A NOTE ON THE OPTIMAL USE OF NONMONETARY SANCTIONS

Louis Kaplow

Discussion Paper No. 61

7/89

Program in Law and Economics
Harvard Law School
Cambridge, MA 02138

The Program in Law and Economics is supported by a grant from the John M. Olin Foundation

A NOTE ON THE OPTIMAL USE
OF NONMONETARY SANCTIONS

Louis Kaplow

A Note on the Optimal Use of Nonmonetary Sanctions

Louis Kaplow

Abstract

When nonmonetary sanctions are used to deter harmful acts, it is understood that the optimal magnitude of the sanction will reflect not only the harmfulness of the act but also the social cost of imposing the sanction. This note explains why it may be optimal for nonmonetary sanctions to equal either zero or the maximum feasible sanction, rather than any intermediate level (deterring some individuals but not others).

A Note on the Optimal Use of Nonmonetary Sanctions

Louis Kaplow*

When nonmonetary sanctions are used to deter harmful acts, it is understood that the optimal magnitude of the sanction will reflect not only the harmfulness of the act but also the social cost of imposing the sanction. This note explains why it may not be optimal for nonmonetary sanctions to be at an intermediate level, deterring some individuals but not others.

One reason that no intermediate sanction may be optimal is that a component of the marginal cost of imposing a higher sanction -- the additional sanction borne by undeterred individuals -- falls as the sanction rises: the higher the sanction, the fewer the individuals who bear a marginal increase in the sanction. A second reason is that a component of the marginal benefit of imposing a higher sanction -- that higher sanctions, by deterring more individuals, result in fewer bearing the sanction -- increases as the sanction rises: the higher the sanction, the greater the savings in sanctioning costs for each additional individual that is deterred. The analysis uses a simple model, similar to those in Polinsky and Shavell (1984) and Shavell (1987), to demonstrate these effects and indicate when they dominate, making it optimal to impose no sanction or to impose the maximum feasible sanction (or, if lower, a sanction sufficiently high to deter all individuals, including those whose acts produce more benefits than the immediate harm they cause). Implications of this analysis for the optimal probability of nonmonetary sanctions are noted in the concluding remarks and an appendix.

---

*Harvard Law School and National Bureau of Economic Research. I am grateful for comments from Lucian Bebchuk, A. Mitchell Polinsky and Steven Shavell and for support from the John M. Olin Foundation.

1. Model and Analysis

Individuals each decide whether to commit an act. Individuals' benefits from acts, which differ, are not observed by the social authority. All acts cause the same level of harm. There is a given probability that the act will be subject to a nonmonetary sanction, the level of which is chosen by the social authority.[1]

The following notation is employed:

h = harm of the act.

b = individuals' benefits from acting.

$f(b)$ = differentiable distribution on $[0,\bar{b}]$; $F(b)$ is the c.d.f. of $f(b)$.

p = probability that harmful acts are detected.

s = nonmonetary sanction for detected harmful acts; $s \leq \bar{s}$.

σ = social cost (per unit) of nonmonetary sanction.

Individuals obtain the benefit b if they act and are subject to the sanction s with probability p, so they will act if and only if

(1) $b > ps$.

The type of individual just indifferent between acting and not acting is thus of type $b = ps$.

Given individual behavior, one can now state social welfare, which is taken to be the sum of individuals' benefits minus the harm of acts and sanctioning costs, where individual behavior is determined by (1).

---

[1] For reasons explored in Shavell (1989), the assumption that the probability is given may be approximately correct for many acts, as when detection resources simultaneously determine the probability of detection for many types of acts. The possibility that the probability of punishment may be selected independently for each particular act is discussed in the concluding remarks and is analyzed in an appendix.

$$(2) \quad W = \int_{ps}^{\bar{b}} (b - h - p\sigma s)f(b)db.$$

That is, each individual who acts affects social welfare by obtaining benefits of b, imposing harm of h, and, with probability p, facing the sanction s, which has a social cost per unit of σ. The first-order condition for the optimal sanction is[2]

$$(3) \quad \frac{dW}{ds} = (h + p\sigma s - ps)pf(ps) - p\sigma(1 - F(ps)) = 0.$$

The first term reflects the benefit from deterring the marginal individual -- which consists of avoiding the harm from the act plus the expected sanctioning cost minus the benefit from committing the act -- weighted by the rate at which the population is deterred as the marginal type, ps, increases. The second term reflects the cost from increasing the sanction applicable to inframarginal individuals, which consists of the marginal sanctioning cost per individual multiplied by the undeterred portion of the population. This expression indicates that an interior extreme point involves equating the (net) marginal and inframarginal effects of the change in sanction. To explore whether (3) involves a minimum or a maximum, it is useful to distinguish three cases, depending on the value of σ.

