

The Fairness of Sanctions: Some Implications for Optimal Enforcement Policy

A. Mitchell Polinsky, *Stanford Law School*, and Steven Shavell,
Harvard Law School

In this article we incorporate notions of the fairness of sanctions into the standard model of public enforcement. When both the probability and magnitude of sanctions may be varied, the usual solution involves a very high sanction and a relatively low probability of enforcement if individuals are risk neutral. When the issue of fairness is added to the analysis, the optimal sanction generally is not extremely high because such a sanction would be seen as unfair. The optimal probability of imposing sanctions may be higher than in the usual case (to offset the lower sanction) or lower than in the usual case (because the lower sanction reduces the effectiveness of enforcement).

1. Introduction

In the standard model of public enforcement of law, sanctions are evaluated in terms of their deterrent benefits and their costs, but there is no

Both authors also are research associates of the National Bureau of Economic Research. Polinsky's research was supported by the Center for Advanced Study in the Behavioral Sciences (through National Science Foundation grant #SBR-960123) and the John M. Olin Program in Law and Economics at Stanford Law School. Shavell's research was supported by the John M. Olin Center for Law, Economics, and Business at Harvard Law School. We received useful comments from Richard Craswell, Nuno Garoupa, Harold Hochman, Eric Rasmusen, Michael Stein, and an anonymous referee, and able research assistance from Paul Riskind.

Send correspondence to: A. Mitchell Polinsky, Stanford Law School, 559 Nathan Abbott Way, Stanford, CA 94305; Fax: (650) 723-3557; E-mail: polinsky@leland.stanford.edu or Steven Shavell, Harvard Law School, Cambridge, MA 02138; Fax: (617) 496-2256; E-mail: shavell@law.harvard.edu.

sense in which sanctions are considered to be fair or unfair.¹ It is apparent, however, that individuals do in fact have opinions about the fairness of sanctions. Notably, they tend to believe that a sanction should not be out of proportion to the gravity of the act committed. For example, if an individual double parks, a fine of thousands of dollars or a jail term is likely to be considered unfair in view of the modest harm caused by the act.²

In this article we incorporate notions about the fairness of sanctions into the standard model of public enforcement. To take beliefs concerning the fairness of sanctions into account, we assume that the utility of individuals depends on the sanctions that are imposed on those who violate laws—specifically, that individuals obtain what we refer to as *fairness-related utility or disutility* from the imposition of sanctions on others. Individuals might feel better off as a result of punishment of violators, perhaps due to retributive satisfaction, or individuals might feel worse off, because of empathy for violators. The use of fairness-related utility or disutility is the natural (indeed, the only) way to include conceptions of the fairness of sanctions in the model of enforcement, presuming that the fairness of sanctions must derive from what individuals want, as opposed to philosophical principles that might or might not be connected to individuals' preferences.³ We refer to the sanction that maximizes fairness-related utility (or minimizes such disutility) as the *fairness ideal*.

To compare our results to those in the usual analysis of enforcement, let us call the sanction that maximizes social welfare when fairness is ignored—namely, the harm caused by an act divided by the probability of catching the violator—the *conventional deterrence ideal*. When fairness-related utility exists, however, the desired level of deterrence changes because the full harm caused by a harmful act includes the expected fairness-related utility or disutility associated with the imposition of sanctions on the violator. For example, if individuals obtain retributive satisfaction from the punishment of violators, the full harm is less than the direct

1. For example, the issue of fairness is not discussed in two recent surveys of the economic literature on public enforcement—Mookherjee (1997) and Polinsky and Shavell (2000)—and is only briefly mentioned in a third—Garoupa (1997, p. 284).

2. That sanctions should reflect the gravity of the acts committed is the predominant view about the fairness of sanctions. See, for example, Benn (1967) and Greenawalt (1983).

