completely independent manner with b and h. First, it will often be inefficient for an enforcement authority to be assigned responsibility to apprehend only one type of party. (It would be inefficient to have a policeman on the beat pursue only one type of lawbreaker; it will be efficient to have him pursue any lawbreaker whom he has a good opportunity to apprehend.) In other words, it is natural to assume that enforcement effort will often affect different types of parties in similar ways. Second, the ability of enforcement authorities to tailor their efforts to particular types of parties will be limited by authorities' ability to distinguish among types of parties *ex ante*.

parties to commit acts if and only if

$$b > h. \tag{4}$$

II. The Optimal System of Deterrence Where Courts Possess Perfect Information

In this case, courts can determine the benefits b and the harm h, so the sanction s may be a direct function $s(b,h)$ of b and h. Thus, we have

PROPOSITION 1: *If courts are able to obtain perfect information about parties who are apprehended, then under the optimal system of deterrence,*

(a) *Parties whose acts are undesirable and who can be deterred (given the probability of apprehension) will be deterred by the threat of a positive sanction—of at least b/p^*; such sanctions will equal the maximum possible sanction, $\bar{s}$, for some parties.*

(b) *Parties whose acts are undesirable and who cannot be deterred, and also parties whose acts are desirable, will face no sanction and will commit their acts.*

(c) *Sanctions will therefore never actually be imposed.*

(d) *The optimal probability of apprehension is such that the marginal cost of raising the probability equals the reduction in harm net of benefits of additionally deterred parties.*

The proof is given in the Appendix. Several characteristics of the optimal system of deterrence may now be noted.

(*i*) *Relation of the optimal sanction to h, b, and p^*.* An increase in h has no necessary effect on $s^*(b,h)$ within the area where parties' acts are undesirable and can be deterred since $s^*(b,h)$ need only satisfy $b \leqq p^*s^*(b,h)$ there; an increase in h could, however, make a desirable act undesirable and thus raise $s^*(b,h)$ from 0 to a positive level. An increase in b increases the minimum $s^*(b,h)$ needed to deter within the area of undesirable deterrable acts, namely, b/p^*; but an increase in b could make a deterrable act undeterrable and thus reduce $s^*(b,h)$ from a positive level to 0. An increase in p^* reduces the minimum $s^*(b,h)$ needed to deter in the area of undesirable deterrable acts, but because it enlarges this area, it can result in an increase in $s^*(b,h)$ from 0 to a positive level.

(*ii*) *Perfect information about b and h is socially valuable.* This is clear since $s^*(b,h)$ depends on both b and h in a nontrivial way.

(*iii*) *Comparison with the first-best situation.* The situation under the optimal system of deterrence is inferior to the first-best situation because undesirable acts that cannot be deterred are committed and because the expense $c(p^*)$ is incurred.

(*iv*) *Example.* To illustrate the optimal system of deterrence, consider the following discrete example in which there are three levels of benefits for each of three levels of harm:[10] parties of type A enjoy benefits of 50 from committing acts, those of type B enjoy benefits of 100, and those of type C benefits of 150; there are 10 A's, 1 B, and 10 C's for each level of harm; the levels of harm are 80, 300, and 900. In addition, $\bar{s} = 1{,}000$, $\sigma = 1$, and $c(p) = 1{,}000p$ for $p \leqq .12$ and $c(p) = 1{,}000{,}000p$ for $p > .12$. In this example, it is easy to verify that deterring the C's would not be optimal since that would require p to be at least .15,[11] but being able to deter A's and B's will be worthwhile since this will require p to be only .1; and $p^* = .1$ since it will clearly not be optimal for p to be any larger than the minimum necessary to deter the A's and B's. Hence, when $h = 80$, s^* will be at least 500 for A's, so they will be deterred, and s^* will be 0 for B's and C's. When $h = 300$ or 900, s^* will be at least 500 for A's and s^* will be 1,000 for B's, so A's and B's will be deterred, and s^* will be 0 for C's. Note that the outcome will differ from the first-best outcome since C's will commit acts when the harm is 300 and 900, and since the expense $c(.1) = 100$ will be incurred.

[10] It will be obvious that nothing rests on the discrete nature of this example; the same example will be reconsidered in the next section.

[11] If p were .15 rather than .1, the probability necessary to deter the B's, costs of apprehension would rise by $c(.15) - c(.1) = 31{,}200 - 1{,}000 = 30{,}200$. This amount clearly exceeds the harm net of benefits done by C's, so it cannot be optimal for p to equal .15.