$\underline{\sigma > 1}$

The second-order condition for a maximum is

$$(4) \quad \frac{d^2W}{ds^2} = p^2(\sigma - 1)f(ps) + p^2(h + p\sigma s - ps)f'(ps) + p^2\sigma f(ps) \leq 0.$$

The first term is the product of the rate of change in the marginal benefit per individual deterred and the rate at which individuals are deterred as the sanction increases. When $\sigma > 1$, this term is always positive. The interpretation is that, the higher the sanction, the greater the benefit per individual deterred as a result of an increase in the sanction because the greater is the savings in sanctioning cost. Of course, as the sanction increases, the benefit of the marginal type, ps, also increases. But $\sigma > 1$

[2] This is analogous to condition (4) in Polinsky and Shavell (1984) and condition (A5) in Shavell (1987).

implies that $ps < p\sigma s$, so the savings in sanctioning cost dominates. The intuition is that the benefit of the marginal type just equals the expected sanction, while, when $\sigma > 1$, the social cost per unit of the sanction exceeds the private cost of the sanction. (Because the harm is the same for all individuals' acts, it does not enter into this term.)

The second term in (4) is the product of the marginal benefit per individual deterred, which is positive -- so long as $h + p\sigma s - ps > 0$, as, from (3), must be true if $dW/ds = 0$ -- and the rate at which the population density changes, which is of indeterminate sign. In the simple case of the uniform distribution, $f'(ps) = 0$, so the second term would equal zero. When the distribution is not uniform, this term will have the sign of $f'(ps)$.

The third term in (4) is the product of the cost from increasing the sanction on each inframarginal individual and the rate at which individuals are deterred as the sanction increases, both of which are positive. The interpretation is that, the higher the sanction, the lower is the fraction of the population that is undeterred, and thus the less is the marginal cost of increasing the sanction on inframarginal individuals. Since the inframarginal sanctioning cost effect is negative in (3) and it decreases as the sanction increases, this term is always positive.

In summary, the first and third terms are positive and the second is of indeterminate sign (it is zero in the simple case of the uniform distribution). Thus, it is plausible that any interior solution satisfying (3) will be a minimum rather than a maximum, in which case the optimum involves a sanction of 0 or $\bar{s}$.[3] In order for the solution to (3), if one exists, to be a maximum, $f'(ps)$ must be sufficiently negative at the sanction satisfying (3) that the magnitude of the second term in (4) exceeds the combined magnitude of the other terms.[4] That is, if the density of the

---

[3] In the latter case, if $p\bar{s} > \bar{b}$, any sanction in the interval $[\bar{b}/p, \bar{s}]$ will deter all individuals and thus produce equivalent welfare.

[4] Thus, for example, for an act where the distribution of the population is uniform in the interval $[0, \hat{b}]$, falling off rapidly just above $\hat{b}$ (although possibly rising again thereafter), any interior optimum will be such that $ps > \hat{b}$.

population at some point decreases rapidly as the benefit increases, it may be that some deterrence would increase welfare but further increases would not produce sufficient benefits to outweigh the additional sanctioning costs.[5]

Note, finally, that it is not only the case that an interior solution to (3) may be a minimum, but the first-order condition (3) need not hold for any $s \in [0,\bar{s}]$, even though (3) is continuous in s. For example, with the uniform distribution, dW/ds is everywhere increasing. And, if $hf(0) > \sigma$, $dW/ds > 0$ at $s = 0$. With the uniform distribution, it is also possible that, at $s = 0$, $dW/ds < 0$ and that it remains negative at the maximum feasible sanction.[6]

$\sigma \in [.5,1]$

In this case, the first term in the second-order condition (4) is negative (or zero if $\sigma = 1$). The intuition is that, when the social cost of the sanction is less than the private cost, the marginal benefit of deterring each additional individual falls as the sanction rises: the forgone private benefit, which for the marginal type just equals ps, is less than the savings in social cost, $p\sigma s$, from not applying the sanction to the marginal individual.