3. We discuss this point further in section 5 below.

harm to the victim, and the desired level of deterrence therefore is less than it would be if this retributive satisfaction were ignored. Accordingly, we refer to the sanction that maximizes social welfare when fairness-related utility is recognized in defining the full harm, but when fairness is otherwise ignored, as the *full deterrence ideal*.⁴

When the fairness of sanctions is taken into account in the simplest model of law enforcement, in which the probability of catching violators is treated as fixed, the optimal sanction no longer equals the conventional deterrence ideal. Rather, because the optimal sanction reflects both fairness and deterrence considerations, it is between the fairness ideal and the full deterrence ideal, and it may be greater than or less than the conventional deterrence ideal.

When both the probability and magnitude of sanctions may be varied, the usual solution to the enforcement problem is also altered because of fairness considerations. As is well known, if individuals are risk neutral, the usual solution consists of the highest possible sanction and a relatively low probability, for this combination allows a suitable level of deterrence to be maintained while conserving enforcement resources.⁵ Thus, for example, even if a law violator caused just \$10 in harm, say from double parking, the usual solution in the risk-neutral case would call for a sanction of \$10,000 if this were the highest sanction that could be imposed; the probability of apprehension could then be very low—1/10%—and still result in an expected sanction equal to the harm of \$10.⁶

When the issue of fairness is added to the analysis, however, the usual solution is generally not optimal because a very high sanction will be seen as unfair, or more precisely, will result in the lowering of individuals' fairness-related utility. With respect to double parking, even a sanction of \$100 might be considered unfair, let alone a sanction of \$10,000.

4. The full deterrence ideal does not take into account the effect of sanctions on social welfare that is unrelated to deterrence—that is, the effect of sanctions on fairness-related utility when the number of individuals who commit the harmful act is held constant.

5. Although optimal sanctions in the usual model are not necessarily extreme if the assumption of risk neutrality is relaxed, there is still an incentive to set sanctions high in order to reduce enforcement costs.

6. In fact, the optimal probability is somewhat lower than 1/10%, but we use 1/10% for simplicity here. (The reason the optimal probability is such that the expected sanction is less than the harm is explained in note 14 below and the accompanying text.)

A consequence of the fairness-related motive to keep the sanction at a relatively low level is that the socially desirable probability of sanctions changes. One possibility is that the optimal probability is higher, perhaps much higher: to achieve a desired level of deterrence with a lower sanction, the probability has to rise. If the sanction for double parking cannot exceed \$100 because of fairness considerations, then, to create an expected sanction of \$10, the probability must be at least 10%, vastly exceeding the 1/10% optimal probability in the usual solution with a sanction of \$10,000. Another possibility, though, is that the optimal probability is lower than in the usual case: the additional deterrence created by raising the probability is relatively low because the sanction is relatively low, and the lower the deterrent benefit from raising the probability, the lower the social incentive to devote resources to enforcement.

In section 2, we present a model of enforcement that includes individuals' fairness-related utility. In section 3 we determine the optimal sanction when the probability of apprehension is fixed. In section 4, we solve for the optimal sanction and probability when both may be varied. In section 5 we comment on several topics as they relate to our analysis: different conceptions of fairness; risk aversion; imprisonment as the sanction; liability insurance; firms as defendants; the motivation of onlookers to cooperate with enforcers; and private enforcement through civil suit. In section 6 we conclude with some brief remarks.⁷

7. Several articles that are related to ours should be mentioned. Wittman (1974) analyzes optimal sanctions solely from the perspective of achieving retributive fairness, ignoring deterrence; he does not consider the optimal probability of catching injurers. Ehrlich (1982) discusses optimal enforcement under alternative concepts of justice; in the part of his analysis most relevant to our article, in which he considers retribution, he assumes that there is a social cost from failing to sanction injurers (those who escape liability), but he does not allow for the possibility that the sanctions that are imposed on injurers (on those who are caught) might be unfair. Miceli (1991) derives the optimal sanction and the optimal accuracy of the prosecutorial process (as opposed to the optimal probability of detection) when deterrence and the fairness of sanctions are taken into account. Waldfogel (1993) studies empirically whether actual sanctions are better explained by considerations of deterrence or fairness (in the sense that sanctions are proportional to the harm caused), and suggests that deterrence concerns better explain sanctioning practice. Diamond (1997) investigates the optimal sanction with respect to both deterrence and fairness when the probability of detection is fixed; he does not consider the optimal probability of detection because his analysis is in the context of private lawsuits for damages. Lando (1997) examines how fairness concerns affect the optimal due care standard under the negligence rule in tort law; he, too, does

