

CRIMINAL LAW AND THE OPTIMAL USE OF NONMONETARY SANCTIONS AS A DETERRENT

*Steven Shavell**

INTRODUCTION

In this Article, I analyze the theoretically optimal use of nonmonetary sanctions as a deterrent, and I then examine their actual domain of use and the major principles and doctrines of the criminal law.¹

The plan of the Article is as follows: Part I describes the model on which the analysis is based. Part II considers the determination of the optimal area of nonmonetary sanctions in the model. This determination is based on a list of factors that may make reliance on monetary sanctions as a deterrent problematic, and that therefore may make resort to nonmonetary sanctions appealing despite the higher social costs presumed to be associated with their imposition. Part II then applies this theory about the model to help explain why the actual domain of nonmonetary sanctions—principally the area of crimes punished by imprisonment—and that of monetary sanctions—including much of the area of tort law—are what they are.

Next, Part III studies the specific nature of the optimal system of deterrence assuming that it is desirable to employ nonmonetary sanctions. The theme that is elaborated in the analysis of the model here is that because nonmonetary sanctions are socially costly to impose, it is best for society to threaten to impose the sanctions only where parties can probably successfully be deterred from acting undesirably. However, it will appear inevitable that sanctions will sometimes be imposed; and from this it is concluded that sanctions should not be too high, and the optimal magnitude of sanctions is deduced. The optimal probability of apprehension of parties is also discussed, and it is emphasized that if the probability is too low, it will not be possible to deter certain parties even with the threat of the highest conceivable sanctions.

Part III then investigates the criminal law in view of the theoretical analysis. For example, one section examines the hypothesis that “intent” may serve as a proxy for factors that raise the optimal sanction,

* Professor of Law and Economics, Harvard University. I wish to thank Albert Alschuler, Lucian Bebchuk, Michael Faure, John Kaplan, Louis Kaplow, Daniel Meltzer, Geoffrey Miller, A. Mitchell Polinsky, Richard Posner, and David Rosenberg for comments, and the National Science Foundation for financial support.

1. I consider only the goal of deterrence even though other goals of criminal law, notably incapacitation, rehabilitation, and retribution, would have to enter significantly into any normative or descriptive theory claiming to be reasonably complete. I restrict attention to deterrence because I believe its role, being complicated, is best understood in isolation from that of other goals. The reader will have to be careful to keep this in mind, since the focus on deterrence may at times seem unnatural.

and thus that the importance given to intent in criminal law may promote the goal of deterrence. In addition to the subject of intent, the subjects of attempt, causation, responsibility, justification and excuse, and strict liability are considered.²

I. THE MODEL

In order to analyze the use of nonmonetary sanctions as a deterrent, it is helpful to set up a simple and concededly stylized model.³ According to the model,⁴ parties may commit acts that can cause harm. An act will be identified with a set of probabilities⁵ of possible consequences; shooting at a person, for instance, will be identified with the probabilities describing where, if at all, a bullet will strike that person. Notice that under this definition of an act, what we would usually call a failure to act (for example, failure to prevent a person from shooting another) will be considered an act, for a failure to act will be associated with some distinctive set of probabilities of possible outcomes. Observe as well that the thoughts⁶ of a party when committing an act are quite different from the act itself, that is, from the probabilities of the consequences identified with the act.

The probability that an act will cause harm multiplied by the magnitude of harm will be referred to as the "expected" harm.⁷ The harm

2. I will describe below, see, e.g., *infra* note 53, the specific contribution of this Article to previous writing on the theory of deterrence and criminal law, including C. Beccaria, *An Essay on Crimes and Punishments* (1872); C. Montesquieu, *The Spirit of Laws* (T. Nugent trans. 1878); Bentham, *An Introduction to the Principles of Morals and Legislation*, in *The Utilitarians* 173 (1973); Becker, *Crime and Punishment: An Economic Approach*, 76 *J. Pol. Econ.* 169 (1968). The main additions, however, may be mentioned now: (a) the argument of Part II relating the domain of nonmonetary sanctions to the factors that reduce the utility of monetary sanctions; (b) the level of detail with which the theory of the optimal use of nonmonetary sanctions is developed in the first section of Part III; and (c) certain points in the subsequent analysis of criminal law (primarily about intent, attempt, and causation). An interesting article by Richard Posner, which appears in the present Review, follows similar lines. See Posner, *An Economic Theory of the Criminal Law*, 85 *Colum. L. Rev.* 1193 (1985). His article and this one were written independently of each other, and the reader will probably wish to examine both.

3. The reason that the study of a simple model can be helpful is, of course, that just because it abstracts from so much of the detail of reality, it may allow one to see more clearly than otherwise the importance of particular factors of interest.

4. I present a formal version of the model and sketch its analysis in an appendix, *infra*.

5. A party's assessment of these probabilities will be assumed to coincide with the courts', except in the discussion of mistake. See *infra* text accompanying notes 82, 91-96.

6. The importance of a party's thoughts will be considered in the discussion of intent. See *infra* subsection III.B.1; *infra* note 41.

7. Thus if there is a 20% probability of causing harm of magnitude 100, the expected harm is $20\% \times 100 = 20$. More generally, if an act can cause different magnitudes of harm, the expected harm equals the sum of the products of the probabilities of the different possible magnitudes of harm with these magnitudes. For instance, if there

itself will ordinarily be interpreted as physical injury or property loss, although it is recognized that the concept of a natural or objective category of harm is not free from difficulty.⁸ The theoretical analysis advanced here, however, is not affected by the definition of harm that one adopts. The analysis assumes only the existence of a category of socially undesirable consequences.

A party may derive private benefits from an act. These benefits may or may not be increased by the occurrence of harm. If a person fires a gun at someone whom he wishes to injure, the benefit that he derives will very much depend on whether harm actually results. A person who drives at high speed to reach a destination on time, however, or an individual who throws heavy shingles from a roof to avoid the inconvenience of carrying them down, will derive private benefits that are not increased by the occurrence of harm.

The benefits that society will be considered to derive from a party's act may be less than or equal to the party's private benefits. In fact, society may derive no benefits at all from a party's act. Allowing for a divergence between social and private benefits gives the analyst greater freedom to describe society's values. In particular, since the analyst can assume that the social benefits from an act are zero, he is able to study a society that finds some acts (rape, for example) objectionable no matter how high the private benefits may be. More generally, allowing social benefits to be less than private benefits permits the analyst to take into account society's apparent tendency to impute little value to acts for which the private benefits inhere in a party's enjoyment of the disutility suffered by the victim. Where the private benefits (for example, benefits from reaching a destination in time) are not obtained from the enjoyment of the victim's disutility, society seems more likely to credit them in the social calculus.⁹

The commission of an act will be said to be socially undesirable if the expected social benefits are exceeded by the expected harm. Thus, certainly where the expected social benefits are zero, a potentially

is a 20% probability of harm of 100 and a 10% probability of harm of 300, the expected harm is $20\% \times 100 + 10\% \times 300 = 50$. Note that the expected harm may be interpreted as the average amount of harm that would be done if an act were carried out repeatedly under similar circumstances. The definition of the term "expected" is conventional and will be used throughout the Article.

8. One can cite examples of societies in which acts that would be seen today as harmful were not regarded as such (the Aztec practice of human sacrifice) and in which acts that we would view today as relatively innocuous were regarded as very harmful (the killing of a cat in ancient Egypt). One can also point to contemporaneous examples in which the harmful qualities of acts are much disputed (homosexual relations between consenting adults).

9. While the possible divergence between social and private benefits affects which acts are considered socially undesirable, it does not affect the qualitative nature of conclusions about optimal deterrence, since that will depend only on the existence of some category of socially undesirable acts.

harmful act will be socially undesirable. Where the expected social benefits are positive, however, whether an act is undesirable will depend on a comparison with the expected harm. Thus, it might be desirable for a person to speed to reach the hospital in an emergency, since the expected social benefits might be higher than the expected harm, whereas it might not be desirable for a person to throw shingles from the roof, since the benefits might be lower than the expected harm.

Whether or not a party will actually commit an act—as distinct from whether or not it is socially desirable that he do so—depends on his perception of the possibility that he will suffer a sanction, either monetary or nonmonetary.¹⁰ A party will commit an act if, and only if, the expected sanction would be less than the expected private benefits.¹¹ If he decides not to commit an act, he will be said to be deterred.

The imposition of monetary sanctions will be assumed to involve lower social costs than the imposition of nonmonetary sanctions. The motivation for this assumption is that the disutility to parties who must make payments may be viewed as roughly balanced by the addition to utility of parties who receive the payments or, where fines are paid to the government, who find their tax burdens reduced.¹² By contrast, the disutility experienced by parties punished by nonmonetary sanctions is not balanced in any automatic way by additions to the utility of other parties.¹³ Furthermore, the imposition of nonmonetary sanctions often involves direct claims on goods and services, as with the building and operation of prisons.¹⁴

Finally, to impose sanctions, society must apprehend parties who commit harmful acts. This requires society to maintain an enforcement apparatus, with accompanying expenses, and these expenses will rise

10. The nonmonetary sanction will usually be interpreted as imprisonment. Other nonmonetary sanctions include loss of reputation, social stigma, and the death penalty.