Note, however, that the first and third terms in (4) can be combined to yield

---

[5] Further illumination in this case can be derived by rearranging the first-order condition (3) to yield:

$$s = \frac{\sigma(1 - F(ps))}{p(\sigma - 1)f(ps)} - \frac{h}{p(\sigma - 1)}.$$

Casual inspection indicates that, when $\sigma > 1$, the optimal sanction would appear to be *higher* the lower is h, the greater is the undeterred portion of the population (on whom application of the sanction is a waste) and the lower is f(ps), the rate at which the benefits of deterrence increase with the sanction. As all three results are counterintuitive, one is led to suspect that such a solution is a minimum rather than a maximum.

[6] It must be that dW/ds is positive at some point because it clearly is positive at $s = \bar{b}/p$: the second term in (3) equals zero while the first term remains positive. There is, however, no guarantee in the case in which $\bar{s} < \bar{b}/p$ that (3) would become positive at any $s \leq \bar{s}$. (Note, however, that if $\sigma < 1$, it is no longer true that dW/ds is necessarily positive at $s = \bar{b}/p$. From (3), it can be seen that a necessary condition for dW/ds to become negative at $s = \bar{b}/p$ -- that is, at $ps = \bar{b}$ -- is that $h < (1-\sigma)\bar{b}$. This requires $h < \bar{b}$ and σ sufficiently small.)

(5) $p^2(2\sigma - 1)f(ps)$.

When $\sigma > .5$, (5) is positive, again indicating that an interior solution would be a minimum unless $f'(ps)$ is negative by a sufficient amount.[7]

$\underline{\sigma < .5}$

In this case, the combined effect of the first and third terms in (4) is negative, as indicated by (5), so an interior solution will be a local maximum unless $f'(ps)$ is sufficiently positive.

2. Concluding Remarks

Although intuition might have suggested that there generally would be an intermediate solution for the optimal level of nonmonetary sanctions -- which would involve equating the marginal benefits of increasing the sanction with the marginal costs -- the analysis here demonstrates that this need not be the case. It is plausible that the optimal nonmonetary sanction will be zero or maximal -- that is, the maximum feasible sanction or, if less, a sanction sufficient to deter all individuals from committing the act. (Note that this result holds even if many individuals' benefits from committing the act exceed the harm caused by the act.) And, when the choice is between deterring everyone[8] and deterring no one, no sanctioning costs are involved with either possibility, so the selection between these extremes would be made based solely on whether the harm exceeds the average benefit from committing the act.[9]

---

[7] If $\sigma = .5$, (5) equals zero. In the uniform case the first derivative (3) is constant -- in particular, $dW/ds = p(fh-.5)$, where f is the density -- so there cannot be an interior optimum (unless $fh = .5$, in which case the level of the sanction has no net effect on welfare).

[8] If the maximum feasible sanction is insufficient to deter everyone -- that is, if $p\bar{s} < \bar{b}$ -- the sanction would be imposed on undeterred individuals, so the sanctioning cost of $p\sigma\bar{s}(1-F(p\bar{s}))$ would have to be taken into account.

[9] One implication is that, if many acts involve the same distribution of individuals' benefits but different levels of harm and the optimum for each level of harm involves extreme sanctions, the optimal sanction will be nondecreasing in the harm: above some level of harm, the optimal sanction will

The existence and strength of the tendency for an intermediate solution to be a minimum depends on the magnitude of the social costs of the nonmonetary sanction relative to the private costs. In this model, the private cost was normalized to one. The most plausible case for the interior solution to be a minimum was when $\sigma > 1$ -- that is, when the social cost exceeds the private cost.[10] This condition would hold if the social cost is thought to include the private cost (deprivation of liberty, in the case of imprisonment) plus additional costs (such as those of running prisons). When this relationship holds, it is beneficial, at the margin, to deter an individual no matter how great the benefit of the act relative to its direct harm because, for the marginal individual, the benefit of the act equals the private cost of the sanction, which in turn is less than the social cost of the sanction (plus the harm of the act). And, for the contrary result, the greater the extent to which the social cost of the sanction is less than the private cost, the more plausible it is that the optimal sanction would be intermediate. The social cost may be less than the private cost if the private cost (loss of liberty) is not deemed a social cost[11] or if the sanction produces social benefits aside from deterrence (e.g., incapacitation, retribution).