III. The Optimal System of Deterrence Where Courts' Information is Imperfect

Now assume that while the courts can determine h, they cannot determine b; in particular, let $r =$ imperfect indicator of b observed by the courts; $0 \leqq r \leqq \bar{r}$.[12] Then social welfare may be written as

$$(5)\quad \int_0^{\bar{h}}\int_0^{\bar{r}}\int_{ps}^{\bar{b}}(b-h-p\sigma s)f(b|r,h) \times f(r|h)f(h)\,db\,dr\,dh - c(p),$$

where $f(b|r,h)$ and $f(r|h)$ are conditional probability densities, $f(h)$ is the unconditional density of h, and s is understood to be a function $s(r,h)$.[13] The solution to the social problem of maximization of (5) over sanctioning functions s and over p is described by

PROPOSITION 2: *If courts cannot obtain perfect information about apprehended parties, then under the optimal system of deterrence,*

(a) *There will generally be parties who commit acts and who are sanctioned. These parties will generally include both those whose acts are and are not desirable, and both those who are and are not deterrable.*

(b) *Also, some parties will generally be discouraged from committing desirable acts.*

(c) *The optimal sanction must be the maximum, $\bar{s}$, for some parties. If the optimal sanction is not 0 or $\bar{s}$, it will be such that the marginal social cost of raising sanctions, in terms of sanctions actually imposed, equals the net marginal social benefits due to deterrence of additional parties.*

(d) *The optimal probability of apprehension is such that the marginal cost of raising the probability plus the cost of imposing sanctions more frequently equals the reduction in harm net of benefits of additionally deterred parties.*

The proof is given in the Appendix. Characteristics of the optimal system of deterrence will now be discussed.

(i) *Relation of the optimal sanction to h, r, and p^*.* An increase in h causes $s^*(r,h)$ to rise, other things equal. To show this, note that the first-order condition (11) (see the Appendix) determining s^* is of the form $g(s^*,h)=0$.[14] Implicitly differentiating this with respect to h, we obtain $s^{*\prime}(h) = -g_h(s^*,h)/g_s(s^*,h)$. Now $g_s(s^*,h)<0$ (the second-order condition for optimality of s^*) and, assuming that the conditional density f does not change with h (to isolate the effect of a change in h), we have $g_h(s^*,h) = pf(ps|r,h)>0$, so that $s^{*\prime}(h)>0$ as claimed. The explanation for this is that if h increases, the marginal social cost of imposing sanctions is unaffected, but the marginal benefits of deterrence are increased.

Suppose that an increase in r indicates a rightward shift in the conditional distribution of b, that is, $\text{Prob}[b \geqq x|r,h]$ rises with r (where the probability is less than 1). Then writing (11) in the form $g(s^*,r)=0$ and proceeding analogously to the previous paragraph, we see that the sign of $s^{*\prime}(r)$ equals the sign of $g_r(s^*,r)$. Now $g_r(s^*,r) = -p^*\sigma d[\text{Prob}[b \geq p^*s^*|r,h]]/dr + p(h+p^*\sigma s^* - p^*s^*)f_r(p^*s^*r|,h)$. By assumption, the derivative of the probability is positive, so the first term is negative; and from (11), $p(h+p^*\sigma s^* - p^*s^*)>0$. Hence, certainly if $f_r(p^*s^*r|,h)<0$, g and thus s^* decrease with r; but if $f_r(p^*s^*|r,h)$ is sufficiently high, s^* could increase. The explanation for the ambiguity of the effect of an increase in r is straightforward: when the conditional probability distribution of b

[12] We can imagine that $r = r(b,h,\theta)$, where θ is a random variable, but we will not need to take the generation of r into explicit account.

[13] The lower limit of integration in (5) is ps because the assumption is that parties commit acts if $b > ps = ps(r,h)$. This means that it has been implicitly assumed that parties can predict what a court's observation r will be. It would be more realistic to assume that parties can only imperfectly predict what a court's observation r will be; but to assume that would complicate the expressions below without changing the essential nature of the results.

[14] It is assumed here and elsewhere in (i) that $s^*(r,h)$ is in $(0,\bar{s})$ and is thus determined by the first-order condition; otherwise, as $s^*(r,h)$ will be a corner solution, it will not change with small variations in h, r, and p^* (and the way it will change with large variations is similar to what was described in (i) following Proposition 1).

shifts to the right, the marginal social cost of sanctions increases (since more parties commit acts, more suffer sanctions); but the marginal social benefits may increase or decrease (because the density of parties just deterred may increase or decrease).

