ISSN 1045-6333

ENFORCEMENT COSTS AND THE OPTIMAL MAGNITUDE AND PROBABILITY OF FINES

A. Mitchell Polinsky

and

Steven Shavell

Discussion Paper No. 74

7/90

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.

Abstract. Some of the costs of enforcing laws are "fixed" -- in the sense that they do not depend on the number of individuals who commit harmful acts -- while other costs are "variable" -- they rise with the number of such individuals. This article analyzes the effects of fixed and variable enforcement costs on the optimal fine and the optimal probability of detection. It is shown that the optimal fine rises to reflect variable enforcement costs; that the optimal fine is not directly affected by fixed enforcement costs; and that the optimal probability depends on both types of enforcement costs.

July 1990

ENFORCEMENT COSTS AND THE OPTIMAL MAGNITUDE AND PROBABILITY OF FINES

A. Mitchell Polinsky
and
Steven Shavell*

1. Introduction and Summary

Whenever the government uses fines to control harmful activities, it incurs certain costs in connection with the enforcement process. This article will show how these costs affect the optimal magnitude of fines and the optimal degree of enforcement effort, that is, the optimal probability of detection of those who commit harmful acts.

We will distinguish between two types of enforcement costs. *Fixed enforcement costs* are costs that do not depend on the number of individuals who commit harmful acts. These costs would include, for example, the expenses borne by a pollution control agency in testing water quality in a lake. *Variable enforcement costs* are costs that do depend on the number of individuals who commit harmful acts, such as the costs of prosecuting and penalizing polluters. In our basic model, we assume that the fixed enforcement costs are the costs of maintaining the probability of detection at a certain level, and that the variable enforcement costs are the costs of fining individuals.

The main conclusions of this article may be summarized as follows. *The optimal fine equals the harm, properly inflated for the chance of not being detected, plus the variable enforcement cost of imposing the fine.* If the fine is of this magnitude, the expected fine will equal the harm caused by the harmful act plus the expected increase in

---

*Stanford University and National Bureau of Economic Research; Harvard University and National Bureau of Economic Research, respectively. Polinsky's research was supported by the John M. Olin Program in Law and Economics at Stanford Law School; Shavell's research was supported by the National Science Foundation (grant SES-8821400). We are grateful for helpful comments from Jeffrey Parker and seminar participants at the University of Michigan and the Bay Area Group in Economics and Law, and for research assistance from Timothy Church and Joel Waldfogel.

enforcement costs occasioned by it. For instance, suppose that the probability of detecting a polluter is 50%, that a polluter would cause $10,000 of harm, and that the cost of imposing a fine on a polluter is $3,000. Then the optimal fine is $23,000: $10,000 multiplied by 2 because of the 50% chance of detection, plus $3,000. Given this fine, the expected fine is $11,500 (50% x $23,000). Thus, if a party pollutes, he will take into account the harm of $10,000 as well as the expected increase in enforcement costs of $1,500 (reflecting the 50% chance that he will be caught and that the state will incur variable enforcement costs of $3,000).

Note that *fixed enforcement costs do not directly affect the optimal fine*. The fixed cost of catching polluters -- such as the expense of testing water quality -- does not directly affect the optimal fine because the fixed cost is, in essence, a sunk cost. When a party pollutes, he does not increase this fixed cost. However, as will be explained, fixed enforcement costs do influence the optimal probability of detection; they therefore indirectly affect optimal fines, since the factor by which the harm is inflated is the reciprocal of the probability of detection.

The optimal probability of detection depends on both types of enforcement costs. If variable enforcement costs are high, the optimal probability will be low because enforcement will be expensive; and for sufficiently high variable enforcement costs, the optimal probability will be zero. Similarly, if fixed enforcement expenditures are not very productive (that is, if it is very costly to raise the probability), the optimal probability will be low; in the extreme it will be zero. If it is inexpensive to raise the probability, the probability will be higher, but may be less than one; for even if the probability could be raised costlessly, it may not be desirable to raise it to one because variable enforcement costs may then become excessive.

