

Optimal Sanctions and Differences in Individuals' Likelihood of Avoiding Detection

LUCIAN ARYE BEBCHUK AND LOUIS KAPLOW*

*Harvard Law School and the National Bureau of Economic Research,
Cambridge, MA 02138, USA*

1. Introduction

Gary Becker (1968) argued that society can economize on enforcement resources by reducing the probability of apprehension while increasing sanctions, thereby maintaining the same deterrent effect at reduced cost. The implication is that optimal sanctions are maximal. This insight has been explored and qualified in a variety of subsequent work.¹

This paper considers the implications for optimal enforcement policy of the possibility that individuals are not all equally easy to apprehend. Upon reflection, it seems clear that the probability of apprehension will depend not only on enforcement effort, but also on particular characteristics of the actor and the situation. For example, experienced actors may be more difficult to detect because of their expertise or easier to identify because of their track record (e.g., fingerprints on file), and an act may have been committed in front of witnesses who know the actor or when no one else was present.

The factors that make an individual easy or difficult to apprehend may be known to the individual but not to the government, particularly before the government expends some enforcement effort or apprehends the individual. Thus, when choosing its level of enforcement effort for a given act, the government may not know the experience of the actor or the identity of witnesses (if any). In such cases, it will not know in advance the effectiveness of a given level of enforcement effort. Rather, it will know, based on experience, only the portion of all committing such an act who will be apprehended. As a result, individuals committing acts that appear the same to the government will face different probabilities of apprehension. We analyze op-

*We thank Howard Chang, Jesse Fried, Marcel Kahan, A. Mitchell Polinsky, Steven Shavell, the editors, and two anonymous referees for helpful comments and the John M. Olin Foundation for financial support. Lucian Bebchuk's work has also been supported by the National Science Foundation.

¹The literature has largely focused on offering explanations for why and when the use of maximal sanctions may be inefficient: risk-aversion (Kaplow, 1992; Polinsky and Shavell, 1979), nonmonetary sanctions (Kaplow, 1990b; Polinsky and Shavell, 1984), avoidance costs (Malik, 1990), marginal deterrence (Shavell, 1992), general enforcement (Mookherjee and Png, 1992; Shavell, 1991), wealth differences (Polinsky and Shavell, 1991), individuals' imperfect information about whether acts are subject to sanctions (Kaplow, 1990a), and individuals' imperfect observation of the probability of apprehension (Bebchuk and Kaplow, 1992).

timal enforcement in such circumstances and demonstrate that optimal sanctions may be less than maximal.

We find it useful to examine three cases. First, we consider the possibility that the enforcement authority knows how difficult an actor would be to identify and apprehend before any resources have to be expended. (We assume throughout that individuals know their own probability of apprehension.) In this case, the problem reduces to the situation where individuals are homogeneous because the probability of apprehension and sanction may be tailored to each actor. Thus, the optimal fine would be maximal and the optimal probability of apprehension would reflect marginal enforcement costs and marginal deterrence gains.

Next, we assume that the difficulty of apprehension cannot be determined until after an individual is apprehended. This case is of interest because the process of apprehension or knowledge of the actor's identity may provide substantial information regarding the *ex ante* probability of apprehension. For individuals who are of types that are difficult to apprehend, it will be optimal to apply the maximum feasible fine, because they will be underdeterred at an optimum. For individuals who are easy to apprehend, however, it may well be optimal to apply lower fines to avoid overdeterrence.

Finally, we examine the case in which, even after the fact, the enforcement authority cannot determine an actor's type. Presumably, some differences in skill or circumstances often will remain unknown to the authority, even after the individual is apprehended. In this case, a single sanction must be applied to individuals who are not equally easy to apprehend, so types that are easier to detect may be overdeterred and those difficult to detect will be underdeterred. As a result, it may be optimal to reduce the fine from its maximum feasible level while raising enforcement effort. The reduced fine will cause a greater reduction in the expected sanction for types that are easy to detect (because their probability of apprehension is higher), while the greater enforcement effort may have a greater effect on types who are more difficult to apprehend.