A small increase in p^* will lower $s^*(r, h)$ proportionately. This is immediately evident from the form of (9) (see the Appendix). Since p and s appear in (9) only as a product, it is clear that if p^* rises, lowering s^* so as to maintain $p^*s^*(r, h)$ constant must continue to maximize (9).[15]

(*ii*) *Imperfect information about parties' benefits is socially valuable*. The imperfect information r is socially valuable because given any h, $s^*(r, h)$ depends on the conditional density function $f(\cdot|r, h)$ in a nontrivial way.

(*iii*) *Comparison with the situation where courts possess perfect information*. The situation here differs because some parties suffer sanctions and because some parties whose acts are desirable may be discouraged from committing them.

(*iv*) *Example*. To illustrate the optimal system of deterrence, reconsider the example from the last section assuming that courts can observe h only; since they have no information about b, the sanction must be a function of h alone. The reader can verify that, as before, $p^* = .1$. The reader can easily see that for an h, p^*s^* must equal either 0, so that no parties are deterred; or 50, so A's only are deterred; or 100, so that A's and B's are deterred; that is, s^* must equal either 0, 500, or 1,000.[16] He can then verify that when $h = 80$, $s^* = 0$, so that A's, B's, and C's will commit acts;[17] when $h = 300$, $s^* = 500$, so that B's and C's will commit acts; and when $h = 900$, $s^* = 1{,}000$, so that C's only will commit acts. Comparing the outcome with that where courts had perfect information, we see that when $h = 80$, A's commit acts here but A's did not before (the problem here, of course, being inability to identify A's). When $h = 300$, B's commit acts here but did not before and B's and C's are sanctioned here but were not before; and when $h = 900$, C's are sanctioned here but were not before.

IV. Concluding Comments

(a) The model could be extended in the following ways.[18]

(*i*) An act could be associated with a probability distribution of harm (rather than with a single and certain level of harm). If so, it would be natural to assume that courts are unable to obtain perfect information about the probability distribution, for while courts can usually determine directly the harm actually done, they may lack important information about the harm that might have been done (as in an attempted crime).

(*ii*) A party's benefits could be allowed to depend on harm done. This assumption is often realistic since a party's object is often to do harm (as in murder or theft) and would furnish an indirect reason why the sanction should rise with harm: the higher the harm, the higher the party's benefits and thus the higher the sanction probably needed to deter him.

(*iii*) A party could be allowed to choose among a set of harmful acts. This would allow study of the issue of "marginal deterrence": discouraging a party who is *not* deterred from committing an undesirable type of act (kidnapping) from doing greater harm while committing the act (killing his victim)

[15] Recall, however, that we are assuming the change in p^* is small enough that s^* is still determined by the first-order condition.

[16] Were s to equal any other amount, deterrence would not be improved but more costs due to imposition of sanctions would be borne. For instance, if $s = 600$, then A's would be deterred, just as they would if $s = 500$, but B's and C's who were apprehended would suffer sanctions of 600 rather than 500.

[17] To see that $s^* = 0$ when $h = 80$, observe that if $s = 0$, the change in social welfare (parties' benefits less the harm from acts and costs of imposing sanctions) will equal $(10\times 50 + 100 + 10\times 150) - (21\times 80) = 420$; whereas if $s = 500$, the change in welfare will equal $(100 + 10\times 150) - (11\times 80) - (11\times 50) = 170$; and if $s = 1{,}000$, the change in welfare will equal $(10\times 150) - (10\times 80) - (10\times 100) = -300$. That s^* will be as claimed when h is 300 and 900 can be similarly verified.

[18] The extensions are discussed informally in my article (1985a).

by making the level of the sanction depend on the level of harm done.

(b) The main points regarding the optimal sanction shed light on the principles and doctrines of criminal law.[19] For instance, the importance given to "intent" in the criminal law may further the purposes of deterrence, for intent can be argued to be a rough proxy for a variety of factors (including the magnitude of the benefits a party derives from committing an act and the expected harmfulness of an act) that raise the optimal sanction. The results concerning the optimal sanction also obviously suggest the rationality of not punishing those who probably cannot be deterred (the insane or the coerced) or those whose acts were not undesirable (those who kill in self-defense).

(c) The optimal use of nonmonetary sanctions may be contrasted with the optimal use of monetary sanctions—Pigouvian taxes—in the classic model of harmful externalities. In that model, of course, if parties pay for harm done, a first-best outcome results, and the only information required by the social authority is the magnitude of the harm. The chief reason that the first-best outcome results without the authority's needing more information is that *imposition of taxes is implicitly assumed to be socially costless*; since the fact that parties for whom benefits exceed harm *pay* taxes does not itself reduce social welfare, there is no need for a social authority to ascertain parties' benefits.