Sections 2 and 3 derive the results just summarized in a model of optimal enforcement. Section 4 discusses several extensions of the model: a stage of investigation and prosecution after detection of an individual; variable enforcement costs that increase with the magnitude of fines; the possibility that the probability of detection falls as the fine rises; the sanction of imprisonment; and the bearing of expenses by individuals in connection with the enforcement process. Section 5 comments on the importance of enforcement costs, the practical ability to use the formulas developed in this article, and the inclusion of enforcement costs in actual sanctions.[1]

2. The Model

In the model, risk-neutral individuals contemplate whether to commit a harmful act that yields benefits to them. Each individual is identified by the harm that he would cause if he commits the act and by the benefit he would obtain from it. If an individual commits the harmful act, he will be fined with some probability. The cost to the state of establishing this probability, which is assumed to be independent of the number of individuals who commit the harmful act, is the fixed enforcement cost. The cost to the state of imposing the fine, which depends on the number of individuals who commit the harmful act, is the variable enforcement cost.[2]

The following notation will be used.

h = harm caused if the harmful act is committed; $h \geq 0$;

$g(h)$ = probability density of h over individuals; g is positive for $h \geq 0$;[3]

b = benefit from committing the harmful act; $b \geq 0$;

$r(b)$ = probability density of b over individuals; r is positive for $b \geq 0$;

c = cost of establishing the probability of detection (fixed enforcement cost); $c \geq 0$;

$p(c)$ = probability of detection; $p(0) = 0$, $p'(c) > 0$, $p''(c) < 0$;

w = wealth of each individual;

$f(h)$ = fine given the harm; $0 \leq f(h) \leq w$;

---

[1]Our article builds on Becker (1968), who recognized that the optimal magnitude and the optimal probability of sanctions reflect not only the level of harm, but also the cost of enforcement. He observed (p. 192), for example, that if the probability of detection is one, the optimal fine equals "the sum of marginal harm and marginal [enforcement] costs." Stigler (1970, p. 533) also mentions that fines should include enforcement costs. However, the precise way that the two types of enforcement cost affect the optimal fine and the optimal probability has not been investigated. But see note 7 below.

[2]This cost can be viewed either as the expense of collecting a fine through a formal enforcement process, or as the expense of settlement negotiations.

[3]If h were the same for everyone, the model would not be interesting because, regardless of the level of enforcement costs, the optimal fine would equal the wealth of each individual.

k = cost of imposing the fine (variable enforcement cost).

The following assumptions are made. The harm h is observable, so that the fine f can be made to be a function of h, but the benefit b is not observable. The variables h and b are independently distributed (this assumption is made for simplicity). The probability of detection p is the same for acts resulting in different levels of harm.[4]

An individual will commit a harmful act if his benefit is greater than or equal to the expected fine:[5]

$$b \geq pf. \tag{1}$$

Social welfare is the sum of the benefits individuals obtain from committing harmful acts, less the harm they do, and less expected enforcement costs. Given (1), social welfare can be expressed as

$$\int_0^\infty \int_{pf}^\infty (b - h - pk)r(b)db\,g(h)dh \; - \; c, \tag{2}$$

where $p = p(c)$ and $f = f(h)$. Note that the cost of maintaining the probability of detection, c, is unaffected by the number of individuals who commit the harmful act. However, the expected cost of imposing fines increases by pk for each individual who commits the act.

Society's problem is to choose c and a schedule $f(h)$ to maximize (2). Let c^* and $f^*(h)$ denote the solution to the social problem (and let $p^* = p(c^*)$).

3. Analysis of the Model

This section derives the optimal fine schedule and the optimal probability of detection, and shows how they depend on fixed and variable enforcement costs.

3.1 Optimal Fines

Given any positive c and p, the social problem for any h is to choose the fine f to

---

[4]A justification for this assumption is provided in Shavell (1989): To the extent that the investment in enforcement effort applies to a wide range of harms, it is appropriate to treat the probability of detection with respect to any one type of harm as fixed. (If, alternatively, p could be chosen independently for each h, the optimal fine would equal wealth for each h.)

[5]For convenience, we assume that when $b = pf$, an individual will commit a harmful act.

maximize

$$\int_{pf}^{\infty} (b - h - pk)r(b)db. \tag{3}$$

The derivative of (3) with respect to f is $-p(pf - h - pk)r(pf)$, which is positive for $pf < h + pk$ or, equivalently, for $f < h/p + k$; zero at $pf = h + pk$ or $f = h/p + k$; and negative for greater pf and f. Hence, $f^*(h) = h/p + k$ if this is feasible, that is, if $h/p + k \leq w$; otherwise, $f^*(h) = w$.