11. That I make this assumption hardly implies that I adopt the easily-caricatured view that in reality individuals are all the time weighing the threat of sanctions against the benefits of contemplated acts. The assumption is meant only to reflect the belief that the threat of sanctions is sometimes an important consideration for most individuals (for example, those deciding whether to report their full income to the tax authorities) and is often an important consideration for some individuals (for example, professional thieves).

12. There still are, of course, social costs associated with the imposition of monetary sanctions. These reside mainly in the expense of reaching parties' assets and in the costs of "risk bearing," that is, the possible dislocative effects on parties who have to pay large amounts.

13. Some individuals, however, may derive pleasure from seeing guilty parties punished.

14. In assuming that the cost of imposing nonmonetary sanctions exceeds that of monetary sanctions, I do not wish to deny that the reverse might occasionally be true. For instance, the costs of reaching parties' assets might be high in certain situations and exceed the expense of carrying out the death penalty. I nevertheless make the assumption because of the need to keep the model simple and because of my presumption that imprisonment usually involves greater costs than does the imposition of monetary sanctions.

with the level of the probability of apprehension.¹⁵

Social welfare will be defined as the expected social benefits associated with the commission of acts, less the expected harm caused by the acts, the social costs of imposing sanctions, and enforcement expenses. The problem for society will be to devise a system of sanctions and to select a probability of apprehension so as to maximize social welfare.

II. THE DOMAIN OF NONMONETARY SANCTIONS

A. *The Theoretically Optimal Domain of Nonmonetary Sanctions*

Since nonmonetary sanctions are socially more costly to impose than monetary sanctions in the above model, it is easy to say where nonmonetary sanctions should be employed. Specifically, nonmonetary sanctions should be employed only where monetary sanctions cannot adequately deter undesirable acts. Social welfare will be greater if only the less costly monetary sanctions are used to deter undesirable acts.¹⁶

Upon reflection, moreover, one can see that this argument actually implies that nonmonetary sanctions should be employed only after a monetary sanction equal to a party's wealth has already been imposed. Otherwise, there would be an opportunity to lower social costs while continuing to deter the same number of parties by simultaneously reducing the nonmonetary sanction and increasing the monetary sanction.¹⁷

In any event, to determine whether monetary sanctions can adequately deter undesirable acts, it proves useful to examine five factors. The initial three factors are important because they bear on the possibility that the monetary sanction needed to deter will exceed a party's assets, and thus that he could not be deterred by threat of a monetary sanction. One of these factors plainly is the size of a party's assets.¹⁸ The smaller a party's assets, the smaller the chance that the prospect of

15. I will assume for simplicity that the probability of apprehension is perceived accurately by parties.

16. This general point has been at least suggested by most writers who have considered the issue of the choice of sanctions from a utilitarian perspective, see Becker, *supra* note 2, at 193; Bentham, *supra* note 2, at 183; 1 C. Montesquieu, *supra* note 2, at 91; but the point has not been developed as below.

17. Suppose, for instance, that a person with assets of \$100,000 is fined \$5000 and is sentenced to four years in prison. The disutility of this combined sanction could be held constant by increasing the fine from \$5000 to \$100,000 and reducing the prison sentence to some appropriate level, say one year. This would save society the cost of imprisoning the person for three years. Of course, if a fine of \$100,000 would involve greater disutility than the original sanction of \$5000 and the four year sentence, then a fine of less than \$100,000 and no prison sentence could be substituted for the original sentence.

18. Because in reality a person can earn money in the future, and the state can, in principle, take away these future earnings, "assets" might be interpreted to mean present assets plus future earnings.

their forfeiture will be enough to outweigh the private benefits a party expects to derive from committing an undesirable act and thus the smaller the ability to deter him by use of monetary sanctions.¹⁹ In the extreme case, for instance, it is impossible to deter a person with no assets by the threat of monetary sanctions.

A second factor is the probability that a party will escape being sanctioned. The greater this probability, the higher is the monetary sanction required to deter. If the probability that a party will escape sanctions is one-half, the magnitude of the monetary sanction would have to be approximately doubled; if the probability is one-third, it would have to be approximately tripled; and so forth.²⁰ Such increases in the monetary sanction needed to deter make it more likely that the sanction will exceed the party's assets.

A third factor is the level of the private benefits a party will derive from an undesirable act. The larger these benefits, the higher the monetary sanction necessary to deter, and again the greater the chance that this amount will exceed the party's assets.²¹

The remaining two factors are the probability that an act will cause harm and the magnitude of the harm. They are relevant because they determine the expected harm that will result if undesirable acts are not deterred.²²

If the combined importance of these five factors is sufficiently high, that is, if the likelihood of failure to deter undesirable acts, together with the expected harm in which the acts would result, is sufficiently high, then resort to nonmonetary sanctions may be desirable despite the greater social costs attending their use. Nonmonetary sanctions

19. This point should not be confused with the point that a monetary sanction of given magnitude that can be paid will probably serve as a greater deterrent to a person with low assets than to a person with high assets. For example, a sanction of \$100 will very likely be a greater deterrent to a person with \$1000 of assets than to a person with \$100,000. My point, by contrast, is that losing his entire wealth will be a greater deterrent to the person with \$100,000 than to the one with \$1000 simply because the former has more at stake than the latter.

20. The relationship would be exact—the sanction would have to be multiplied by precisely $1/p$, where p is the probability of escaping the sanction—under the simplifying assumption that parties are “risk neutral,” i.e., that the marginal utility of a dollar does not depend on the level of wealth. Under the more plausible assumption that parties are “risk averse” and value dollars more the lower their wealth, the sanction would not have to be multiplied by as high a number as $1/p$ for the expected disutility of the uncertain sanction to be maintained.

21. It should be noted that because I am assuming that the act is undesirable, I am implicitly assuming that increases in the level of private benefits do not imply that social benefits become large enough to make the act desirable.

It should also be observed that it could be impossible to deter a party from committing an undesirable act by use of a monetary sanction alone even if his wealth were unlimited. This is because the private benefits he would obtain from the act might be nonmonetary and exceed the disutility of losing any amount of money.

22. Of course, the two factors also help to determine whether an act is undesirable.

may be helpful of course because their threatened use might deter parties who could not be deterred by monetary sanctions alone.

B. The Actual Domain of Nonmonetary Sanctions

The simple theory just discussed helps to explain the basic choices that society has made concerning where to employ nonmonetary sanctions. In the core area of crimes—the crimes that we frequently punish by imprisonment²³—a brief assessment of the five factors indicates that the level of deterrence would often be inadequate were we to rely on monetary sanctions alone.

The first factor, the level of parties' assets, is relevant because criminals as a class seem to have relatively little wealth.²⁴ And this is not unexpected, since many personal characteristics that are correlated with criminality (for example, poor education and drug and alcohol abuse) are also correlated with inability to earn. Moreover, a primary motivation for some crimes, particularly theft, robbery, and burglary, is presumably that individuals have little money of their own.²⁵ The second factor, the likelihood of escaping apprehension, is clearly relevant, for criminal acts frequently go unpunished.²⁶ The third factor, the level of private benefits obtained by wrongdoers, often seems large, at least by comparison to its size outside the area of crime of present con-

23. For ease, I focus here on imprisonment as a nonmonetary sanction even though there are other forms of nonmonetary sanctions.

24. In one Department of Justice report, statistics are presented showing that criminals have little or no legally obtained income. See Bureau of Justice Statistics, U.S. Dep't of Just., Report to the Nation on Crime and Justice 38 (Oct. 1983) [hereinafter cited as Report]. For instance, according to one survey, almost half of all male inmates in prison had annual incomes under \$3000 prior to arrest, and the median income for both male and female inmates was at the poverty level (as defined by the United States government) prior to arrest. Also, 40% of all males in prison had been unemployed when arrested, and of those who were working, about one fifth were working part-time. This compares to an employment rate of 84% for the United States male population in the 18-54 age group, with only 3% working part-time.

25. But needless to say, there are many reasons why individuals who are not poor might steal; a well-paid accountant might embezzle from his employer merely because he has a good opportunity to do so and wants to live the high life. I am speaking here only about general tendencies.