2. The Model

We assume that there is a population of risk-neutral potential violators who obtain gains from committing a harmful act. The gains vary among these individuals. If an individual commits the act, he will be caught with some probability and made to pay a monetary sanction that cannot exceed his wealth.⁸ The probability of apprehension depends on the state's investment in enforcement effort. We use the following notation:

- g = gain of an injurer;
- $f(g)$ = density of g on $[0, \infty]$;
- h = harm;
- w = wealth of individuals;
- s = sanction, $s \leq w$;
- e = enforcement expenditures; and
- $p(e)$ = probability of catching violators.

We also assume that individuals derive utility from the imposition of sanctions on violators because of individuals' concern about fairness; let

$v(s)$ = individuals' fairness-related utility if s is imposed.

In other words, $v(s)$ represents the aggregation over all individuals of their individual utilities from the sanction being imposed.⁹ The term $v(s)$ will be more important the greater the number of individuals who care about punishment, the more each individual wants there to be punishment, and so forth. Also, $v(s)$ could be positive or negative for the reasons suggested in the introduction. The typical shape of $v(s)$ might be an inverted-U, but $v(s)$ could also steadily fall (suppose people empathize with the violator) or steadily rise (suppose the opposite). The interpretation of $v(0)$ is the utility (or disutility) that individuals derive when a violator is not sanctioned, including as a result of his not being caught.

not study the likelihood of sanctions. In Lando's analysis, the fairness-related utility or disutility experienced by the injurer and the victim are taken into account, but not, as here, that of other parties as well.

8. The assumption that sanctions are monetary is made mainly for convenience; we comment on the sanction of imprisonment in section 5 below.

9. There is no reason for our purposes to single out the victim's fairness-related utility (or disutility), even though the victim might care much more about the punishment of the injurer than do other individuals. Note, too, that some individuals might obtain utility due to spite or other factors not described as deriving from notions of fairness; $v(s)$ would include such utility as well.

We assume that $v(s)$ is concave in s , so that there is a unique sanction that maximizes $v(s)$, and refer to this sanction as:

$$s_f = \text{the fairness ideal.}$$

This is the sanction that would be socially best to impose if the only concern were the fairness of sanctions, that is, if sanctions did not also have a deterrence function in reducing the number of violators.¹⁰

Social welfare is assumed to equal the sum of utilities. Equivalently, social welfare equals the gains that violators derive from committing the harmful act, minus the harm they cause, plus individuals' fairness-related utility, and minus enforcement costs:¹¹

$$W = \int_{ps}^{\infty} [(g - h) + pv(s) + (1 - p)v(0)]f(g)dg - e. \quad (1)$$

Note that the lower limit of the integral is ps because individuals commit the harmful act when $g \geq ps$. The term in brackets reflects that if a person commits the harmful act, he obtains a gain of g and causes harm h , and he also is responsible for expected fairness-related utility of $pv(s) + (1 - p)v(0)$. The social problem is to choose e (and thus p) and s to maximize (1). Optimal values of these variables are indicated by an asterisk.

We consider the following example below: g is distributed uniformly on $[\$0, \$2,000]$, in which case $f(g) = .0005$ on that interval; $h = \$1,000$; $w = \$10,000$; $p(e) = .001e$ for e in the interval $[\$0, \$1,000]$; and $v(s) = -.0001(s - s_f)^2$, where $s_f = \$1,500$ (thus, $v(s)$ is an inverted-U with a maximum at a sanction of $\$1,500$).

3. The Optimal Sanction when the Probability Is Fixed

We first derive the optimal sanction s^* when enforcement effort and the probability of detection are fixed. This sanction is assumed to be unique and to be identified by the first-order condition from differentiation of (1). Differentiating (1) with respect to s to yields

$$\begin{aligned} W_s = & -p[(ps - h) + pv(s) \\ & + (1 - p)v(0)]f(ps) + pv'(s)[1 - F(ps)], \end{aligned} \quad (2)$$

10. We defer until the next section the formal introduction of the conventional deterrence ideal and the full deterrence ideal.