In other words, the optimal fine is

$$f^*(h) = h/p + k \tag{4}$$

if the harm caused by the harmful act is less than or equal to $p(w - k)$. Therefore, if $w > k$, there will be a range of harms for which the optimal fine is $h/p + k$.

Observe that, for harms in this range, the expected fine equals $h + pk$, which is the increase in expected social costs if an individual commits a harmful act -- that is, the harm caused plus the expected cost of imposing the fine. Thus, deterrence is socially optimal. For harms greater than $p(w - k)$, the expected fine is pw, which is less than $h + pk$, so that there is underdeterrence.[6]

Note too that the fixed enforcement cost c of establishing the probability of detection does not enter into the optimal fine formula (except indirectly through its influence on the probability of detection p).

In summary,

[6]In Polinsky and Shavell (1982), we studied a model in which firms causing an externality were detected with probability one and enforcement costs were associated with the imposition of Pigouvian taxes. In the notation of the present model, then, the only enforcement cost was k. We derived the optimal tax, but it did not equal what (4) implies, the harm plus the enforcement cost k. The reason is that we assumed that the sanction took the particular form of a tax per unit of a firm's activity. (To be precise, if a firm's level of activity is x, it causes harm hx, so that, according to (4), the sanction should be $hx + k$. But we assumed that the tax was a constant t per unit of activity, so it took the form tx.)

Proposition 1. For harms below a threshold ($h \leq p(w - k)$), the optimal fine $f^*(h)$ is $h/p + k$; the expected fine is $h + pk$; and deterrence is optimal. The optimal fine includes the variable enforcement cost k of imposing the fine, but not the fixed enforcement cost c of establishing the probability of detection. For higher harms, the optimal fine $f^*(h)$ is at its maximum, w; the expected fine is $pw < h + pk$; and there is underdeterrence.◂

3.2 Optimal Probability of Detection

Given Proposition 1, social welfare (2) can be expressed as

$$\int_0^{p(w-k)} \int_{h+pk}^{\infty} (b - h - pk)r(b)db\,g(h)dh + \int_{p(w-k)}^{\infty} \int_{pw}^{\infty} (b - h - pk)r(b)db\,g(h)dh - c. \tag{5}$$

The first term is associated with individuals who cause harm that is below the threshold and whose expected fine is h + pk, whereas the second term applies to individuals who cause higher harms and whose expected fine is pw.

Differentiating (5) with respect to c and setting the result equal to 0, we obtain (after cancelling several terms) the first-order condition

$$r(pw)p'(c)w \int_{p(w-k)}^{\infty} (h + pk - pw)g(h)dh = 1 + p'(c)k\left[\int_0^{p(w-k)} (1 - R(h + pk))g(h)dh + (1 - R(pw))(1 - G(p(w - k)))\right], \tag{6}$$

where R and G are the cumulative distribution functions of r and g. The optimal probability p^* is determined implicitly by (6). The left-hand side of (6) is the marginal benefit of raising c and therefore p: for harms above the threshold $p(w - k)$, there is underdeterrence; raising p just deters those who would obtain benefits of pw and therefore raises welfare by $h + pk - pw$ for each such individual. The right-hand side is the marginal cost of raising c and p, which has two components: the direct cost of raising c, 1; and the increase in variable enforcement costs accompanying more frequent imposition of fines.

How p^* depends on variable and fixed enforcement costs is summarized in the next two propositions.

Proposition 2. As the variable enforcement cost of imposing the fine increases from $k = 0$, the optimal probability p^* may increase or decrease. Eventually, however, p^* must decrease and equal 0; in other words, if k is sufficiently large, no enforcement is optimal.◂

The proof of this proposition is contained in the Appendix. The reason that p^* initially may increase or decrease as k increases can be explained as follows. The marginal benefit of raising p grows because, as k increases, the social loss from underdeterrence increases. However, the marginal cost of raising p may either rise or fall: it tends to rise because, as k increases, the variable enforcement costs incurred when more individuals are fined increase; but it tends to fall because, as k increases, fewer individuals commit the harmful act (since the expected fine rises with k for those individuals whose harm is below the threshold level of harm), thereby lowering variable enforcement costs. If the marginal cost of raising p is greater than the marginal benefit of raising p, p^* will fall; otherwise p^* will rise. The reason that p^* must approach zero as k grows large is that, otherwise, a positive fraction of individuals would be fined, and the variable enforcement costs of imposing fines would grow without bound. Why p^* must equal zero for k sufficiently large (rather than just approach zero) is more difficult to explain; see the proof in the Appendix.