(d) While in the present article, it was assumed that the form of sanctions was nonmonetary, the social decision whether to employ such sanctions rather than only monetary sanctions could also have been studied. Were this done, the presumed conclusion would be that nonmonetary sanctions would not be optimal to use unless the socially costless (or at any rate less costly) monetary sanctions could not adequately deter parties. That in turn would be more likely to be the case where the harm parties might do is great in relation to their assets and where it would be difficult to identify or apprehend parties who do harm.[20]

Appendix

PROOF OF PROPOSITION 1:

(a) The claim is that if $b \leqq h$ and $b \leqq p\bar{s}$, then $s^*(b,h)$ is any s satisfying $b \leqq ps$. This is obviously so, for such an s will prevent a decline in social welfare (see (3)) of $h-b$ and not involve the actual imposition of sanctions since the party will be deterred.

To show that $s^*(b,h)=\bar{s}$ for some parties, assume otherwise. Then it would be possible to lower p^* slightly and raise all positive $s^*(b,h)$ so as to maintain $p^*s^*(b,h)$ constant. This would result in the same behavior of parties, so that social welfare would not change on that account, but welfare would rise since $c(p)$ would fall. Hence, social welfare could not have been maximized, a contradiction.

(b) The first claim is that if $b \leq h$ and $b > p\bar{s}$, then $s^*(b,h)=0$. This is clearly so, since imposing a positive sanction would not deter the party but would lower social welfare by $p\sigma s$. The other claim is that if $b > h$, then $s^*(b,h)=0$. This is true because with $s^*=0$, the party will commit his act, increasing social welfare by $b-h>0$; whereas with a positive sanction, social welfare would rise by only $b-h-p\sigma s$ if the party was not deterred and would not rise at all if he was deterred.

(c) This is immediate from (a) and (b).

(d) From (a) and (b), we know that social welfare can be written

$$\text{(A1)}\quad \int_0^{\bar{b}}\int_h^{\bar{b}}(b-h)f(b,h)\,db\,dh - \int_{p\bar{s}}^{\bar{h}}\int_{p\bar{s}}^{\min(h,\bar{b})}(h-b)f(b,h)\,db\,dh - c(p).$$

The first term reflects the commission of socially desirable acts and the second, unde-

[19]See generally Richard Posner (1985) and my paper (1985a).

[20]This is a theme discussed informally in Posner and my paper (1985a); closely related points are made in Becker (pp. 190–93) and Polinsky and Shavell (p. 95).

sirable acts. The derivative of (A1) with respect to p is

$$\text{(A2)}\quad \bar{s}\int_{p\bar{s}}^{\bar{h}}(h-p\bar{s})f(p\bar{s},h)\,dh + \bar{s}\int_{p\bar{s}}^{p\bar{s}}(p\bar{s}-b)f(b,p\bar{s})\,db - c'(p) = \bar{s}\int_{p\bar{s}}^{\bar{h}}(h-p\bar{s})f(p\bar{s},h)\,dh - c'(p),$$

from which it follows that p^* is determined by

$$\text{(A3)}\quad c'(p) = \bar{s}\int_{p\bar{s}}^{\bar{h}}(h-p\bar{s})f(p\bar{s},h)\,dh.$$

The left-hand side is the marginal cost of raising p and the right-hand side is the marginal gain due to deterring additional parties.

PROOF OF PROPOSITION 2:

(a) and (b) are obviously true since the sanction $s^*(r,h)$ is not a function of b.

(c) The reason that $s^*(r,h)$ must equal $\bar{s}$ for some parties is essentially the reason given in the proof to Proposition 1: if not, p^* could be lowered slightly and all positive $s^*(r,h)$ could be raised so as to maintain $p^*s^*(r,h)$ constant. Then the integral in (5) would not change since p and s enter only as a product, yet $c(p)$ would fall, a contradiction.