11. Social welfare can be expressed in this way because, if individuals are risk neutral, the payments of monetary sanctions by violators who are caught are simply transfers of money, as are taxes that are needed to finance the cost of enforcement (net of the revenue collected from sanctions).

where $F(\cdot)$ is the cumulative distribution of $f(\cdot)$. The term in the first set of brackets is the net social gain that is foregone as a result of deterring the marginal person: the gain of the marginal person is ps , the harm done is h , and expected fairness-related utility is $pv(s) + (1 - p)v(0)$. The second term is an inframarginal effect: it is the influence of a higher sanction on the fairness-related utility associated with persons who are not deterred. The first-order condition corresponding to (2) is

$$[(ps - h) + pv(s) + (1 - p)v(0)]f(ps) = v'(s)[1 - F(ps)]. \quad (3)$$

Note that if individuals do not care about fairness, so that $v(s) \equiv 0$, then (3) reduces to $ps - h = 0$, or $s = h/p$. Because the sanction determined by this familiar formula is based solely on promoting deterrence, we refer to it as:

$$h/p = \text{the conventional deterrence ideal.}$$

As observed in the introduction, when fairness-related utility is taken into account, the desired level of deterrence changes because the full harm caused by a harmful act includes the expected fairness-related utility (or disutility) associated with it. Because fairness-related utility itself depends on the sanction, the desired level of deterrence depends on the sanction. Accordingly, let

$$s_d(s) = \text{the full deterrence ideal given sanction } s,$$

where

$$s_d(s) = \{h - [pv(s) + (1 - p)v(0)]\}/p. \quad (4)$$

Thus, if individuals obtain expected utility from imposing sanctions (the term in brackets is positive), the full harm is lower than the direct harm h , in which case the full deterrence ideal is less than the conventional deterrence ideal. Conversely, if individuals suffer expected disutility from imposing sanctions, the full harm is higher than h , and the full deterrence ideal exceeds the conventional deterrence ideal.

We now show that the optimal sanction is between the fairness ideal and the full deterrence ideal, where the latter is evaluated at the optimal sanction. In other words, the optimal sanction can be interpreted as a compromise between the sanction called for on deterrence grounds (properly interpreted to reflect fairness-related utility) and the sanction called for on fairness grounds.

Suppose first that $s^* < s_f$. Then the right-hand side of (3) is positive at s^* because $v'(s^*) > 0$ if $s^* < s_f$. Hence, the left-hand side of (3) must be

positive at s^* , implying that $(ps^* - h) + pv(s^*) + (1 - p)v(0)$ is positive, or $ps^* > h - [pv(s^*) + (1 - p)v(0)]$. Dividing this last expression by p gives

$$s^* > \{h - [pv(s^*) + (1 - p)v(0)]\}/p = s_d(s^*). \quad (5)$$

Hence, $s_d(s^*) < s^* < s_f$. By similar reasoning, it can be shown that if $s^* > s_f$, then $s_f < s^* < s_d(s^*)$. Thus, s^* is always bounded by the corresponding full deterrence ideal and the fairness ideal.

The preceding discussion also can be used to show that the optimal sanction can be less than or greater than the conventional deterrence ideal, h/p . In the case in which s^* is less than s_f , suppose that expected fairness-related utility, given s^* , is negative—individuals suffer disutility from the imposition of sanctions:

$$pv(s^*) + (1 - p)v(0) < 0. \quad (6)$$

Because $(ps^* - h) + pv(s^*) + (1 - p)v(0)$ is positive in this case, (6) implies that $ps^* - h$ must be positive or, equivalently, that $s^* > h/p$. By analogous reasoning, if s^* exceeds s_f and expected fairness-related utility is positive, then $s^* < h/p$.