Next, let us examine how the optimal probability of detection p^* is influenced by fixed enforcement costs. It will be helpful to define the "productivity" of fixed enforcement costs in the following way: the probability of detection equals $p(\lambda c)$, where $\lambda > 0$ is the productivity of costs c. Thus, the higher is λ, the higher is p for any given c. How p^* depends on λ is summarized in the next proposition.

Proposition 3. If the productivity λ of fixed enforcement costs in establishing the probability of detection is sufficiently low, the optimal probability p^* is zero. Thereafter, as λ increases, p^* may rise or (if positive) fall. However, as λ grows large, p^* tends toward $\hat{p}$, the p that would be optimal if the probability could be increased at no cost; $\hat{p}$ may be less than 1.◂

The proof of this proposition also is provided in the Appendix. The reason that p^* equals zero for λ sufficiently low is evident; for λ very low, the marginal cost of raising the probability exceeds the deterrence benefits. The explanation for why p^* may either increase or decrease with λ thereafter is analogous to the corresponding explanation following Proposition 2. That p^* tends toward $\hat{p}$ makes sense because, as λ grows large, it becomes inexpensive to alter p. The reason $\hat{p}$ may be less than 1 is that, even if there are no fixed enforcement costs associated with raising p, there are increased variable enforcement costs since a greater fraction of those who commit the harmful act are caught.

4. Extensions of the Model

This section discusses several extensions of the model. Because they do not substantially alter the manner in which the optimal probability of detection p^* is determined, the focus will be on how the extensions affect the optimal fine.

4.1 Investigation-Prosecution Following Detection

It was implicitly assumed that if an individual who commits a harmful act is detected, a fine would be imposed with certainty. Suppose, however, that detection is followed by a second stage during which the state investigates and prosecutes an individual, and at the end of which a fine is imposed only with a probability. Let

s = cost of the investigation-prosecution stage;

q = probability of a fine being imposed after the investigation-prosecution stage.

Hence, the probability that an individual will have to pay a fine is pq.

The investigation-prosecution costs are another type of variable enforcement cost because they depend on the number of individuals who commit harmful acts. With these costs included, expected variable enforcement costs become ps + pqk. The first term is the expected cost of the investigation-prosecution stage and the second term is the expected cost of imposing the fine.

Since an individual will now commit a harmful act if $b \geq pqf$, the optimal fine $f^*(h)$ is determined by maximizing

$$\int_{pqf}^{\infty} (b - h - ps - pqk)r(b)db \tag{7}$$

with respect to f. Differentiation yields the result that $pqf = h + ps + pqk$ if the wealth constraint on the fine is not binding, so that

$$f^*(h) = h/pq + s/q + k. \tag{8}$$

Recall that, in the absence of investigation-prosecution costs, the optimal fine was $f^*(h) = h/p + k$ (see (4)).

Formula (8) illustrates a general principle: the optimal fine equals the costs incurred by society as a result of the harmful act -- the harm itself and any variable enforcement costs -- divided by the probability, *at the time the cost is incurred*, that the injurer will have to pay the fine. Thus, h is divided by pq because, when the harm occurs, the probability of having to pay the fine is pq; s is divided by q because, when the investigation-prosecution costs are incurred, the probability of having to pay the fine is q; and k is divided by 1 because, when the cost of imposing the fine is incurred, the probability of having to pay the fine is 1.[7]

If the fine is computed according to this principle, the expected fine will equal the expected social costs due to an individual committing a harmful act, including the harm caused and the expected variable enforcement costs. Consequently, deterrence will be socially optimal.

4.2 Variable Enforcement Costs Increase with the Fine

The variable enforcement cost k of imposing fines was presumed in the analysis to be independent of the size of fines. More generally, however, it would be reasonable to suppose that the cost of imposing fines increases with their size because individuals would more strongly resist their imposition (e.g., by attempting to conceal assets).