After presenting our model and analysis for each case, we offer brief concluding remarks concerning the importance of our results, how they relate to common intuitions concerning the undesirability of extreme sanctions, and directions for future research.

2. The Model

A. Framework for Analysis

Risk-neutral individuals may commit an act that imposes a social cost of h . The benefit to the actor, b , varies among individuals according to the distribution $f(\cdot)$, which is positive on $[0, \bar{b}]$; the cumulative distribution function is $F(\cdot)$. We assume that \bar{b} is sufficiently large that some acts are socially beneficial, as the desire to avoid overdeterrence will be an important feature of the analysis.²

The probability of apprehension is determined jointly by individuals' characteristics (which they know), θ , and the government's expenditure on enforcement in re-

²As will be apparent from Proposition 2, this requires that \bar{b} exceed the sum of the direct social cost of the act and the enforcement cost per act.

sponse to each act, x .³ We assume for simplicity that θ is distributed uniformly on $[0, 1]$ and independently of b . The resulting probability of apprehension is $p(x, \theta)$, where $p_x > 0$ and $p_{xx} < 0$; that is, there are diminishing returns to enforcement expenditures. In addition, for $x > 0$, $p_\theta > 0$; that is, a higher θ indicates that an individual is more likely to be apprehended for a given enforcement expenditure.

Finally, after an individual is apprehended, the government chooses the sanction, s , which cannot exceed \bar{s} . We interpret the sanction as a fine that is socially costless to impose, while the maximum feasible fine may be interpreted as individuals' wealth.

Individuals will choose to act whenever their benefit, b , exceeds their expected sanction, $p(x, \theta)s$. The optimal enforcement policy is defined as that which maximizes the sum of the benefits from acts minus the harm caused by acts and enforcement costs. We examine optimal enforcement for three cases:

- The government observes individuals' types, θ , ex ante—that is, before apprehension resources are expended. In this case, both x and s may differ among individuals of different types.
- The government observes θ ex post—that is, after individuals are apprehended. In this case, s may differ among individuals but x must be the same for everyone.
- The government does not observe θ . In this case, both x and s must be identical for everyone.

B. Individuals' Types Observable Ex Ante

The government can individualize x and s for each type. Thus, one can solve the government's optimization problem separately for each θ . Because this problem is qualitatively the same as one involving no heterogeneity with regard to apprehension, the familiar Becker argument applies.

For any θ , the government chooses x and s to maximize

$$\int_{p(x, \theta)s}^{\bar{s}} (b - h)f(b)db - x(1 - F(p(x, \theta)s)), \quad (1)$$

subject to the constraint that $s \leq \bar{s}$. Let s^* denote the optimal sanction.

Proposition 1: *When the government observes how difficult individuals are to apprehend (θ) before expending enforcement resources, the optimal sanction for all individuals is maximal. That is, $s^* = \bar{s}$ for all θ .*

Proof: As one increases s and reduces x so as to keep $p(x, \theta)s$ constant, the first term in (1) is unaffected and the magnitude of the second term falls [x falls and $F(p(x, \theta)s)$ does not change], so the optimum is where the constraint is binding. *Q.E.D.*

³This model employs an investigation technology, in which the government makes expenditures in response to individual acts, rather than the monitoring technology employed in much of the law enforcement literature, in which the government's single expenditure (independent of the number of acts) determines the detection rate for all acts. We choose an investigation technology because in our problem it is natural to consider how much the government would spend after observing acts, and our first result depends on the government's being able to spend different amounts for different acts it observes. For a discussion of different enforcement technologies, see Mookherjee and Png (1992) and Shavell (1991).

C. Individuals' Types Observable Ex Post

In this case, the government can individualize s but not x . Thus, the government chooses x and $s(\theta)$ to maximize

$$\int_0^1 \left[\int_{p(x,\theta)s(\theta)}^{\bar{b}} (b-h)f(b)db - x(1 - F(p(x,\theta)s(\theta))) \right] d\theta, \quad (2)$$

subject to the constraint that $s(\theta) \leq \bar{s}$ for all θ . Let $s^*(\theta)$ denote the optimal sanctioning function and x^* denote the optimal enforcement expenditure.