26. According to Federal Bureau of Investigation, U.S. Dep't of Just., 1981 Uniform Crime Reports, Crime in the United States 152 (Aug. 1982), the percentage of reported murders that resulted in arrest was 72%; the figure for aggravated assault was 58%; for rape, 48%; for robbery, 24%; for larceny-theft, 19%; for burglary, 14%; and for motor vehicle theft, 14%. Because many crimes are not fully reported and because arrests do not always result in convictions, the above percentages must be taken as upper bounds for the probability of imposition of sanctions. For example, in 1981 only 26% of household larceny thefts and only 67% of motor vehicle thefts were reported, see Report, *supra* note 24, at 24, and only about 50% of arrests for all serious crimes resulted in convictions, see Report, *supra* note 24, at 45. Hence, the probability of imposition of a sanction for larceny-theft may have been not 19% but approximately $26\% \times 19\% \times 50\% = 2.47\%$, and the probability for motor vehicle theft not 14% but about $67\% \times 14\% \times 50\% = 4.69\%$.

cern. Furthermore, plausible calculations indicate that the magnitude of the private benefits can easily outweigh the highest possible expected monetary sanction.²⁷ The other two factors, the magnitude and probability of harm, seem to figure predominantly in the core area of crime. The harm caused by serious crimes against persons, such as murder and rape, is obviously great, and the harm resulting from other "lesser" crimes in the core area—*theft and counterfeiting*, for instance—appears, upon reflection, to be also quite large.²⁸ Likewise, it is evident that parties who set out to commit crimes that fall in the core area will cause harm with high probability. In fact, we ordinarily *define* crime in a way that implies that the probability of harm is high.²⁹ All this suggests the validity of the claim that as a general matter something more than monetary sanctions must be employed to achieve an adequate degree of deterrence in the core area of crimes.³⁰

For purposes of comparison, it is illuminating to consider the importance of the five factors in the area of unintentional torts, where of course only monetary sanctions, namely judgments paid to victims, are used. First, the assets of parties—especially corporate entities—are probably higher on average in the area of unintentional torts than in the core area of crimes. Second, the likelihood of escaping sanctions seems significantly smaller in the area of unintentional torts than in the core area of crime, for usually a party causing tortious harm cannot avoid being identified as the responsible party or does not try to avoid identification. Indeed, if he does try to avoid identification (as when a driver who strikes a pedestrian leaves the scene of the accident), his act may be converted into a crime. Third, the benefits parties derive from

27. Suppose, for example, that a potential thief with reachable assets of \$500 contemplates stealing an automobile and knows that he would be convicted with probability 5%. This would be financially attractive so long as he could obtain more than \$25 for the automobile, since the expected sanction he would face is $5\% \times \$500 = \25 .

28. Theft can be seen to be very harmful on general utilitarian grounds because it leads individuals to invest resources (for example, in burglar alarms), in order to protect property and to use property in constrained ways (for example, not leaving a new car in an unattended lot). Moreover, if theft is a serious enough problem, work effort and production might suffer because individuals would not be able to count on enjoying their purchases. Counterfeiting, at a minimum, induces individuals to check currency to make sure that it is genuine. And if widespread, counterfeiting would destroy confidence in money as a medium of exchange; transactions would be reduced to barter, and productivity would fall greatly.

29. Before labeling a party's act as criminal, we either insist directly that the act made harm highly likely (as when the party behaved with extreme recklessness) or else that the party wished a harmful outcome to occur—meaning that he probably behaved in a way that raised substantially its chance of occurring. For discussion of intent, see *infra* Subsection III.B.1.

30. A thorough examination of the claim would discuss how its basis and the relative importance of the different factors varies by the circumstances of the case and by the category of crime. Murder, for example, results in greater harm than the theft of radios from automobiles, but the likelihood of apprehension for murder is greater. See *supra* note 26.

committing unintentional torts, and thus the difficulty of deterring them, often seem to be of a lower order than in the core area of crime. This is because the benefits are comprised only of the costs avoided by not taking safety precautions.³¹ Last, although the magnitude of harm may certainly be high in the area of unintentional torts, the probability of potentially tortious acts resulting in harm appears systematically smaller than the probability of criminal acts resulting in harm.³² In sum, it seems that, in the area of unintentional torts, the use of monetary sanctions should produce a much better level of deterrence than in the core area of crime.

There are, of course, several areas other than that of unintentional torts where monetary sanctions but not imprisonment are employed: intentional torts either that cannot be punished under the criminal law, such as defamation, or that the state often decides not to prosecute criminally, such as assault and battery; and acts punished by fines only, including many violations of tax, business, and driving regulations. While the point will not be pursued here, it seems clear that the chief difference between these areas and the core area of crime is that the magnitude or the probability of harm is lower. Thus, it is not altogether surprising that monetary sanctions are used in these areas.

At the same time, I should emphasize that the consistency I see between theory and reality is very approximate. It is not difficult to adduce examples of disagreement between the actual use of monetary and nonmonetary sanctions and their theoretically optimal use as deterrents. For instance, my impression is that there are many occasions where a person with not insubstantial assets is sentenced to prison but pays no fine or only a modest one.³³ If so, then as explained previously, a savings in social resources could be achieved by reducing prison sentences and making greater use of fines. An example of plausibly insufficient use of imprisonment may also be mentioned. Where firms might cause harms much greater than their worth or where the harms would be difficult to trace, tort liability may not create an adequate level of deterrence. Hence, more frequent imposition of criminal

31. Contrast the small benefits a person obtains from not shoveling ice from the sidewalk, or the savings a firm enjoys *per unit* of its product for not including a safety feature, with the private benefits the typical murderer derives from his act, or the embezzler or the traitor from his.

32. Contrast the usual likelihood of harm from an intentional criminal act such as shooting at someone or entering a home to steal, with the likelihood of harm from failing to clear ice from the sidewalk, or from negligently failing to signal when changing lanes on the highway. Indeed, that the likelihood of harm is ordinarily small in the context of unintentional torts is reflected by our use of the term "accident" to refer to adverse outcomes.

33. For example, a fine of \$5000 and an 18-month prison sentence was recently imposed on R. Foster Winans, a former Wall Street Journal reporter, for his part in a scheme to trade securities based on inside information. One supposes that Winans might have been able to pay a higher fine. See N.Y. Times, Aug. 7, 1985, at D1, col 6.

liability on the officers of firms may be necessary to achieve an acceptable level of safety. These examples illustrate the possibilities for the courts and the legislature to achieve social gains through an altered use of monetary and nonmonetary sanctions.

III. THE USE OF NONMONETARY SANCTIONS

Having discussed why nonmonetary sanctions, rather than less costly monetary sanctions, may be required to deter undesirable behavior, I will analyze in this Part the theoretically optimal use of nonmonetary sanctions³⁴ and then examine the doctrines of criminal law.

A. *The Theoretically Optimal Use of Nonmonetary Sanctions*

To determine the optimal use of nonmonetary sanctions, it is useful to consider the hypothetical situation where the courts³⁵ can obtain perfect information about parties who are apprehended, before studying the case where the courts' information is imperfect.

1. *The Situation Where Courts Can Obtain Perfect Information.* — Assume that the courts can obtain truly complete information about parties and their acts, once the parties have been brought before them. The courts will then be able to determine two things about any act that a party has committed.³⁶ First, they will be able to say whether the act was or was not socially undesirable. They will know, for instance, whether the death of *X*'s business partner on a hunting trip was the result of *X*'s deliberate act of shooting his partner to take control of their business, or was just an unfortunate accident.

Second, the courts will be able to tell whether a party could possibly have been deterred from committing an undesirable act *given* the probability of apprehension. A party could not possibly be deterred from committing an act if his expected private benefits exceed the disutility of the highest conceivable *expected* sanction—the highest conceivable sanction (perhaps the death penalty) discounted by the probability of its imposition.³⁷ The courts will be able to determine whether a

34. For simplicity, I will consider the use of nonmonetary sanctions unaccompanied by monetary sanctions. Hereinafter, "sanctions" will mean nonmonetary sanctions.

35. I use "courts" to refer to the legal entities that determine sanctions.

36. Note that since the courts have perfect information, they will not make a mistake as to whether a particular party has truly committed a particular act.

37. Suppose the benefit a party would derive from committing an act is 200, that the disutility of the maximum sanction is 1000, and that the probability of imposition of sanctions is 10%. Then the party could not possibly be deterred, for the maximum expected sanction would be $10\% \times 1000$ or 100, which is less than 200. It should be clear to the reader that as a formal matter, such a possibility of inability to deter rests on the notion that the disutility of sanctions is bounded. That the disutility of sanctions must be bounded is, in fact, a conclusion deducible from the usual postulates of expected utility theory. See K. Arrow, *Essays in the Theory of Risk Bearing* 63–69 (1971); L. Savage, *The Foundations of Statistics* 81–82 (2d rev. ed. 1972).

If the reader is troubled by the idea that deterrence may be impossible, it is likely

party could possibly have been deterred because they will be able to ascertain the benefits the party expected to derive from his act, as well as the disutility of the maximum sanction and the probability of its imposition.³⁸

Using these observations about the courts' knowledge, it is straightforward to describe the optimal system of sanctions. Under the optimal system, no sanction will be imposed on a party who committed a socially desirable act, for the courts will recognize the act as desirable and thus have no reason to employ a sanction. Also, no sanction will be imposed on a party who committed an undesirable act if he could not possibly have been deterred given the probability of imposition of sanctions, for the court will know that he could not have been deterred. Therefore, it would serve no purpose and yet be socially costly to impose a sanction. The remaining situation is that in which a party committed an undesirable act and could have been deterred. But this situation will never arise—and sanctions will never in fact be imposed—because it will clearly be optimal to set sanctions high enough to deter an undesirable act wherever deterrence is possible.³⁹

In this last case, where undesirable acts can be deterred, there is no single optimal level of the sanction. The sanction could as well be extremely high as barely sufficient to deter since an extremely high sanction is no more costly than a low one if neither will be imposed.⁴⁰ Moreover, the magnitude of the expected harm does not affect the size of the sanction (presuming that the expected harm is high enough to make the act undesirable). The sanction is dependent on the level of expected private benefits, however, for the expected sanction must be sufficient to offset those benefits.⁴¹ Similarly, the sanction is dependent

that he is thinking that most individuals would do almost anything to avoid a *certain* death. While this is doubtlessly true, it is in no way illogical for the same individuals to act in ways that raise the *probability* of death by only a small amount. It is a familiar fact that individuals subject themselves to increased risks of death for even rather trivial benefits (when they jaywalk for instance). They would not do so were the disutility of death infinite.