To allow for this possibility, let the variable enforcement cost k be a function k(f) of the fine, with $k'(f) > 0$. Then, the first-order condition from differentiating (3) with respect to f becomes

$$(9) \qquad -p(pf - h - pk)r(pf) - pk'(f)(1 - R(pf)) = 0.$$

Hence, $pf = h + pk - k'(f)(1 - R(pf))/r(pf)$, or

$$(10) \qquad f^*(h) = h/p + k - k'(f)(1 - R(pf))/pr(pf) < h/p + k.$$

In other words, taking account of the possibility that higher fines result in higher variable enforcement costs lowers the optimal fine. This is because reducing the fine has the beneficial effect of reducing variable enforcement costs.

---

[7]An essentially identical formula is contained in Cohen (1988, pp. 5-7). (His article did not distinguish between fixed and variable enforcement costs and did not investigate the effects of enforcement costs on the optimal probability of detection.)

4.3 Probability of Detection Falls with the Fine

The probability of detecting individuals also may be influenced by the fine. Because individuals presumably would do more to avoid detection as the fine increases, the probability of detection would be expected to fall.[8] Hence, let the probability of detection be a function of f as well as c; since $f = f(h)$, $p = p(c, f(h))$, where $p_f < 0$. Then, given c and h, the optimal f maximizes (3), where p is treated as a function of f. The first-order condition becomes (after some substitution)

$$f^*(h) = h/p + k - [p_f k(1 - R(pf))]/[(p_f f + p)pr(pf)]. \tag{11}$$

It is evident from (11) that $f^*(h)$ exceeds $h/p + k$ if pf rises with f (since $d(pf)/df = p_f f + p$) and that $f^*(h)$ is less than $h/p + k$ if pf falls with f. Thus, taking account of the possibility that higher fines reduce the probability of detection has an ambiguous effect on the optimal fine. This is not surprising because, on one hand, a fall in the probability of detection saves variable enforcement costs, which is beneficial; but on the other hand, raising the fine may reduce the expected fine, which may be detrimental.[9]

4.4 The Sanction of Imprisonment

The sanction of imprisonment can be analyzed like the extension discussed in subsection 4.2 in which the cost of imposing fines rises with their level. Here the cost of imprisonment rises with its duration. In particular, let

z = length of imprisonment;

---

[8]Another possibility is that the probability of detection rises with the fine. For example, since the number of individuals who commit harmful acts falls with the fine, the chance that any one of them is detected might rise.

[9]An alternative explanation of why the optimal fine might be higher or lower than $h/p + k$ is as follows. First, rewrite (11) in the form $-p_f k(1 - R(pf)) = (p_f f + p)r(pf)(pf - h - pk)$. The left-hand side is the marginal benefit of raising f, which is positive, and consists of the reduction in variable enforcement costs due to the fall in p. The right-hand side is the marginal cost of raising f, and must be positive since the marginal benefit is positive. If $d(pf)/df = p_f f + p$ is positive, then fewer individuals commit the harmful act as f increases, so for the marginal costs to be positive, each such individual must have been adding to social welfare. Because pf is the benefit of the marginal individual and $h + pk$ is the increase in expected social costs due to his committing the act, this means that $pf - h - pk$ must be positive or, equivalently, $f > h/p + k$. Similarly, if $p_f f + p$ is negative, $pf - h - pk$ must be negative, or $f < h/p + k$.

σ = per unit cost to the state of imposing the imprisonment sanction; $\sigma > 0$.[10]

It will be assumed that the disutility to an individual of an imprisonment term of length z equals z.

Since an individual will commit the harmful act if $b \geq pz$, the optimal sanction $z^*(h)$ maximizes

(12) $$\int_{pz}^{\infty} (b - h - p\sigma z)r(b)db$$

with respect to z. The first-order condition from (12) is, after dividing by p,

(13) $$-(pz - h - p\sigma z)r(pz) - \sigma(1 - R(pz)) = 0.$$

This condition can be rewritten as

(14) $$z^*(h) = h/p + \sigma z - \sigma(1 - R(pz))/pr(pz) < h/p + \sigma z,$$

which is similar to (10). As with fines in subsection 4.2, the optimal imprisonment sanction depends directly on variable enforcement costs, but not all such costs are added to the sanction because reducing the length of imprisonment has the beneficial effect of reducing variable enforcement costs. Note also that the optimal imprisonment sanction z^* does not depend directly on the fixed enforcement costs c required to maintain the probability of detection (although z^* is influenced indirectly through the effect of c on p^*).

4.5 Individuals Bear Costs in Connection with Enforcement

Thus far it has been assumed that all of the costs of the enforcement process are borne by the state. In practice, of course, individuals who commit harmful acts also bear costs in connection with enforcement -- in evading detection, defending against conviction, or contesting the level of fines.