We may summarize our discussion by observing that there are two basic factors that determine the optimal sanction s^* . On one hand, there is optimal deterrence. This consideration suggests that the expected sanction should equal the full social cost of the harmful act: not just h , but rather h minus individuals' expected fairness-related utility. That implies that the expected sanction should tend toward $h - [pv(s) + (1 - p)v(0)]$, which could exceed h or be less than h . On the other hand, there is optimal sanctioning of those who are not deterred, and this consideration suggests that the sanction should tend toward the fairness ideal s_f . The optimal sanction s^* is the resultant of these two factors.¹²

In the example, in which the harm is \$1,000 and the fairness ideal is \$1,500, suppose the probability of enforcement is fixed at .15. Then the conventional deterrence ideal is \$6,667 ($= \$1,000/.15$). It can be calculated that the optimal sanction, however, is \$3,946.¹³ The full deterrence

12. An important difference between Diamond's (1997) analysis and ours is that we show that the optimal sanction is between the full deterrence ideal and the fairness ideal; another difference, as previously noted, is that we analyze the optimal probability of imposing sanctions (see the next section).

13. The optimal sanction was determined by calculating social welfare for values of the sanction between \$1 and \$10,000 in increments of a dollar. The full deterrence ideal was determined from (4) with $s = \$3,946$.

ideal, which includes the expected fairness-related disutility associated with imposing this sanction, is \$8,540. Thus, as must always be the case, the optimal sanction is between the fairness ideal and the full deterrence ideal, and, in this example, the optimal sanction is significantly below the conventional deterrence ideal.

4. The Optimal Probability and Sanction

To determine the optimal probability p^* and the optimal sanction s^* when both can be varied, we maximize W given by (1) over enforcement costs e and the sanction s , or equivalently, we maximize W just over e , substituting $s^*(p(e))$ for s , where $s^*(p)$ is the optimal sanction given p , as determined in section 3. Because $W_s = 0$ at $s^*(p(e))$, $dW/de = W_e$, and

$$W_e = -sp'(e)[(ps - h) + pv(s) + (1 - p)v(0)]f(ps) \\ + p'(e)[v(s) - v(0)][1 - F(ps)] - 1. \quad (7)$$

The term in the first set of brackets is the net social gain that is forgone from deterring the marginal person. The second term is an inframarginal effect associated with sanctioning a greater fraction of individuals who are not deterred; with respect to each such individual, fairness-related utility changes by $v(s) - v(0)$.

It is instructive first to compare s^* to its value in the usual case in which there is no fairness-related utility. In the usual case, the optimal sanction is maximal, equal to wealth. The argument is that if s were less than w , then s could be raised to w and p lowered to p^o such that $p^ow = ps$; thus, the number of violators would be the same but social welfare W would rise because e is lower. In the present case, however, in which there is fairness-related utility, the foregoing argument does not apply: although $p^ow = ps$ implies that the number of violators is the same and $p^o < p$ means that e is lower, the increase in the sanction from s to w may lower social welfare by reducing individuals' fairness-related utility. Indeed, for most harmful acts the fairness ideal presumably is far less than the maximal sanction, in which case raising the sanction to its maximum could cause a substantial reduction in social welfare due to fairness considerations. In the present case, therefore, s^* may well be less than w .

Next consider the optimal probability in the usual case in which fairness-based utility is absent. If $v(s) \equiv 0$, then the first-order condition

corresponding to (7) becomes

$$-sp'(e)(ps - h)f(ps) = 1, \quad (8)$$

which implies the well-known result that $ps < h$, including at $s^* = w$. In other words, in the absence of fairness considerations, the optimal probability and sanction are such that the expected sanction is less than the harm, meaning that there is underdeterrence—some individuals are induced to commit the harmful act even though their gains are less than the harm.¹⁴

When individuals' fairness-related utility is taken into account, the optimal probability p^* could be higher or lower than in the usual case for the reasons suggested in the introduction. Namely, if $s^* < w$, as would often be true, then to achieve any level of deterrence now requires a higher probability. This leads to the possibility that p^* exceeds the optimal probability in the usual case. However, if $s^* < w$, the effect on deterrence of an increase in the probability is smaller: the derivative of the expected sanction now is $p'(e)s^* < p'(e)w$. This suggests that p^* may be less than the optimal probability in the usual case.