38. While I have assumed that the party accurately perceives this probability, in a more realistic model he might not. In such a model, the courts would have to determine what was the party's own assessment of the probability of apprehension to know whether he could have been deterred.

39. Suppose, for example, that as in note 37 *supra*, the maximum possible sanction is 1000 and the probability of its imposition is 10%, but that a party would obtain private benefits of only 50 from committing an undesirable act. This party could and thus would be deterred by the threat of a sanction, and therefore would never suffer a sanction.

40. The party of the previous note, who would obtain private benefits of 50, could be deterred by any sanction higher than 50—for that would result in an expected sanction higher than 50—and the sanction may as well be the maximum of 1000, since the sanction will not actually be imposed.

41. If the party of note 39 *supra* would obtain private benefits of 60 rather than 50, the sanction would have to be at least 600 to deter him. And notice that such depen-

on the probability of apprehending parties.⁴²

To determine the optimal probability of apprehending parties, two opposing effects must be taken into account. As the probability is raised, enforcement expenses increase, but the maximum possible expected sanction, and thus the number of parties who can be deterred from committing undesirable acts, also increases.⁴³ The optimal probability will reflect an appropriate balancing of these effects, and in general will be higher the greater the harm that would be done by parties. Furthermore, efficiencies that decrease enforcement expenses will tend to raise the optimal probability.

2. *The Situation Where Courts Cannot Obtain Perfect Information.* — Realistically, the courts cannot obtain perfect information about parties and their acts. The courts therefore cannot employ sanctions as just described, and their use of sanctions will differ from that under perfect information in three significant ways. First, the possibility that the courts will not recognize certain acts as socially desirable and will impose sanctions will deter some socially desirable acts. For instance, a person may decide not to break into a cabin when he is in great need of shelter for fear that the courts will not understand the true character of his situation and will punish him. Second, some socially undesirable acts may not be deterred even though a sufficiently high sanction would provide an effective deterrent. Thus, a person might decide to commit a robbery knowing that the sanction, if he is caught, will be a year in prison, while he would not act if the sanction were two years. Third, sanctions may actually be imposed: some parties who commit socially desirable acts will be punished (those who do break into cabins when they need shelter and are erroneously punished), some parties who commit undesirable acts and who could have been deterred will be punished, and some parties who commit undesirable acts and who could not have been deterred will be punished (for example, those who kill others in moments of uncontrollable rage).⁴⁴

The magnitude of each of these problems varies with the level of sanctions. As the level of sanctions rises, more undesirable acts will be deterred, but the social cost of imposing sanctions in a given instance becomes greater, as does the problem of discouraging socially desira-

dence of the sanction on the private benefits a party would obtain means that the sanction depends on his thoughts (since they reflect the benefits he anticipates).

42. If the probability of apprehension of the party who would obtain private benefits of 50 were 20% rather than 10%, the size of the sanction necessary for deterrence would be 250 rather than 500, since $20\% \times 250 = 50$.

43. If p is the probability of apprehension and 1000 is the maximum sanction, then the maximum expected sanction equals $p \times 1000$. Thus, the parties who can be deterred are those who would obtain expected private benefits less than $p \times 1000$. The higher the value of p , the greater the number of parties who can be deterred.

44. A fourth reason why sanctions will sometimes be imposed is that some parties who committed no act at all will mistakenly be thought to have committed a punishable act. I will not be discussing this problem here, however.

ble acts. The optimal level of the sanction will be that which makes the best compromise between these competing effects.⁴⁵

The nature of this best compromise and thus the level of the optimal sanction depends on the characteristics of parties and their acts (as they are perceived by the courts). It is clear, first, that the higher the expected private benefits, the higher will be the optimal sanction—unless the benefits are so large that deterrence is probably not possible. This is because the higher the expected private benefits, the higher is the expected sanction needed to accomplish deterrence where deterrence is possible.⁴⁶

It is also true that the greater the expected harmfulness of an act, the higher will be the optimal sanction. This is for three reasons. First, the greater the expected harm, the more society values increased deterrence, and therefore the more willing society should be to bear the costs of imposing higher sanctions.⁴⁷ In addition, an increase in ex-

45. To illustrate the process of determining the optimal sanction, suppose again that the maximum possible sanction is 1000, that the likelihood of apprehension is 10%, that there are two parties *X* and *Y* who each might commit an undesirable act causing expected harm of 800, that *X* would obtain a private benefit of 50 from committing the act (so that he can be deterred only by a sanction of at least 500), that *Y* would obtain a private benefit of 200 from the act (so that he cannot be deterred even by the highest sanction), and that courts are unable to distinguish between *X* and *Y*—the respect in which the courts' information is imperfect—and thus must use the same sanction for each.

It follows that if a positive sanction is optimal, it will be a sanction of 500. In particular, a positive sanction less than 500 would deter neither *X* nor *Y*, and thus sometimes would be imposed but to no purpose. A sanction of 500 (or slightly more) would deter *X* but not *Y*. A sanction above 500 would only increase the sanction that would be imposed on *Y*, since he will not be deterred, and therefore would be inferior to a sanction of 500.

To determine whether a positive sanction of 500 rather than a sanction of zero is optimal, a comparison of levels of social welfare must be made. In computing social welfare, assume for simplicity that the private benefits that *X* and *Y* obtain from committing their acts have no social value and that the social costs of imposing sanctions equal the disutility experienced by the punished parties, and for simplicity ignore society's enforcement expenses. Thus, maximization of social welfare will correspond to a minimization of expected harm plus expected sanctions imposed. If the sanction is 500, these expected social costs are $800 + 10\% \times 500 = 850$ (since *Y* alone commits the undesirable act), and if the sanction is zero, the social costs are $800 + 800 = 1600$ (since both *X* and *Y* commit the act). Hence a sanction of 500 is optimal.

46. In the example of the previous note, if the expected benefit of *X* were 60 rather than 50, the optimal sanction would be 600 rather than 500.

47. To illustrate this, the example in note 45 *supra* must be modified and made somewhat more complicated. Suppose that there is a third party, *Z*, who would obtain a private benefit of 70 from committing the act—and thus who could be deterred, but only by a sanction of at least 700. Suppose also that there are 1000 *X*'s, one *Z*, and 100 *Y*'s. Then it is straightforward to verify along the lines of note 45 that the optimal sanction is still 500. At 500, the expected social costs due to the imposition of sanctions would be $101 \times 50 = 5050$ since *Z* and the 100 *Y*'s would commit the act. At 700, the expected social costs due to the sanctions would be $100 \times 70 = 7000$ since the 100 *Y*'s but not *Z* would commit the act. The reason a sanction of 500 is socially superior to one of 700 is

pected harm may be associated with an increase in expected private benefits, for, as previously noted, the private benefits often depend on the harm, as when a party robs or intentionally injures another.⁴⁸ And since the optimal sanction rises with the expected private benefits—because parties become harder to deter—an increase in expected harm will tend indirectly to increase the optimal sanction. Finally, raising the sanction with the expected harmfulness of acts gives parties who are *not* deterred incentives to do less harm. Specifically, it encourages them to choose initially to commit less harmful acts (to plan to beat up an enemy rather than knife him to death) and, during the commission of acts, to refrain from increasing the level of harm done (to avoid killing a person in the course of a robbery). Such incentives will be said to reflect “marginal deterrence.”

Although the optimal sanction rises with the expected harm associated with an undesirable act, that does not imply that the sanction necessarily rises with the *actual* harm done. Of course, this is because in calculating expected harm, one discounts possible levels of actual harm by their probabilities.⁴⁹ For example, act *A* may result in higher actual harm than act *B*, yet *A* may have created a much lower probability of harm and consequently be associated with a lower expected harm.⁵⁰ Therefore, the optimal sanction for *A* could be lower than the optimal sanction for *B*. Nevertheless, the actual harm done is relevant to the optimal sanction where, as will often be true, the courts’ knowledge of the expected harmfulness of an act is poor. Since they do not directly observe the probability of harm or the different levels of harm that might have resulted, courts must infer these from, among other evidence, the occurrence and magnitude of the actual harm. Thus, the sanction can reasonably rise with the actual harm. In addition, marginal deterrence might not be adequate unless the sanction rises with the level of actual harm.

The relationship between the probability of apprehending parties and the optimal sanction is, one would expect, that the lower the probability, the higher the optimal sanction. This would likely be required to maintain the expected sanction needed for proper deterrence.⁵¹

that it would not be worthwhile to incur additional social costs of 1950 just to deter the one *Z* from committing an act whose expected harm is 800. If, however, the expected harm associated with the undesirable act were above 1950, it would be optimal to deter *Z* by raising the sanction to 700.

48. See *supra* p. 1234.

49. See *supra* note 7.

50. Suppose act *A* causes harm of 1000 with probability of 10%, and act *B* causes harm of 500 with probability of 50%. Then the expected harm associated with *A* is 100 and is lower than that associated with *B*, namely, 250.