The potential optimality of an extreme policy of no deterrence or complete deterrence (to the extent feasible) arose in a model in which the probability of being sanctioned was fixed,[12] but this assumption is not essential, as demonstrated in the Appendix. In this alternative case, the optimal sanction, if any is to be imposed, is the maximum feasible sanction, so the choice of the probability determines the expected sanction. Because raising expected sanctions in this modified model involves additional expenditures of enforcement resources, a lower expected sanction is dictated when an intermediate solution is optimal and it is more plausible that an intermediate

---

be $\bar{b}/p$ and, below that level, the optimal sanction will be zero. Considerations of marginal deterrence, however, might justify modifying this simple result.

[10] This was the assumption used in Polinsky and Shavell (1984).

[11] That is, a society might make the judgment that the utility of criminals is of little or no weight, at least with respect to the sanction they bear for committing their criminal act.

[12] See note 1.

solution will be optimal, assuming that enforcement is subject to diminishing returns.

Finally, note that the model of nonmonetary sanctions analyzed here is applicable to any sanction where the social cost is (roughly) some fixed proportion of the private cost of the sanction. For example, if monetary sanctions are imposed on risk-neutral individuals and there is some collection cost that is proportional to the fine, this model would apply. Since the cost often might be a small fraction of the fine revenue, it seems plausible in this case that an interior optimum would exist. If, however, individuals are risk-averse, the social cost of the sanction includes risk-bearing costs, which could be a substantial portion of the private cost. Thus, the discussion here is suggestive of the result that there may not be an intermediate optimum.[13]

---

[13] The private risk-bearing cost is necessarily less than the private cost of the sanction because the private cost is the sum of the expected penalty, ps, and the risk-bearing cost. Thus, in terms of this model (which is not strictly applicable due to the nonlinearities that arise with risk aversion), one necessarily has $\sigma < 1$, but, without further restrictions on the parameters, σ could take on any value in the interval (0,1). For analysis of the case when individuals are risk-averse, see Kaplow (1989), Polinsky and Shavell (1979).

Appendix: The Optimal Probability of Nonmonetary Sanctions

Assume that the social authority can choose the probability of sanctions, where the enforcement cost is c(p), with $c'(p) > 0$, $c''(p) > 0$. Individual behavior remains as stated in (1) and social welfare can be expressed as

$$(2')\quad W = \int_{ps}^{\bar{b}} (b - h - p\sigma s)f(b)db - c(p).$$

Because behavior and sanctioning costs are determined by ps, it is clear, contrary to the suggestion of Polinsky and Shavell (1984), that the optimal sanction in this case is $\bar{s}$: for any lower s, one could raise the sanction and reduce the probability to keep ps constant, which would not affect the integral in (2') but would reduce c(p). See Shavell (1989).

The first-order condition for the optimal probability is[14]

$$(3')\quad \frac{dW}{dp} = (h + p\sigma s - ps)sf(ps) - s\sigma(1 - F(ps)) - c'(p) = 0.$$

The only differences between (3') and (3) are that the first two terms in (3') are weighted by s rather than by p (taken alone, this has no effect) and there appears a third term in (3') reflecting the marginal cost of raising the expected sanction, which implies that, if an interior optimum exists, it involves a lower expected sanction than indicated by (3), as one would expect.

The second-order condition for a maximum is

$$(4')\quad \frac{d^2W}{dp^2} = s^2(\sigma - 1)f(ps) + s^2(h + p\sigma s - ps)f'(ps) + s^2\sigma f(ps) - c''(p) \le 0.$$

The only differences between (4') and (4) are that the first three terms in (4') are weighted by s^2 rather than by p^2 (taken alone, again of no effect) and there appears a fourth term in (4') reflecting that the marginal cost of raising the expected sanction is increasing, which implies that an interior optimum is more plausible than when s was the choice variable. In all other respects, the analysis for the case where the expected sanction is determined solely by selecting s rather than by selecting p is applicable.

---

[14] This is analogous to condition (5) in Polinsky and Shavell (1984).

References

Kaplow, Louis, The Optimal Probability and Magnitude of Fines for Acts That Are Definitely Undesirable, Harvard Program in Law and Economics Discussion Paper No. 55, 1989.

Polinsky, A. Mitchell and Shavell, Steven, The Optimal Tradeoff Between the Probability and Magnitude of Fines, American Economic Review 69, 1979, 880-891.

__________ and __________, The Optimal Use of Fines and Imprisonment, Journal of Public Economics 24, 1984, 89-99.

Shavell, Steven, The Optimal Use of Nonmonetary Sanctions as a Deterrent, American Economic Review 77, 1987, 584-592.

__________, Specific Versus General Enforcement of Law, Harvard Program in Law and Economics Discussion Paper No. 58, 1989.