A further reason the optimal probability may be higher or lower than in the usual case relates to the desired level of deterrence when individuals care about fairness. If their expected fairness-related utility is positive, then the full social cost from a harmful act is less than the direct harm h , so it may be desirable to deter harmful acts less than in the usual case by using a lower probability of detection. Conversely, if expected fairness-related utility is negative, the optimal probability may be higher than in the usual case.

Another difference when fairness matters is that it may no longer be desirable to set the sanction and the probability so as to cause some underdeterrence. For example, if individuals' expected fairness-related utility is negative, it may be desirable to set the probability and sanction such that the expected sanction exceeds the harm, so as to deter violations of the law in order to avoid both the harm caused and the fairness-related

14. To understand this result, suppose that p were such that $ps = h$. Then there would be no first-order loss of social welfare from lowering p because the individuals who would be induced to commit the harmful act would obtain gains equal to harm. But enforcement costs would be saved, making it desirable to lower the probability. See Polinsky and Shavell (2000, pp. 53–54).

disutility associated with imposing sanctions. Conversely, if individuals' fairness-related utility is positive, greater underdeterrence than in the usual case may be optimal.

The comparison of p^* and s^* when there is fairness-related utility with their values in the usual case can be illustrated in the example. In the usual case, in which fairness is ignored, the optimal sanction is \$10,000—the maximal sanction—and the optimal probability is .08. When fairness is taken into account, however, the optimal sanction declines to \$3,827 and the optimal probability rises to .17. Thus, consideration of fairness more than halves the optimal sanction and more than doubles the optimal probability. Note, too, that the expected sanction in the usual case is \$800 ($= .08 \times \$10,000$), whereas it is \$651 ($= .17 \times \$3,827$) when fairness is taken into account. The lower level of deterrence apparently reflects the fact that the effectiveness of the probability in achieving deterrence is greatly reduced due to the significantly lower level of the optimal sanction.

5. Extensions and Discussion

In this section, we discuss a number of extensions of the model and comment on some issues bearing on the analysis.

Individual preferences regarding fairness. The key assumption underlying our analysis is that individuals' utilities are affected by the punishment imposed on others and that social welfare, and thus optimal enforcement policy, should reflect this source of utility—just as any other source of utility. The importance of taking fairness concerns into account in designing enforcement policy therefore depends on the empirical significance of the taste that individuals have for fairness of punishment. In contrast, individual preferences are not relevant to the weight accorded to notions of fairness by most philosophers of punishment. Rather, such philosophers tend to attach importance to notions of fairness *per se*, and thus would want policy guided by these notions regardless of whether individuals have a taste for fairness.¹⁵ Moreover, for reasons elaborated upon elsewhere, this view of fairness is problematic because, if followed, it can lead to outcomes that make every individual in society worse off (Kaplow and Shavell, 1999).

15. As a formal matter, the $v(s)$ term in our analysis could be interpreted as an external measure of the fairness of sanctions, independent of individuals' utilities.

Different conceptions of fairness. Although we assumed that individuals' fairness-related utility v is a function only of the sanction s , in general v may be a function of other variables too, including the following: (1) the harm h , allowing the fairness ideal to depend on the magnitude of the harm (or on whether harm occurred);¹⁶ (2) an index of fault, allowing the fairness ideal to depend on the extent of the wrongdoer's fault; and (3) the probability of detection p , allowing the fairness ideal to depend on the degree to which an injurer who has been caught has been singled out.¹⁷ Also, fairness concerns may relate to the imposition of sanctions on innocent parties. It would be straightforward to consider these different conceptions of fairness, and although the conclusions would change in various ways, the two main points we made would continue to hold: that s^* generally is less than the maximal sanction, and that p^* generally is different from the optimal probability in the usual case, for essentially the reasons we discussed.