51. It is theoretically conceivable, however, for the optimal sanction to fall if the probability falls. This can happen if, when the probability falls, many parties become undeterrable and the remaining parties can be deterred easily. To illustrate, consider

Finally, with regard to the determination of the optimal probability of apprehending parties, the main point is the one from before, namely, that the probability should not be too low lest too many parties become impossible to deter. There is now, however, a further reason why the probability should not be too low: the sanction for acts would then generally have to be very high to accomplish adequate deterrence, but that would undermine marginal deterrence because the sanctions for acts of differing severity would be similar.⁵²

This completes the theoretical discussion of the optimal use of nonmonetary sanctions as a deterrent. In thinking about the subject, the reader should keep in mind that it cannot be equated with the simple problem of deterring a person with known characteristics from committing an undesirable act. The proper view of the subject takes into account the imperfect nature of the courts' information and the errors that result relative to the situation under perfect information.⁵³

another modification of the example in note 45 *supra*. There are three parties: not only *X* who would obtain a benefit of 50 and *Y* who would obtain a benefit of 200, but also *W* who would obtain a benefit of 2. If the probability of imposition of sanctions is 10%, then both *X* and *W* can be deterred, and the optimal sanction would be 500. But if the probability of sanctions were 1%, only *W* could be deterred (as the maximum expected sanction would be 10), and the optimal sanction would fall from 500 to 200.

52. For instance, if the probability of apprehension were very low, the sanction for both robbery and murder might have to be near life imprisonment to create adequate deterrence. Then only a small difference in the sanctions for the two crimes would exist, and robbers would have little reason not to murder their victims.

53. This view of the subject has not been developed by other writers on deterrence and criminal law, although, of course, their contribution is great, and it is worthwhile to describe it briefly. The most important writers on the theory of the use of sanctions as a deterrent are probably Charles Montesquieu, Cesare Beccaria, Jeremy Bentham, and Gary Becker. Montesquieu and Beccaria were among the first to adopt a utilitarian style of analysis. While they were often vague and unanalytical, they suggested that, because nonmonetary sanctions are costly to impose (they considered mostly the disutility to punished parties), sanctions should be used sparingly and only when likely to accomplish deterrence. See C. Beccaria, *supra* note 2, at 17-19, 21; 1 C. Montesquieu, *supra* note 2, at 84. They occasionally raised points of considerable sophistication about deterrence, as for instance with respect to marginal deterrence. See C. Beccaria, *supra* note 2, at 32; 1 C. Montesquieu, *supra* note 2, at 90. Bentham's work built on Montesquieu and Beccaria (whom he frequently cites) and generally stands out as incomparable. He was the first to state carefully and investigate the assumption that potential wrongdoers weigh possible benefits against the probability and magnitude of sanctions. Of particular interest is his analysis of the optimal magnitude of sanctions. See Bentham, *supra* note 2, at 169-77. Becker was the first to attempt a formal model of sanctions and deterrence, and was also the first to study the probability of apprehension as an instrument of social choice.

The present Article goes beyond these works and explicitly takes into account the imperfect nature of the courts' information. Only by doing so can one explain why, if a particular level of the sanction will deter, just that sanction, and not a higher one, should be used. Without this, one's theory of optimal sanctions is not really logical, or is at least incomplete. Also, this Article stresses the implications of the possible inability to deter parties given the probability of apprehending them. Only by considering these implications can one explain why extremely high sanctions should not often be used,

B. *The Principles of Criminal Law and the Use of Nonmonetary Sanctions*

The major principles and doctrines of criminal law will now be considered in view of the theory of optimal deterrence.

1. *Intent*. — A central feature of the criminal law is the emphasis it places on intent. To analyze intent, it is best to begin by making several definitions. Let us say that a party “desires” a result if it would either directly or indirectly raise his utility.⁵⁴ Let us also say that a party “intends” a result if he (a) desires the result and (b) acts in a way that he believes will raise the probability of the result.⁵⁵ This definition of intent seems to comport with its ordinary meaning.⁵⁶ According to the definition, we would say that *X* intended that *Y* die if he desired *Y*’s death and shot at *Y* and killed him. If, however, *X* shot at *Y* but instead struck *Z* whose death he did not desire, we would not say that he intended that *Z* die. Also, we would not say that *X* intended that *Y* die if *X* played golf with *Q*, and *Y* happened to be killed in an automobile accident (since the round of golf did not increase the probability of *Y*’s death).

In the criminal law, the role of intent, as I have defined the term, may be summarized by several statements.⁵⁷ First, intent to do harm is ordinarily a principal factor in determining liability and the severity of punishment. Second, the effect of intent on liability and punishment is generally the same whether an intended harmful result is directly or indirectly desired. Whether *X* shot *Y* because *Y* was his enemy or only because *Y* stood in the way of an inheritance will not ordinarily affect the punishment of *X* under the law. Third, whether a harmful result different from the desired result occurs does not usually influence a party’s legal treatment. Where *X* aimed at *Y* but shot *Z* instead, it is murder just as if he aimed at *Z*.

These features describing the role of intent are roughly consistent with the purposes of deterrence, for intent appears to be linked to the factors that, according to theory, call for, or increase the level of, sanc-

and why it is not desirable to allow the probability of apprehending parties to be very low in order to save enforcement expenses.

54. In the language of utility theory, a result is desired if it is an argument in the individual’s utility function that would raise his utility or, if the result is not such an argument in his utility function, if its occurrence would lead to an increase in his expected utility.

55. The significance of erroneous beliefs is discussed *infra* notes 80–82, 91–96 and accompanying text. For now, I assume the party’s beliefs about the probability are correct.

56. The traditional definition of intent in the criminal law is broader: a party “intends” a result even where he does not desire it if he acts in a way that makes it highly probable (rather than only more probable).

57. See W. LaFare & A. Scott, *Handbook on Criminal Law* § 28 (1972) (describing the role of intent). See generally H. Hart, *Intention and Punishment in Punishment and Responsibility* 113 (1968) (a philosophically-oriented discussion of intent and criminal law) G. Fletcher, *Rethinking Criminal Law* § 6.5 (1978) (same).

tions.⁵⁸ Intent is, first of all, positively related to the probability of harm, for by definition, where a party intends to do harm, his act raises the probability of harm. This factor is particularly significant when the courts' direct evidence about the probability of harm is limited, since they can often make inferences about the probability from knowledge of a party's intent. For instance, where the courts hear little testimony about *X*'s shooting of *Y* but know that *X* had the purpose of killing *Y*, they might infer that *X* carefully drew a bead on *Y*.⁵⁹ Second, intent may be correlated with the likely magnitude of harm because a party who desires a harmful result is prone to do greater harm than one who does not. *X* may more often hit *Y* in a vital spot than in an arm or a leg if *X* desires to harm *Y*. Intent is also closely related to the private benefits parties expect to derive from their acts. By definition, the utility of parties who intend harm is raised by the occurrence of harm, and as just indicated, both the probability and magnitude of such desired harm tend to be higher where there is intent. Thus, parties who intend to do harm should be more difficult to deter.⁶⁰ In addition, intent to do harm may be associated with the absence of social benefits, for as pointed out earlier, society often appears reluctant to value private benefits that are based on the enjoyment of harm.⁶¹ Finally, intent may be linked to the probability that a party will escape a sanction, since a party who intends to commit a harmful act is more likely to choose a particular place and time to avoid identification and arrest, or to take steps thereafter to do so.⁶²

58. Oliver Wendell Holmes, Jr., was apparently the first to try to establish a connection between intent and factors that ought to increase the sanction appropriate for deterrence. His discussion was largely limited, however, to the relationship between intent and the probability of doing harm. See O. Holmes, *The Criminal Law, in The Common Law* 34 (M. deWolfe ed. 1963).

59. It is worth developing this example in more detail. Suppose that *X* and *Y* were hunting together when *X* shot *Y*. *X* claims that he fired at a deer running between him and *Y* and unfortunately did not see *Y*. A witness who was standing at some distance confirms that there was a deer running between *X* and *Y*, but he is not able to say whether *X* noticed *Y* or aimed at him rather than at the deer. With only this very imperfect knowledge of *X*'s act, the courts could value highly information about *X*'s intent (for example, evidence that he planned to kill *Y* at a good opportunity) or lack thereof (for example, evidence that *X* had nothing to gain from *Y*'s death). Thus, knowledge of his intent may alter the courts' assessment of the probability of harm due to *X*'s act. If, however, the courts' direct knowledge of *X*'s act were complete (for example, suppose they had a close up movie of his behavior), they would not need to know anything about intent to assess the probability of harm. But, as will be seen, they might well find knowledge of intent valuable for other reasons.

60. If *X* intends to kill *Y*, it will be difficult to deter him, because he wants *Y* dead and because shooting at *Y* makes this result likely. By contrast, if *X* is a true friend of *Y*, to deter a negligent or reckless shot will not require a substantial sanction (if it requires any sanction at all).