Such costs do not affect the formula for optimal fines for the simple reason that individuals will properly take these costs into account because they bear them. To demonstrate this, let

a_1 = cost of trying to evade detection borne by each individual who commits the harmful act;

---

[10]One could also interpret σ as including the cost borne by the individual punished.

a_2 = cost of trying to avoid paying the fine borne by each individual who is detected.

These costs affect social welfare (2) and the expression for determining the optimal fine (3) in two ways. First, since an individual now will commit the harmful act if $b \geq a_1 + p(f + a_2)$, the latter expression replaces pf as the lower limit of integration with respect to b. Second, the part of the integrand representing the benefit to the individual less the harm and less enforcement costs now becomes $(b - h - a_1 - p(k + a_2))$. It is easily verified that these changes do not affect the first-order condition from maximizing (3), so that it is still true that $f^*(h) = h/p + k$, as claimed.[11]

Although the enforcement costs borne by individuals do not affect the formula for optimal fines $f^*(h)$, these costs do affect the optimal level of enforcement effort c^*, and thus p^*. Presumably, as in Proposition 2, p^* initially may increase or decrease as a_1 or a_2 rises, but p^* eventually must decrease to zero as a_1 or a_2 becomes large.

5. Concluding Remarks

This section discusses the magnitude of enforcement costs, the ability to apply the formulas in this article, and the incorporation of enforcement costs in actual sanctions.

(a) *Importance of enforcement costs.* Enforcement costs often are not inconsequential relative to the sanctions imposed. For example, the enforcement costs of the Internal Revenue Service in its screening program to detect discrepancies between income reported on tax returns and income reported by payers amount to approximately 10% of the additional revenue collected; and its enforcement costs in its examination program to audit returns are in the 20-35% range for most categories of taxpayers.[12]

[11]However, this result does not hold if the costs borne by individuals rise with the magnitude of fines (because individuals spend more resisting their imposition). In that case, it can be shown that $f^*(h) < h/p + k$ for reasons analogous to those in subsection 4.2. We are grateful to Jeffrey Parker for suggesting this point.

[12]The Information Returns Program of the IRS, which screens returns to detect discrepancies, cost $54 million in FY 1981 and yielded $530 million in additional revenues, resulting in a cost/yield of 10.2%. See President's Private Sector Survey on Cost Control (1983, p. 79). The examination program of the IRS, which engages in audits of returns, had the following average costs per examination, additional yields, and resulting cost/yield percentages in FY 1985: (a) for individuals earning between $25,000 and $50,000, cost per return $235, additional yield $678, cost/yield 34.7%; (b) for non-farm

The enforcement costs of the Securities and Exchange Commission in its fraud enforcement program are about 10% of total penalties and disgorgements.[13] In general, it seems plausible that the enforcement costs associated with many minor violations, such as parking and traffic violations, are large in relation to the fines.[14] And, presumably, when penalties are high, the state's prosecution, settlement, and collection costs often will be substantial because of defendants' strong incentives to resist incurring such penalties.

(b) *Practical ability to use the formulas.* The formulas in our article are not difficult to employ. All that an enforcement agency would need to do is to add to the fine that it otherwise considers appropriate the costs of investigation and prosecution, multiplied by the reciprocal of the probability that a fine will be imposed after investigation and prosecution, plus the cost of imposing the fine itself. (More generally, the agency should include any type of variable enforcement cost, multiplied by the reciprocal of the probability -- at the time the cost is incurred -- that the injurer will have to pay the fine.) Thus, the information needed by an agency is simply its own enforcement data. For instance, suppose that an enforcement agency knows the following: only one-third of the cases it investigates and prosecutes result in the imposition of a fine; the average cost of investigation-prosecution is $1,000; and the average cost of collecting the fine is $500. Then the amount that the agency should add to the fine to reflect enforcement costs is ($1,000 x 3) + $500, or $3,500.

(c) *Inclusion of enforcement costs in actual sanctions.* There presently are circumstances in which enforcement costs can be included in penalties. For example,

---

businesses earning over $100,000, cost per return $1,224, additional yield $5,187, cost/yield 23.6%; and (c) for corporations earning between $1 million and $5 million, cost per return $2,534, additional yield $12,383, cost/yield 20.5%. See Steuerle (1986, p. 28).