Imprisonment as a sanction. We assumed that the sanction was monetary and costless to apply. If instead the sanction were imprisonment, and thus socially costly to impose, there would be no qualitative differences in our conclusions. Again, the optimal sanction in the case without fairness-related utility would be maximal if individuals are risk-neutral,¹⁸ and again consideration of fairness often would result in s^* being less than maximal and in p^* being different from that in the usual case. The importance of fairness may be greater when imprisonment is the sanction, because individuals' feelings about the acts calling for imprisonment, and about the sanction, are likely to be stronger than when the sanction is monetary.

16. Formally, v may be written as $v(s, h)$; because we held h fixed, our analysis is consistent with this formulation.

17. Note, however, that it does not make sense for v to be a function of p alone, for the following reason. If s were very high in relation to the harm and considered to be unfair (suppose the sanction were capital punishment and the harm were littering), then individuals would feel worse off the higher is p , because an unfair sanction would then be more likely to be imposed. In contrast, if s were thought to be appropriate in relation to the harm, then individuals would feel better off the higher is p . Thus, when v depends on p , v must also involve s (and h if h varies).

18. The optimal sanction also would be maximal if individuals are risk averse with respect to the imprisonment term (but not if they are risk preferring). See Polinsky and Shavell (2000, pp. 54–55).

Attitudes toward risk. We assumed that individuals are risk neutral with respect to monetary sanctions. If they are risk averse, our conclusions would not be affected in any essential way. However, in that case the optimal sanction in the absence of fairness concerns is typically less than maximal (Polinsky and Shavell, 2000, p. 54), so that consideration of fairness will tend not to lower s^* as much, nor to affect p^* as much, as it did in the case in which individuals are assumed to be risk neutral.

Liability insurance. The possibility that individuals have liability insurance bears on our analysis, for the effective sanction suffered by a person with such insurance is less than the sanction imposed: the effective sanction is only the amount that is not covered by the insurance policy (plus any increase in future premiums). One presumes that concerns about the fairness of sanctions relate more to the effective sanction than to the nominal sanction. Suppose therefore that fairness-related utility is not $v(s)$, but rather $v(s - y(s))$, where $y(s)$ is the amount covered by the liability insurer when the sanction is s . Consequently, the fairness ideal with respect to the nominal sanction would be higher than in the absence of insurance, and generally so would the optimal sanction s^* .

Firms as violators. If sanctions are imposed on firms rather than on individuals, then fairness-related utility v may have to be reconsidered, presuming that what matters in terms of fairness is that culpable *individuals* bear sanctions, as opposed to the artificial legal entity of a firm. Then, one would want to identify the sanction actually suffered by a culpable person within a firm if the firm bears a sanction s . For example, if a firm demotes a person who caused the firm to incur the sanction s , then the person's loss from the demotion would be the relevant sanction, not s . Thus, $v(s)$ would become $v(z(s))$, where $z(s)$ is the sanction suffered by the person within the firm responsible for causing the firm to be sanctioned.

Private role in public enforcement. Although we assumed that the probability of sanctions depended only on public effort e , in fact private parties often play an important role in public enforcement by supplying information to enforcement authorities, testifying in court, and so forth. Notably, the ability of police to catch criminals often depends crucially on witness cooperation. Why should private parties rationally aid in public enforcement, given the costs to them of doing so (and the possibility of their suffering reprisal from violators)? An important part of the answer may be that they obtain utility from seeing violators punished. Plausibly, private

parties are likely to be more helpful if the sanction imposed on the violator is fair. Thus, the fairer are sanctions, the more private parties will aid in enforcement and the greater the probability of enforcement.¹⁹ If the choice of the sanction affects the probability in this way, the optimal sanction s^* will tend to be closer to the fairness ideal than in our analysis.²⁰

Private enforcement. Although our analysis has been concerned with public enforcement, we could also analyze private law enforcement through civil suits along similar lines. One observation in this regard is that because the probability of a private suit is often high, due to the gains from suit, the sanction that is ideal in terms of deterrence might be less than—not more than—the sanction that is ideal in terms of fairness. For example, if a party acts in a way that is considered outrageous, or that does great harm, a sanction approximating the harm might accomplish proper deterrence, but the fair sanction may be much higher. Another observation is that, whereas in the context of public enforcement the probability of a sanction being imposed is a function of the government's decision about enforcement effort, in the private context it is a function mainly of the sanction that the private plaintiff collects. Hence, the choice of the sanction affects the expected sanction ps , and thus deterrence, through p as well as through s , thereby complicating the determination of the sanction that best balances deterrence and fairness considerations.