61. See *supra* p. 1234.

62. In some cases, however, the factor of intent could increase the probability of sanctions because a person's motives might be discoverable and lead police to investigate him. The importance of this consideration depends on the type of crime and the

These arguments suggest why intent, though mainly a mental factor, ought to influence liability and punishment according to deterrence theory. Moreover, it should be noted that the arguments do not depend on whether the intended harm is directly or indirectly desired. In either case, the probability and magnitude of harm, the expected private benefits, and the likelihood of escaping sanctions are likely to be higher than for unintentional conduct.⁶³ Moreover, the arguments are largely unaffected by whether the actual result was the desired result. It therefore makes sense that such distinctions usually do not affect a party's punishment.

Consideration of situations in which a party is not liable despite his intent to do harm sheds further light on intent and the theory of deterrence.⁶⁴ A party may intend to do harm but escape liability because circumstances make his act socially desirable; for example, a party forced to choose between two harmful acts may invoke the defense of necessity if he chooses the less harmful.⁶⁵ In addition, a sanction is not imposed where deterrence may be impossible, as in cases of insanity⁶⁶ or duress.⁶⁷

Conversely, parties are sometimes punished despite their lack of intent to do harm. Where a party does not desire a harmful result but acts in such a way that serious harm becomes very likely, he may be punished. Imposition of sanctions here may be justified because the expected harm is high;⁶⁸ the fact that the party does not desire the harm does not make his behavior less dangerous. Similarly, where a party does not desire harm but commits a strict liability crime, his punishment may be justified in principle if the courts find it very difficult to differentiate between desirable and undesirable acts.⁶⁹

2. *Attempt.* — The criminal law punishes attempts to do harm.⁷⁰ To understand why punishment of attempts may also be desirable

particular case. It might be significant in some cases of murder, for instance, but would probably not be in most cases of theft and robbery.

63. The only argument that is affected concerns the discounting of the private benefits. If the intended result is directly desired, the private benefits seem more likely to be discounted than otherwise.

64. That there should be such situations is not surprising a priori, for intent was only said to be linked with factors leading to the optimality of liability.

65. See *infra* note 102 and accompanying text.

66. See *infra* notes 89–90 and accompanying text.

67. See *infra* notes 100–01 and accompanying text.

68. Notice here the consistency with deterrence theory. Where the result is not desired, the private benefits are lower and thus deterrence is easier, other things being equal. Therefore, to warrant liability, the expected harm associated with the act must be unusually high.

69. See *infra* notes 109–12 and accompanying text.

70. See W. LaFare & A. Scott, *supra* note 57, § 59. It is interesting to note that this is in contrast to the situation in tort law, where there is no liability unless harm is done. See W. Keeton, D. Dobbs, R. Keeton & D. Owen, *Prosser and Keeton on the Law of Torts* § 30, at 165 (5th ed. 1984) [hereinafter cited as *Prosser & Keeton*].

under the theory of deterrence, it is necessary to consider why punishment of parties only when they actually do harm might not create an effective deterrent. One reason is that even if the sanction were set at the maximum level, the probability of its imposition might not be sufficient to make the expected sanction high enough to deter.⁷¹ In addition, even where it would be possible to deter most parties by punishing them only when they do harm, the necessary sanctions would generally be high (because of a low probability of their imposition) and conflict with the goal of marginal deterrence, since marginal deterrence requires a certain degree of gradation in the schedule of sanctions.⁷²

The punishment of attempts may serve to solve these problems by raising the expected sanction without increasing the magnitude of sanctions, for the punishment of attempts in effect increases the probability of sanctions.⁷³ Moreover, the punishment of attempts is a socially inexpensive means of increasing the probability, since opportunities to punish attempts often arise as a byproduct of society's investment in apprehending parties who actually do harm.⁷⁴

The force of this argument for sanctioning attempts clearly increases with the probability that a party will be apprehended for an attempt. Where an act takes a long time to carry out (especially where it requires preparations) or where it has a substantial chance of not succeeding (for example, shooting from a distance), the probability of being caught for an attempt will be higher than otherwise. Therefore, the deterrent value of punishing attempts will also be higher.

The possible desirability of punishment of attempts according to deterrence theory does not imply that there is any advantage in punishing attempts in the way the criminal law does, namely, less severely

71. This reason does not apply in the usual model of torts, because it is assumed that parties do not escape suit and have assets sufficient to pay for harm done. In that model, making parties pay money damages only when they actually do harm creates adequate incentives to reduce risk. See Brown, *Toward an Economic Theory of Liability*, 2 J. Legal Stud. 323 (1973).

72. See *supra* note 52 and accompanying text.

73. An illustration may be helpful. Suppose a party who attempts a crime will succeed in doing harm half the time, that he will be apprehended with probability 30% if he does harm and 20% if he does not, that the disutility of the maximum sanction is 1000, and that the expected private benefit from his act is 200. If sanctions are not imposed for unsuccessful attempts, the probability of imposition of sanctions is just $\frac{1}{2} \times 30\% = 15\%$, so the maximum expected sanction is only 150. Thus, since the party's expected benefits are 200, he cannot be deterred. But if the party is punished whether or not his attempt is successful, the probability of imposition of sanctions is $\frac{1}{2} \times 30\% + \frac{1}{2} \times 20\% = 25\%$, so the maximum expected sanction is 250; hence the party can be deterred.

74. Given that the police stand ready to apprehend those who do harm (by giving chase, investigating reports, etc.), it may not be that much more expensive for them to apprehend those who commit unsuccessful attempts. At least it would probably be much more expensive for the police to increase the probability of apprehension comparably by making greater efforts to apprehend only those parties who do harm.

than acts that result in harm. In discussing this feature of the criminal law, it is useful to consider separately "interrupted attempts"—acts discovered before they could have succeeded—and "completed attempts"—unsuccessful acts not discovered until later.

With respect to interrupted attempts, the following argument is sometimes made.⁷⁵ If the sanction for an attempt is lower than that for doing harm, a party who begins an attempt might be induced to reevaluate and abandon it, since he then will be punished less. If, however, the sanction for the attempt is the same as for doing harm, he may as well continue. As stated, this argument fails to recognize the possibility of treating the abandoned attempt leniently, while imposing a full sanction on attempts interrupted by others. Suppose for instance that no sanctions are imposed for abandoned attempts and the sanction for an interrupted attempt is the same as the sanction for an act that causes harm. Then the party who sets out to commit a harmful act will certainly have reason to abandon it; not only will he escape sanctions, but he will otherwise face a sanction for a later interrupted attempt equal to the sanction for doing harm.⁷⁶

Nevertheless, punishing interrupted attempts less severely than acts that result in harm may be advantageous. Since interrupted attempts may later be abandoned or fail, there is less evidence of the dangerousness of interrupted attempts and less reason for sanctioning them than acts that actually do result in harm.⁷⁷ The significance of this argument plainly depends on the character of the attempt and the point at which it is interrupted. If an attempt is nearly complete and is likely to succeed, the argument does not carry much weight. That would be so, for instance, where a person had already dropped a lethal dose of poison into his intended victim's drink. An attempt interrupted

75. This argument is usually attributed to Beccaria. See C. Beccaria, *supra* note 2, at 138.

76. However, difficulties courts may experience in distinguishing between interrupted and abandoned attempts lead to qualifications in this line of reasoning. Parties may pretend to abandon their attempts if they suspect that they are going to be interrupted (for example, individuals who are about to rob a bank may notice police approaching and leave). To the extent that parties' acts are mistakenly seen by the courts as abandoned attempts, it would be desirable to impose a positive sanction for apparent abandoned attempts. The opposite mistake occurs when parties who truly abandon attempts are thought by the courts to have stopped only because they believed they were about to be interrupted (for example, suppose the individuals in the bank really did not realize the police were approaching but were thought by the courts to have thought so). To the extent that this mistake arises, it would be desirable to impose a lesser sanction for an apparently interrupted attempt. For a description of difficulties in distinguishing abandoned attempts from interrupted attempts, see W. LaFare & A. Scott, *supra* note 57, § 60; and for a general discussion of the subject, see G. Fletcher, *supra* note 57, § 3.3.8.

77. This the reader will recognize as a version of the general argument advanced earlier, see *supra* text following note 50, that the actual harm done might influence the sanction because of the courts' incomplete information about the dangerousness of an act.

further from completion might properly be sanctioned less severely, however. Indeed, an attempt might escape a sanction altogether if it is interrupted so early that there is great doubt whether and in what manner it would have been continued.⁷⁸ Thus, if the person was apprehended merely when leaving a drugstore with poison, it might be unclear whether he would have used the poison, and unclear too whether his behavior would satisfy the definition of attempt in criminal law.

Two arguments analogous to those just discussed are often advanced to justify imposition of a lower sanction for unsuccessful completed attempts than for successful attempts. First, it is asserted that if the sanction for an unsuccessful completed attempt is equal to the sanction for a successful attempt, a party whose initial attempt fails will have nothing to lose by trying again. This argument, however, overlooks the fact that the sanction for an initial unsuccessful attempt may equal the sanction for an initial successful attempt, and yet the party *will* have something to lose by trying a second time so long as the sanction for a second successful attempt is higher. For example, if the sanction for an initial attempt is five years whether or not it succeeds, but the sanction for causing harm on a second attempt is ten years, a party who at first fails will clearly have reason not to try a second time.