[13]Total SEC expenditures on prevention and suppression of fraud were $46.6 million in 1989 and yielded $482.7 million, so the percentage of costs to yield was 9.7%. Since the SEC obtained unusually high amounts from defendants in 1989, this cost/yield percentage was much lower than normal. This information was obtained from an unpublished SEC document entitled "Budget Estimate Fiscal 1991," pp. II-i & II-10.

[14]Consider, for example, the enforcement costs associated with a $10 fine for an expired parking meter. A ticket has to be written and recorded in a central information system; the check the person sends must be cashed and a record made of that. If the total time involved is even 6 minutes and the average hourly wage plus fringe benefits is $20, the processing cost would be $2, or 20% of the $10 fine.

federal district courts are permitted to add the costs of prosecution to the sanctions they would otherwise impose.[15] This provision, however, does not inflate the prosecution costs to account for the fact that the imposition of a sanction is not certain, and it ignores other components of enforcement costs (such as the cost of collecting a fine).

A second example is that the United States Sentencing Commission recently has required the inclusion of certain enforcement costs in penalties. The sentencing guidelines applicable to individuals provide that "the court shall impose an additional fine amount that is at least sufficient to pay the costs to the government of any imprisonment, probation, or supervised release ordered."[16] According to our analysis, it is appropriate that these costs are not inflated, because they are borne only if the defendant is convicted. However, the sentencing guidelines omit other kinds of enforcement costs that should be included, such as the costs of investigation and prosecution.

Other, more particular, examples of the inclusion of enforcement costs in sanctions may be given. Three such examples, one federal, one state, and one local, are:[17] The United States Department of the Interior's natural resource damage assessment procedures make a polluter pay for the government's cost of determining the pollution damages. The Massachusetts Department of Environmental Quality Engineering has proposed a plan to bill owners of polluted property for the regulators' time, at $67 an hour. And several police departments in the San Francisco area make it a policy to charge drunken drivers for the time officers spend arresting them.

These examples serve to illustrate the practicality of including enforcement costs in sanctions.

---

[15]Section 1918(b) of Title 28 of the United States Code states: "Whenever any conviction for any offense not capital is obtained in a district court, the court may order that the defendant pay the costs of prosecution." See generally O'Malley (1987).

[16]See United States Sentencing Commission (1989, p. 5.20). This provision may not be applicable if the defendant cannot pay all or part of the fine, or if the payment would unduly burden the defendant's dependents.

[17]The examples listed below are based on the following sources: State of Ohio v. U.S. Depart. of the Interior, 880 F.2d 432, 468 (D.C. Cir 1989); William M. Bulkeley, "Massachusetts Wants to Bill Owners Of Polluted Land for Regulators' Time," *Wall Street Journal*, January 10, 1989, p. A20; "Editorial: Paying Crime's Price," *Peninsula Times Tribune*, February 25, 1989, p. A-9.

Appendix

Proof of Proposition 2. In this proof, let $p^*(k)$ denote the optimal p given k, and consider a k such that $p^*(k) > 0$. To see that $p^{*\prime}(k)$ may be positive or negative, first subtract the right-hand side of (6) from the left-hand side. The resulting equation is of the form $W(c, k) = 0$, where W is the derivative of (5) with respect to c. $W(c, k) = 0$ determines c as a function of k, and differentiating $W(c, k) = 0$ with respect to k and solving for $c'(k)$, we obtain $c'(k) = -W_k/W_c$. Since $W_c < 0$ (the second-order condition for (5) to be maximized), the sign of c' equals the sign of W_k. But it is straightforward to verify that the sign of W_k can be positive or negative. Hence, c^*, and thus p^*, may either rise or fall as k increases.

It remains to show that $p^*(k) = 0$ for all k sufficiently large. Let us first establish that $p^*(k) \to 0$ as $k \to \infty$. If this were not true, there would be an $\varepsilon > 0$ and an increasing sequence k_i, where $k_i \to \infty$ as $i \to \infty$, such that $p^*(k_i) > \varepsilon$ for every i. Note that there will always be some fraction of the population who commit harmful acts: the highest possible expected fine is bounded from above by w, so at least $1 - R(w) > 0$ of the population will commit such acts. Hence, the variable enforcement costs are at least $(1 - R(w))\varepsilon k_i$. But this grows unboundedly as $k_i \to \infty$. In contrast, it is always possible to choose $c = 0$, in which case, since everyone commits harmful acts, social welfare is $E(b) - E(h)$, where E stands for expected value. This level of welfare must be higher than welfare under the $p^*(k_i)$ for large k_i, which contradicts the optimality of the $p^*(k_i)$.