To derive the first-order condition determining $s^*(r,h)$ when it is in $(0,\bar{s})$, note from (5) that $s^*(r,h)$ must maximize

$$\text{(A4)}\quad \int_{ps}^{\bar{b}}(b-h-p\sigma s)f(b|r,h)\,db$$

over s. The derivative of this expression with respect to s equals

$$\text{(A5)}\quad -p\sigma\,\mathrm{Prob}[b \geqq ps|r,h] + (h+p\sigma s-ps)pf(ps|r,h),$$

so that if $s^*(r,h)$ is in $(0,\bar{s})$, it is determined by

$$\text{(A6)}\quad p\sigma\,\mathrm{Prob}[b \geqq ps|r,h] = (h+p\sigma s-ps)pf(ps|r,h).$$

The left-hand side of (A6) is the marginal cost of increasing s—the expected marginal social cost $p\sigma$ of increasing the sanction per party who commits an act multiplied by the conditional likelihood $\mathrm{Prob}[b \geqq ps|r,h]$ that a party will do this; and the right-hand side is the marginal benefit of increasing s—the expected benefit per party just deterred multiplied by the rate of increase in these parties.

(d) For a given r and h, let $w(s,p)$ denote the value of (A4) given s and p, and write $s^* = s^*(p)$ to show the dependence of s^* on p (r and h are suppressed in the notation for now). We claim first that if s^* is determined by (A6), then $dw(s^*(p),p)/dp = 0$. To see this, observe that $dw(s^*(p),p)/dp = w_s s^{*\prime}(p) + w_p = w_p$ since $w_s = 0$ (this is (A6)); and w_p is easily verified to be a positive multiple of w_s, so that $w_p = 0$. If s^* is not determined by (A6), one possibility is that $s^* = 0$ and that the nonnegativity constraint is binding. In this case as $s^{*\prime}(p) = 0$, $dw(s^*(p),p)/dp = w_p$; but as w equals $\int_0^{\bar{b}}(b-h)f(b|r,h)\,db$, we have $w_p = 0$, so that the derivative is again 0. The remaining possibility is that $s^* = \bar{s}$ and the constraint $s \leqq \bar{s}$ is binding. In this case, as $s^{*\prime}(p) = 0$, we again know that $dw(s^*(p),p)/dp = w_p$; and as $w = \int_{p\bar{s}}^{\bar{b}}(b-h-p\sigma\bar{s})f(b|r,h)\,db, w_p = -\sigma\bar{s}\,\mathrm{Prob}[b \geqq p\bar{s}|r,h] + \bar{s}(h+p\sigma\bar{s}-p\bar{s})f(p\bar{s}|r,h)$. It follows from this, the fact that the integral in (5) may be rewritten as $\int_0^{\bar{h}}\int_0^{\bar{r}} w(s^*(p,r,h),p;r,h)f(r|h)f(h)\,dr\,dh$, and the fact that $w_p = 0$ on the boundary of the set $Z(p)$—defined to be the set of r and h where $s \leqq \bar{s}$ is binding—that p^* is determined by

$$\text{(A7)}\quad \sigma\bar{s}\int\int_{(r,h)\varepsilon Z(p)}\mathrm{Prob}[b \geqq p\bar{s}|r,h] + c'(p) = \bar{s}\int\int_{(r,h)\varepsilon Z(p)}(h+p\sigma\bar{s}-p\bar{s})f(p\bar{s}|r,h) \times f(r|h)f(h)\,dr\,dh.$$

The left-hand side of (A7) is the marginal social cost of increasing p, comprising both additional sanctions suffered (since more parties are apprehended) and the direct expense of raising p. (By comparison, in (A3) the marginal cost equaled only $c'(p)$, for no parties actually suffered sanctions.) The right-hand side is the marginal benefits in terms of increased deterrence.

REFERENCES

Arrow, Kenneth, J., *Essays in the Theory of Risk Bearing*, Chicago: Markham Press, 1971, 63–69.

Becker, Gary S., "Crime and Punishment: An Economic Approach," *Journal of Political Economy*, March/April 1968, *76*, 169–217.

Carr-Hill, R. and Stern, Nicholas, *Crime, the Police, and Criminal Statistics*, London: Academic Press, 1979, 280–309.

Polinsky, A. Mitchell and Shavell, Steven, "The Optimal Use of Fines and Imprisonment," *Journal of Public Economics*, June 1984, *24*, 89–99.

Posner, Richard A., "An Economic Theory of the Criminal Law," *Columbia Law Review*, October 1985, *85*, 1193–1231.

Shavell, Steven, (1985a) "Criminal Law and the Optimal Use of Nonmonetary Sanctions as a Deterrent," *Columbia Law Review*, October 1985, *85*, 1232–62.

______, (1985b) "The Optimal Use of Nonmonetary Sanctions as a Deterrent," Discussion Paper No. 10, Program in Law and Economics, Harvard Law School, September 1985.