The other argument for punishing unsuccessful completed attempts less severely than those resulting in harm is that the failure of an attempt may constitute evidence that it was a less dangerous act. As with interrupted attempts, the strength of this evidentiary rationale depends on the nature of the attempt. And while one can think of situations where the rationale would be important, in most situations that come readily to mind, it does not seem so.⁷⁹

Finally, it is interesting to consider attempts that cannot possibly succeed.⁸⁰ There are two types of such attempts. The first, for which it is often said liability should not be imposed, is exemplified by the case of a person who sticks pins in a voodoo doll, intending to kill his enemy.⁸¹ Here an objective observer might say that the type of act com-

78. See O. Holmes, *supra* note 58, at 56.

79. For instance, if a person puts poison in his intended victim's drink but the victim fails to succumb, it is true that this is some evidence that the act was less dangerous than one which produced a death—perhaps because the dosage of poison was too low. But this would not ordinarily seem to constitute enough evidence to lower the sanction significantly. In any event, there is less reason to lower the sanction than if the person had been interrupted before he completed the attempt, for then there would have been doubt about whether the attempt would have been completed, as well as whether it would have been successful.

80. See W. LaFare & A. Scott, *supra* note 57, § 60; G. Fletcher, *supra* note 57, § 3.3.3.

81. Note that because this person *believes* he is raising the probability of his enemy's death, we would say he intends his death under my definition.

mitted never causes harm, so that there is no reason to deter the act.⁸² The second kind of attempt, for which there would be liability, is illustrated by the case of a person who shoots a bullet into a dummy that he thinks is his enemy. In this instance, an objective observer would certainly say that the type of act committed creates positive expected harm, for shooting into objects that appear to be individuals will usually result in harm. Hence, according to deterrence theory, the act should be punished.

3. *Causation.* — When a party's act is followed by harm, two causal issues may arise.⁸³ The first concerns the question whether the act was the "necessary" cause (the "but for" cause) of the harm, that is, whether the harm would still have occurred in the absence of the act. Thus, if X poisons Y's drink and Y then dies, but an autopsy reveals that Y died of a heart attack, X's act would not be the necessary cause of Y's death. The criminal law ordinarily treats an act that was not the necessary cause of harm as if it were an attempt: the party is punished for the act, but less so than if the act were the necessary cause of the harm.⁸⁴

This outcome makes sense—though only partial sense—according to deterrence theory. It makes sense that acts which are not the necessary causes of harm are punished, for this enhances deterrence in the same way that punishment of attempts does, namely, by increasing the probability of sanctions.⁸⁵ The reason for the imposition of lesser sanctions, however, is not apparent. That acts sufficient to cause harm turn out not to be the necessary causes of harm is happenstance; it does not mean that the expected harmfulness of the acts was any lower. Therefore, the sanctions for such acts ought not to be diminished, according to deterrence theory.⁸⁶

If a party's act was the necessary cause of harm, the issue arises whether his act was the "proximate cause" of harm. Generally, acts said to be the proximate cause of harm can be recognized as acts which

82. This presumes that the person who failed with the voodoo doll would not have tried other ways of killing his enemy, such as shooting at him. If there is evidence that the person would have done this, then his act would appropriately be defined as "trying to kill an enemy by some means" rather than "trying to kill an enemy using a voodoo doll," and there would be a reason to punish him.

83. See W. LaFare & A. Scott, *supra* note 57, § 35.

84. This situation is different under tort law, where a party usually escapes liability if his act was not the necessary cause of harm. See Prosser & Keeton, *supra* note 70, § 41, at 265.

85. In the usual model of torts, there is no reason to enhance deterrence by use of sanctions where a party's act is not the necessary cause of harm. Assuming that parties have the assets to pay for harm done, and that they would not escape suit, the threat of liability only when their acts are necessary causes of harm is enough to induce parties to take adequate care. See Shavell, *An Analysis of Causation and the Scope of Liability in the Law of Torts*, 9 J. Legal Stud. 463 (1980).

86. The fact that Y died of a heart attack does not cast doubt on the potency of the poison. Note that this is in contrast to a failed or interrupted attempt to murder Y by poisoning.

substantially increased the probability or magnitude of the type of harm that occurred. To illustrate, if *X* severely beats *Y* and *Y* later dies in the hospital from internal injuries, we would probably say that *X*'s actions were the proximate cause of *Y*'s death. If, however, *Y* dies on the way to the hospital in an automobile accident, we would probably not say that *X*'s actions were the proximate cause of *Y*'s death.⁸⁷ In determining punishment, the criminal law usually takes into account whether or not the harm was proximately caused. *X* might be held liable for murder when *Y* dies from his internal injuries, but not when *Y* dies in the automobile accident.

Punishing a party for harm that he proximately caused might sometimes be justified in view of the evidentiary value of the actual harm done. *Y*'s death from internal injuries may be indicative of the severity of the beating he received, whereas his death in an automobile accident would not convey such information. In addition, considerations of marginal deterrence may justify greater punishment for proximately caused harms. *X* might decide against beating *Y* very badly if he knows that should *Y* die in the hospital from his injuries, *X*'s punishment would increase. No such incentive would result were *X* to be punished for *Y*'s death in an automobile accident.⁸⁸

4. *Responsibility*. — The criminal law sanctions a party who commits a socially undesirable act only if he is "responsible," in whole or in part, for his behavior. A party may not be responsible due to insanity, automatism, involuntary intoxication, or youth.⁸⁹ That defenses based on these conditions eliminate or diminish liability has an obvious potential justification according to deterrence theory, since the conditions make it unlikely that the use of sanctions would accomplish significant deterrence.⁹⁰ An insane or involuntarily intoxicated person, for example, presumably cannot be deterred from committing certain acts by the threat of punishment, so that the elimination of punishment might be appropriate. And it might be proper to impose diminished sanctions on young persons where, although there may be some possibility of deterring them, the dangerousness of their acts and their ability to escape sanctions is less than that of adults.

At the same time, deterrence theory suggests two general reasons for restricting such limitation of liability. First, parties might success-

87. The probability of dying in an automobile accident on a single trip (even if by ambulance) is small, and is therefore not much increased by *X*'s beating. In any event, *Y* might have been going somewhere else by automobile if he had not been going to the hospital.

88. That the logic of this paragraph applies generally and is not a feature of my examples can be appreciated from the characterization of proximately caused harms as those of a type whose probability or severity was increased in a substantial way by a party's act. It is exactly where this is true that the occurrence of the outcome may convey useful information about how much the party's act increased the expected harm.

89. See W. LaFare & A. Scott, *supra* note 57, ch. 4.

90. This is stressed by Bentham, *supra* note 2, at 164.

fully feign the conditions which limit responsibility. This possibility is more significant with respect to some of the defenses (for example, insanity) than to others (for example, youth). Second, parties might act in ways that make their conditions especially dangerous. An epileptic might drive an automobile, or a person subject to insane rages might decide to purchase a gun. The imposition of liability could induce these parties to act differently and thereby reduce dangers over which they later would have no control.

5. *Justification and Excuse.* — In addition to the defenses that fall under the rubric of responsibility are defenses ordinarily placed under the heading of justification and excuse. They will be considered separately since the issues they present are different.

a. *Ignorance and Mistake.* — If a party claims that he was unaware that his act was unlawful, he will ordinarily be found liable anyway. He may escape liability if he had little opportunity to learn about the law (as with an unpublished or little-known ordinance) or if he was acting in reliance on a mistaken interpretation of the law made by a court or an appropriate government officer.⁹¹

Such an approach is clearly suggested by the theory of deterrence.⁹² If a party is held liable for violating widely-known laws or laws that can be learned through reasonable effort, whether or not he in fact claims to know such laws, he will have an incentive to learn the law and adhere to it. If, however, a reasonable effort is insufficient to learn a law, it is best to permit parties to escape liability, since a party can be deterred by possible sanctions only if he knows to which acts the sanctions will be applied.

A second kind of mistake arises where a party commits an act that he believes to be different in some material respect from what it is in fact. A mistake of this kind is where a party commits an act that he believes to be innocent although it is actually harmful. In the classic case, a person takes an umbrella from a restaurant assuming that it is his, when it really belongs to someone else. As deterrence theory suggests is desirable, there is no criminal liability in instances like this.⁹³ Assuming that acts believed to be harmless usually are harmless, the expected harm associated with the acts is too low to warrant use of sanctions.

Another category of mistake arises where a party knowingly commits an undesirable act but believes it to be either more or less harmful than it is in fact. For example, an individual might steal a valuable piece of jewelry, thinking it a mere bauble, or he might shoot to kill the "person" who turns out to be a dummy. The legal principles employed in these situations are not uniform; although sanctions are frequently

91. See generally W. LaFare & A. Scott, *supra* note 57, § 47 (general discussion of mistake of law doctrine).

92. See Bentham, *supra* note 2, at 164.

93. See W. LaFare & A. Scott, *supra* note 57, § 47, at 357.

based on what a party did in fact, many times they are affected by what he thought he was doing.⁹⁴

There are two reasons why the harmfulness of the act the party thought he was committing should influence the sanction. First, the benefits a party expects to derive from committing an act and thus the ability to deter him, depend on what he thinks he is doing, not on what he is in fact doing.⁹⁵ Second, the expected harm associated with an act may be more closely related to what the party thinks it is than to what it turns out to be in the particular instance. The act of taking what one thinks is a bauble might usually mean that only a bauble is missing; the act of shooting at what one thinks is a person will usually be very harmful.