Proposition 2: *When the government observes how difficult individuals are to apprehend (θ) only after expending enforcement resources, the optimal sanction is maximal only for individuals who are the most difficult to apprehend. In particular,*

$$\begin{aligned} s^*(\theta) &= \bar{s}, \text{ if } p(x^*,\theta)\bar{s} \leq h + x^*, \text{ and} \\ s^*(\theta) &= (h+x^*)/p(x^*,\theta), \text{ if } p(x^*,\theta)\bar{s} > h + x^*. \end{aligned}$$

Proof: For each θ , the first-order condition⁴ for the optimal sanction is

$$\begin{aligned} (h - p(x^*,\theta)s(\theta))p(x^*,\theta)f(p(x^*,\theta)s(\theta)) + x^*p(x^*,\theta)f(p(x^*,\theta)s(\theta)) \geq 0, \text{ or} \\ h + x^* - p(x^*,\theta)s(\theta) \geq 0, \end{aligned} \quad (3)$$

where the expression holds as an equality unless the constraint is binding. *Q.E.D.*

The intuition of Proposition 2 is as follows. Ideally, the expected sanction should equal the expected harm. The expected harm of an act includes the social cost of the act itself, h , and the cost that the social authority will expend attempting to apprehend the actor, x^* . Thus, for any θ it is optimal for the government to select $s(\theta)$ so that $p(x^*,\theta)s(\theta) = h + x^*$.⁵ This will be feasible for all individuals who are sufficiently easy to apprehend at x^* . For individuals who are more difficult to apprehend (types for whom θ falls below a critical level, recalling that $p_\theta > 0$), underdeterrence is unavoidable, so it is optimal to set the sanction as high as possible.

A further increase in sanctions for those who are easier to apprehend would be desirable only if the enforcement effort were correspondingly reduced, but any reduction in effort will worsen the problem of underdeterrence for those who are more difficult to apprehend. The optimal level of enforcement resources, x^* , thus will reflect a balance between greater enforcement costs and reducing the problem of underdeterrence for difficult-to-apprehend individuals.⁶

⁴The second-order condition holds.

⁵The optimal sanction would not equal, but would be less than, expected harm in a model involving marginal deterrence—that is, when individuals choose among harmful acts. See Shavell (1992).

⁶The result in this case—where the government does not observe types ex ante but does ex post—is similar in spirit to that in Shavell (1991) concerning optimal sanctions when enforcement effort is “general”—that is, when enforcement effort affects the probability of apprehension for acts of differing harm. (See also Mookherjee and Png, 1992.) More harmful acts are underdeterred at the optimum and thus should be subject to the maximal sanction, while less harmful acts should be subject to lesser sanctions that achieve optimal deterrence. Also related is Polinsky and Shavell (1991). They examine variations in individuals' wealth, where wealth is the upper limit on the maximum feasible fine. In their model, individuals with low wealth are just as easy to apprehend but are underdeterred at the optimum because they can only pay low fines. Raising the probability of apprehension reduces this underdeterrence.

D. *Individuals' Types Unobservable*

In this case, the government must select the same x and s for all individuals. It chooses x and s to maximize

$$\int_0^1 \left[\int_{p(x,\theta)s}^{\bar{s}} (b - h)f(b)db - x(1 - F(p(x,\theta)s)) \right] d\theta, \quad (4)$$

subject to the constraint that $s \leq \bar{s}$.