To complete the argument, it will suffice to show that for all k sufficiently high, the derivative of social welfare with respect to c is negative for all c in a neighborhood $[0, \delta]$ of 0. For this fact and the fact that $p^*(k) \to 0$ imply that $p^*(k) = 0$ for all k sufficiently high. Now, for all k large, w is less than k, so that the optimal fine is w for all h, and social welfare is

$$\int_0^\infty \int_{pw}^\infty (b - h - pk)r(b)db\,g(h)dh - c, \tag{15}$$

rather than (5). The derivative of this with respect to c is

$$-p'(c)w(pw - E(h) - pk)r(pw) - p'(c)k(1 - R(pw)) - 1. \tag{16}$$

This expression involves terms not including k plus

$$kp'(c)[pwr(pw) - (1 - R(pw))]. \tag{17}$$

As c approaches 0, (17) tends to kp'(0), which falls without bound as k increases. Hence, (16) must be negative within a suitably small interval $[0, \delta]$ for all k sufficiently large.

Proof of Proposition 3. Since $p = p(\lambda c)$, $dp/dc = \lambda p'(\lambda c)$. Hence, the derivative of (5) with respect to c is

$$r(pw)\lambda p'(\lambda c)w \int_{p(w-k)}^{\infty} (h + pk - pw)g(h)dh$$

$$(18) \qquad - 1 - \lambda p'(\lambda c)k\Big[\int_{0}^{p(w-k)} (1 - R(h + pk))g(h)dh$$

$$+ (1 - R(pw))(1 - G(p(w - k)))\Big].$$

This is less than or equal to $-1 + \lambda m$, where m is the maximum over c of the terms multiplying λ. (It is easily seen that the maximum exists.) Hence, if λ is sufficiently small, (18) is negative for all c, implying that $c^* = 0$.

When c^* is positive, the sign of the derivative of (18) with respect to λ is ambiguous, so that c^* may rise or fall with λ.

As λ grows large, the cost of setting p equal to $\hat{p}$ becomes arbitrarily small. Therefore, p^* must tend toward $\hat{p}$ as $\lambda \rightarrow \infty$.

$\hat{p}$ may be an interior solution -- that is, be less than 1 -- in which case it is determined by differentiating the first two terms of (5) with respect to p and setting the result equal to 0; this yields

$$r(pw)w \int_{p(w-k)}^{\infty} (h + pk - pw)g(h)dh$$

$$(19) \qquad = k\Big[\int_{0}^{p(w-k)} (1 - R(h + pk))g(h)dh$$

$$+ (1 - R(pw))(1 - G(p(w - k)))\Big].$$

References

Becker, Gary S., "Crime and Punishment: An Economic Approach," *Journal of Political Economy*, Vol. 76, No. 2 (March/April 1968), pp. 169-217.

Cohen, Mark A., "Optimal Criminal Fines for Organizations," A Report to the U.S. Sentencing Commission, Draft, January 11, 1988.

O'Malley, Kevin F., "The Assessment of Costs in Federal Criminal Prosecutions," *St. Louis University Law Journal*, Vol. 31, No. 4 (October 1987), pp. 853-874.

Polinsky, A. Mitchell, and Steven Shavell, "Pigouvian Taxation with Administrative Costs," *Journal of Public Economics*, Vol. 19, No. 3 (December 1982), pp. 385-394.

President's Private Sector Survey on Cost Control, *Report on the Department of the Treasury* (Washington, D.C.: U.S. Government Printing Office, 1983).

Shavell, Steven, "Specific versus General Enforcement of Law," Discussion Paper No. 58, Program in Law and Economics, Harvard Law School, June 1989.

Steuerle, C. Eugene, *Who Should Pay for Collecting Taxes? Financing the IRS* (Washington, D.C.: American Enterprise Institute, 1986).

Stigler, George J., "The Optimum Enforcement of Laws," *Journal of Political Economy*, Vol. 78, No. 3 (May/June 1970), pp. 526-536.

United States Sentencing Commission, *Guidelines Manual* (November 1, 1989).