6. Conclusion

We have incorporated notions of the fairness of sanctions into the standard theory of enforcement by including individuals' fairness-related utility. A concern for fairness not only has a direct effect on the choice of sanctions—often leading to a lower sanction than otherwise—but also

19. Formally, letting d represent private enforcement effort, suppose that $p = p(e, d)$, where p is increasing in both arguments. Also assume that an individual chooses d so as to maximize his expected utility from the imposition of sanctions on violators, namely, that he maximizes $p(e, d)v(s) + (1 - p(e, d))v(0) - d$. Then our point in the text is that the higher is $v(s)$, the greater will be the choice of d , and thus the higher will be p .

20. Andreoni (1991) develops an analogous point: Jurors are more likely to convict the defendant if the sanction is not extreme, due to their concerns about mistakenly imposing sanctions on innocent defendants. Consequently, lowering the sanction will raise the probability of conviction.

influences the optimal probability of enforcement. Notably, if the fairness ideal is less than the conventional deterrence ideal, the optimal probability may be higher than would be called for on the basis of deterrence alone because the optimal sanction may be lower than would be desirable solely from the perspective of deterrence. These observations may help to explain what one often observes in public enforcement: sanctions that are quite low in relation to potential violators' capacity to pay, and accompanying probabilities of detection that require substantial enforcement resources to maintain.

References

- Andreoni, James. 1991. "Reasonable Doubt and the Optimal Magnitude of Fines: Should the Penalty Fit the Crime?" 22 *RAND Journal of Economics* 385–95.
- Benn, Stanley I. 1967. "Punishment," in Paul Edwards, ed., *The Encyclopedia of Philosophy*, Vol. 7. New York: Macmillan and Free Press.
- Diamond, Peter A. 1997. "Integrating Punishment and Efficiency Concerns in Punitive Damages for Reckless Disregard of Risks to Others," Working Paper No. 97-19, Department of Economics, Massachusetts Institute of Technology.
- Ehrlich, Isaac. 1982. "The Optimum Enforcement of Laws and the Concept of Justice: A Positive Analysis," 2 *International Review of Law and Economics* 3–27.
- Garoupa, Nuno. 1997. "The Theory of Optimal Law Enforcement," 11 *Journal of Economic Surveys* 267–95.
- Greenawalt, Kent. 1983. "Punishment," 74 *Journal of Criminal Law and Criminology* 343–62.
- Kaplow, Louis, and Steven Shavell. 1999. "The Conflict between Notions of Fairness and the Pareto Principle," 1 *American Law and Economics Review* 63–77.
- Lando, Henrik. 1997. "An Attempt to Incorporate Fairness into an Economic Model of Tort Law," 17 *International Review of Law and Economics* 575–87.
- Miceli, Thomas J. 1991. "Optimal Criminal Procedure: Fairness and Deterrence," 11 *International Review of Law and Economics* 3–10.
- Mookherjee, Dilip. 1997. "The Economics of Enforcement," in Amitava Bose, Mihir Rakshit, and Anup Sinha, eds., *Issues in Economic Theory and Policy, Essays in Honor of Tapas Majumdar*. New Delhi: Oxford University Press.
- Polinsky, A. Mitchell, and Steven Shavell. 2000. "The Economic Theory of Public Enforcement of Law," 38 *Journal of Economic Literature* 45–76.
- Waldfoel, Joel. 1993. "Criminal Sentences as Endogenous Taxes: Are They 'Just' or 'Efficient'?" 36 *Journal of Law and Economics* 139–51.
- Wittman, Donald. 1974. "Punishment as Retribution," 4 *Theory and Decision* 209–37.