To consider whether the optimal sanction is maximal, we ask whether welfare rises or falls when the maximal sanction is reduced slightly, with the probability raised in a corresponding manner. Let \bar{x} be the optimal enforcement effort when the sanction is \bar{s} . In particular, we decrease s and increase x such that $p(x,\hat{\theta})s = p(\bar{x},\hat{\theta})\bar{s}$, where $\hat{\theta}$ is such that $p(\bar{x},\hat{\theta})\bar{s} = h$ —that is, so that $ds/dx = -sp_x(x,\hat{\theta})/p(x,\hat{\theta})$.⁷ The derivative of expression (4) with respect to x with s stipulated to change in the prescribed manner is as follows:

$$\begin{aligned} & \int_0^1 (h - p(x,\theta)s)f(p(x,\theta)s)s(p_x(x,\theta) - p_x(x,\hat{\theta})\frac{p(x,\theta)}{p(x,\hat{\theta})})d\theta \\ & + xs \int_0^1 (p_x(x,\theta) - p_x(x,\hat{\theta})\frac{p(x,\theta)}{p(x,\hat{\theta})})f(p(x,\theta)s)d\theta - (1 - \int_0^1 F(p(x,\theta)s)d\theta). \end{aligned} \quad (5)$$

The first term indicates the effect on welfare resulting from the change in behavior—isolating the harm and benefits of acts, independent of any effect on enforcement costs. The sign of this term is determined by the sign of $p_{x\theta}$. For the case in which $p_{x\theta} < 0$ for all θ when evaluated at $x = \bar{x}$, this term is positive: The left and right factors are both negative for those who are overdeterred ($\theta > \hat{\theta}$) and they are both positive for those who are underdeterred.⁸ This case corresponds to the one in which increased enforcement effort has a relatively greater effect on those more difficult to apprehend (those with low θ).

The remaining terms indicate the change in enforcement costs. The second term reflects the increase or decrease in enforcement costs resulting from the fact that different individuals act. (It parallels the first term: When $p_{x\theta} < 0$, as we assume here, the underdeterred commit fewer acts and the overdeterred commit more acts; the net effect on enforcement costs is ambiguous.) The final term indicates the added enforcement cost per act committed. The total effect on enforcement costs is ambiguous, but they may increase.

In general, which effect dominates will depend on the strength of the behavioral effect (indicated most directly by the magnitude of $p_{x\theta}$) and the enforcement technology. Examples can be constructed in which the behavioral effect dominates; then,

⁷The manipulation assumes that such a $\hat{\theta}$ exists, which it must unless all types are overdeterred or all are underdeterred. This assumption is used to illuminate the intuition, but it is not necessary for the proposition to be valid. The example in note 9, which constitutes a proof, does not rely on the assumption.

⁸The left factor is obvious by inspection, and the right factor follows from the assumption that $p_{x\theta} < 0$.

the derivative is positive, demonstrating that welfare may be greater with less than the maximal sanction.⁹ Thus, we have the following result.

Proposition 3: *When the government does not observe how difficult individuals are to apprehend (θ) either before or after expending enforcement resources, the optimal sanction (which is the same for all individuals) may be less than \bar{s} .*

The intuition of proposition 3 relates to that of proposition 2. In the earlier case, in which the sanction could be tailored to individuals' types, overdeterrence of those more easily apprehended could be avoided. When a single sanction must be applied, the optimum involves a compromise between the problem of overdetering those readily detected and underdetering those more difficult to detect.

In this case, reducing the sanction has the desirable feature that it reduces the expected sanction more for the readily apprehended types, because their probability of apprehension is greater. Whether increasing enforcement effort to maintain the same average expected sanction will improve behavior depends on $p(x, \theta)$. The preceding analysis invoked the assumption that $p_{x\theta} < 0$: Increased effort causes a greater increase in the probability of apprehension for those who are more difficult to apprehend (and thus who have a lower probability of apprehension). When this assumption holds, the combination of the lower sanction and higher probability reduces the expected sanction for those readily apprehended—who were overdeterred—and increases the expected sanction for those more difficult to apprehend—who were underdeterred. If this benefit with regard to behavior is greater than any required increase in expenditures on enforcement, a lower sanction is desirable. On

⁹Constructing an example is complicated by the fact that \bar{x} , optimal enforcement effort at \bar{s} , is itself affected by the technology as well as by the fact that assuming $p_\theta > 0$ rules out the possibility that $p_{x\theta} < 0$ throughout (assuming that $p(0, \theta) = 0$). Consider the following enforcement technology:

$$\text{For } x < \varepsilon: p(x, \theta) = \frac{x}{\varepsilon}(\alpha + \beta\theta).$$

$$\text{For } x \geq \varepsilon: p(x, \theta) = \alpha + \beta\theta + \frac{x - \varepsilon}{x - \varepsilon + 1/k}(1 - \alpha - \beta\theta).$$

As one makes ε arbitrarily small and k arbitrarily large, the character of this technology is such that at an arbitrarily low cost one has all the relevant properties satisfied, as well as the requirements that now follow. In particular, choose α , β , and an $\bar{s} > \bar{b}$ such that at $x = \varepsilon$ no individual is overdeterred in the sense that $p(\varepsilon, 1)\bar{s} \leq h$. (Note that this relationship will be preserved as ε is made smaller.)

Consider the level of enforcement effort that maximizes the value of the first term in expression (4), given \bar{s} . Clearly, the value of this term will be less than its first-best value in the absence of enforcement costs—that is, when all types of individuals act if and only if their benefit exceeds h . (For any given s , if some type acts only when $b > h$, easier-to-apprehend types are overdeterred and harder-to-apprehend types are underdeterred.) Let the difference in these values be ΔW .

This first-best value of the first term in (4) can be approximated arbitrarily closely, say, within an amount $\delta < \Delta W$, by some values, \bar{x} and \bar{s} . (In particular, if one chooses $\bar{s} = h$, as x approaches infinity, first-best behavior can be approximated arbitrarily closely.) Now, as one allows ε to approach zero and k to approach infinity, ΔW is unaffected, and for the same \bar{s} the requisite \bar{x} becomes arbitrarily small—in particular, less than $\Delta W - \delta$. The result is that for this technology the policy (\bar{x}, \bar{s}) improves the value of the first term in (4) by more than the total enforcement costs it incurs, so welfare must be greater than at (\bar{x}, \bar{s}) .

the other hand, if $p_{x_0} > 0$, reducing the sanction and raising enforcement effort would make behavior worse.¹⁰

3. Concluding Remarks

We have examined how optimal sanctions are affected by the fact that not all individuals are equally easy to apprehend. When the government can observe all differences before expending enforcement resources, the result is the familiar one that the optimal sanction is maximal, because an optimal probability can be set for each actor. When the government can observe the difficulty of apprehension only after individuals are apprehended, the character of the optimal sanction changes. For those most difficult to apprehend, a maximal sanction will be optimal, while for those relatively easy to apprehend, a lower sanction will be optimal in order to avoid overdetering them. Finally, when the government cannot observe the difficulty of apprehending individuals even after they are apprehended, a single sanction must be applied. The optimal sanction may be less than maximal because reducing the actual sanction reduces the expected sanction more for those who are easiest to apprehend and thus may be overdeterred than for those who are more difficult to apprehend and thus are underdeterred.

These results are of obvious practical significance because the government rarely will have complete knowledge of the difficulty of apprehending an actor before any enforcement activity takes place. For example, higher sanctions might be appropriate when it can be determined after apprehension that an act was premeditated because such actors generally would have more opportunity to hide their tracks. Similarly, additional penalties might be applied when individuals destroy evidence. It also might be the case that an individual's prior record would constitute information about the difficulty of apprehension, in which case sanctions should be different for repeat offenders.¹¹ Of course, even after apprehension many circumstances of the act will never become known to enforcement authorities. This might help explain why less than maximal sanctions are the rule rather than the exception.

Our analysis may be related to a common intuition that applying maximal sanctions with a very low probability is unfair. When enforcement is probabilistic, those apprehended may worry that they were treated unequally to others who committed

¹⁰It seems plausible that either assumption could hold. (Also, p_{x_0} need not have the same sign for all θ .) While for very low x it seems most plausible that $p_{x_0} > 0$ —that is, small initial enforcement effort would have the greatest effect on those easy to apprehend—it also seems plausible that $p_{x_0} < 0$ for greater levels of enforcement effort, as diminishing returns might be more severe for those easiest to apprehend. (As the probability for those easy to apprehend approaches 1, it must be the case that marginal returns approach zero.)

¹¹Perhaps those caught repeatedly are those who are easy to apprehend, or those whose prior record (fingerprints) makes them easier to apprehend, in which case a lower penalty would be justified. Alternatively, repeat offenders may have additional knowledge of the enforcement system or be those who find the activity more profitable because they are intrinsically more difficult to apprehend (and thus they engage in the activity frequently), in which case a higher penalty would be optimal. (Related, many first offenders may be individuals who were easy to catch, but did not know this initially; after learning from their first adverse experience, they may no longer engage in the activity, so that repeat offenders would be primarily those who are more difficult to apprehend.) For a discussion of optimal penalties for repeat offenders, see Polinsky and Rubinfeld (1991).

the same act but were not apprehended. The standard response is that treatment is equal because all actors faced the same expected sanction *ex ante*. But it might be difficult for apprehended individuals to verify that they were subject to the same probability of apprehension as those who were not apprehended. If, as in our model, individuals' probabilities differ and they know this, some of those apprehended will in fact know that they faced a higher expected sanction.¹²

Our model is designed to examine one important way in which actors may differ: their actual probability of apprehension. Useful further research might consider other differences.¹³

References

- BECHUK, LUCIAN A. AND LOUIS KAPLOW, "Optimal Sanctions When Individuals Are Imperfectly Informed About the Probability of Apprehension," *Journal of Legal Studies* 21, 365–370 (1992).
- BECKER, GARY S., "Crime and Punishment: An Economic Approach," *Journal of Political Economy* 76, 169–217 (1968).
- KAPLOW, LOUIS, "Optimal Deterrence, Uninformed Individuals, and Acquiring Information about Whether Acts Are Subject to Sanctions," *Journal of Law, Economics, and Organization* 6, 93–128 (1990a)
- , "A Note on the Optimal Use of Nonmonetary Sanctions," *Journal of Public Economics* 42, 245–247 (1990b).
- , "The Optimal Probability and Magnitude of Fines for Acts that Definitely Are Undesirable," *International Review of Law and Economics* 12, 3–11 (1992).
- MALIK, ARUN S., "Avoidance, Screening and Optimum Enforcement," *Rand Journal of Economics* 21, 341–353 (1990).
- MOOKHERJEE, DILIP AND I. P. L. PNG, "Monitoring vis-à-vis Investigation in Enforcement of Law," *American Economic Review* 82, 556–565 (1992).
- POLINSKY, A. MITCHELL AND DANIEL L. RUBINFELD, "A Model of Optimal Fines for Repeat Offenders," *Journal of Public Economics* 46, 291–306 (1991).
- AND STEVEN SHAVELL, "The Optimal Tradeoff Between the Probability and Magnitude of Fines," *American Economic Review* 69, 880–891 (1979).
- AND ———, "The Optimal Use of Fines and Imprisonment," *Journal of Public Economics* 24, 89–99 (1984).
- AND ———, "A Note on Optimal Fines When Wealth Varies Among Individuals," *American Economic Review* 81, 618–621 (1991).
- SHAVELL, STEVEN, "Specific versus General Enforcement of Law," *Journal of Political Economy* 99, 1088–1108 (1991).
- , "A Note Marginal Deterrence," *International Review of Law and Economics* 12, 345–355 (1992).

¹²Indeed, those apprehended would systematically tend to have a higher *ex ante* expected sanction than those who are not, because those who are easier to catch will be caught more often.

¹³For example, even individuals facing the same probability of apprehension and sanction may vary in how informed they are about the probability or sanction. Bebchuk and Kaplow (1992) consider individuals who imperfectly observe the probability of apprehension.