Nevertheless, one important factor suggests that the actual harm should influence the sanction. Parties may be able to convince the courts that they thought they were doing little harm when in truth they knew they were doing greater harm. If so, and if the sanction is based on the courts' erroneous assessment of parties' beliefs, sanctions will be too low and diminished deterrence will result.⁹⁶ Hence, there is some reason to raise the sanction where the actual harm exceeds what the party claims to have thought it would be. But note that there is no corresponding argument for lowering the sanction where the actual harm turns out to be less than what the party thought it would be, since a party will have no incentive to exaggerate the harm he thought he was doing.

b. *Entrapment*. — A party may raise the defense of entrapment if a law enforcement official induces the party to commit a criminal act that he would not ordinarily commit.⁹⁷ Where, for instance, a game warden induces a hunter to shoot at bald eagles and the hunter would not otherwise have done this, the hunter can assert the defense of entrapment.

The argument for this defense on grounds of deterrence theory is that if parties would not ordinarily commit criminal acts, there is no behavior that needs to be deterred. Thus, any punishment of the parties and any effort devoted to their entrapment must be considered a social waste;⁹⁸ moreover, their entrapment might also result in the ac-

94. See *id.* § 47, at 360.

95. The individual who thinks he is stealing a bauble will be easy to deter, but not the individual who knows he is stealing valuable jewelry.

96. Individuals may not be properly deterred from stealing valuable jewelry if they know they can convince the courts that they thought the jewelry was only a bauble.

97. See generally W. LaFare & A. Scott, *supra* note 57, § 48 (general treatment of entrapment).

98. To clarify this point, consider a situation where a person would *never* commit a criminal act if not entrapped. Here, plainly, punishing the person and devoting effort to entrap him is wasteful, since otherwise the person would not cause harm. It is irrelevant under deterrence theory that he might be thought a "bad" person because he could be induced to commit a criminal act in certain contrived circumstances.

tual doing of harm.⁹⁹ Hence, it is best not to punish the parties, and to discourage entrapment. The former is directly accomplished by allowing the defense; and the latter is indirectly accomplished, since enforcement officials do not derive the benefit of having been responsible for additional criminal convictions.

On the other hand, the defense of entrapment is not justifiable where parties would ordinarily commit criminal acts. In such a case, it is desirable to deter the parties. Therefore, it may be useful to employ certain law enforcement activity, including deception and subterfuge, to raise adequately the probability of their apprehension.

c. *Duress*. — A party will not be held liable for a harmful act if he committed the act only because of a threat of death or serious injury. To invoke the defense of duress, the threat must have been both imminent and credible, and the party must not have killed another person (although the sanction may still be mitigated in that case). Whether or not the defense is available, the threatening party will be liable for the act he induced.¹⁰⁰

The defense of duress is obviously desirable if the threatened party truly cannot be deterred from committing the act.¹⁰¹ Hence, the law's insistence on the imminence and credibility of the threat, and on its being one of serious injury or death seems understandable. But its refusal to allow the defense where the threatened party has killed someone does not seem rational according to deterrence theory, since it is quite possible that the threatened party could not have been deterred from killing; after all, he will often be comparing an immediate threat to a sanction that will not be immediate, if it is imposed at all. In any case, it is desirable that a party who makes a threat be held liable for crimes committed as a result of that threat, for this will be necessary to deter parties from using threats as a device to circumvent the law.

d. *Necessity*. — The defense of necessity may be asserted when a party, forced by circumstances to choose between two harmful acts, chooses the less harmful act.¹⁰² This makes clear sense according to deterrence theory, since it is socially desirable for a party to minimize harm.

e. *Defense of Self, of Another, or of Property*. — The law regarding self-defense and protection of others and of property is, roughly, that one may use the amount of force apparently necessary to ward off an aggressor whose threat one believes is unlawful and immediate, and who cannot be stopped by police intervention.¹⁰³ Plainly, allowing the use of force will enhance deterrence of aggression. Limiting the justified

99. For example, the game warden might not be able to take the hunter into custody before he shoots an eagle.

100. See W. LaFare & A. Scott, *supra* note 57, § 49.

101. This is pointed out by Bentham, *supra* note 2, at 165.

102. See W. LaFare & A. Scott, *supra* note 57, § 50.

103. See *id.* §§ 53–55.

use of force makes sense under the presumption that the courts are better able than threatened parties to decide on sanctions.

f. *Consent*. — A party may sometimes escape criminal liability if the person affected by his act had consented to its commission. The defense of consent, however, is not available where serious bodily injury is done.¹⁰⁴ The justification for the defense, under deterrence theory, centers on the concept of harm: if consent is taken to mean that there is no harm, then there is no reason to deter acts to which someone has consented.¹⁰⁵ According to this reasoning, of course, a person's consent even to serious bodily injury apparently ought to be allowed as a defense. Yet consideration must be given to the counterarguments that the person's family and friends (or society at large) may feel that harm has been done, and also to the possibility that the injuring party could deceive the courts about the person's consent.

g. *Condonation and Settlement*. — The fact that a person who has suffered harm may later condone or settle with the responsible party¹⁰⁶ may not be used as a defense against criminal liability.¹⁰⁷ If a person is robbed, for example, and he then discovers the robber and forgives him, the robber will still be liable. According to deterrence theory, the major reason for not allowing condonation as a defense is that deterrence would be diluted. Were the defense allowed, the "sanctions" imposed by victims—usually some form of apology, the return of property, or a payment, but never imprisonment—would be less than the sanctions the courts would otherwise impose.¹⁰⁸ Moreover, victims might many times wish to condone injuring parties, for there is no reason to believe that a victim's personal interest in punishing an injuring party would correspond to the social interest in deterrence. Finally, were the defense allowed, there might be a real danger that victims would be coerced into "condoning" injuring parties.

6. *Strict Liability*. — As has been emphasized throughout this discussion of criminal law, much more than the mere doing of harm is ordinarily required before punishment is imposed. In other words, strict liability is an unusual form of liability.¹⁰⁹ It is also true that according to the theory of deterrence, strict liability cannot usually be justified, since under strict liability, socially desirable acts are punished if they happen to result in harm.¹¹⁰ For strict liability to be justified,

104. See *id.* § 57.

105. See Bentham, *supra* note 2, at 163.

106. Condonation is distinct from giving the party prior consent.

107. See W. LaFare & A. Scott, *supra* note 57, § 57.

108. It should be observed that in tort law—where of course settlement is allowed—payments made in settlement would often approximate the expected court-determined monetary sanctions, since otherwise victims might well prefer not to settle.

109. However, liquor and narcotics laws, and pure food laws, often impose liability without fault. See W. LaFare & A. Scott, *supra* note 57, § 31.

110. This is undesirable of course both because it might result in deterrence of desirable acts and because the use of sanctions is socially costly. Suppose, for instance,

two conditions must hold. First, the courts must find it difficult or impossible to distinguish between desirable and undesirable acts. This means the courts must in effect opt between strict liability or no liability at all. Second, if there is no liability, the resulting level of undesirable acts must be unacceptably high.¹¹¹ Whether strict liability is actually used only where these two conditions are approximately met is not clear.¹¹² But the basic fact that strict liability is not the common form of liability makes sense.

CONCLUSION

I do not think that any real comment is required concerning the theoretical analysis of optimal deterrence in this Article.¹¹³ However, my discussion of the many ways in which the criminal law furthers the purposes of deterrence may have created a false impression about the nature of my beliefs in optimal deterrence as the basis for a descriptive theory of the criminal law. I realize that examples of inconsistency of the criminal law with optimal deterrence are not difficult to adduce (as I occasionally noted), that there are differences in the criminal law among countries that cannot be conveniently attributed to differences relevant to deterrence,¹¹⁴ that incapacitation and other goals omitted from the analysis help to explain many features of the criminal law,¹¹⁵ and that juries and the judiciary hardly think exclusively in terms of deterrence. Nevertheless, it certainly does appear that deterrence plays an important part in explaining the criminal law.

that druggists are strictly criminally liable for narcotics sales to unauthorized individuals. Then a druggist might decide against selling narcotics even though it might be socially desirable that he perform that service. And if the druggist does sell narcotics, he might be sanctioned even if through no fault of his own, he makes a sale to an unauthorized person, such as an individual whom he could not have determined was only posing as a physician.

111. It is interesting to note that these two conditions are needed because sanctions are assumed to be nonmonetary and socially costly. The conditions are not required to justify strict liability where sanctions are monetary and socially costless. Indeed, where that is the case, parties are in theory induced to act desirably under strict liability; thus, since the use of sanctions is costless, there is no special reason to avoid adopting strict liability. See, e.g., Shavell, *Strict Liability Versus Negligence*, 9 J. Legal Stud. 1 (1980).

112. At least the first condition may be roughly met. That the courts find it difficult to obtain information about a party's conduct is often stated to be a reason for use of strict liability. See W. LaFare & A. Scott, *supra* note 57, § 31, at 218.

113. See sections II.A. and III.A.

114. For instance, the United States appears to be the only country that allows the defense of entrapment. See G. Fletcher, *supra* note 57, § 7.3, at 541.

115. To give only one illustration, the fact that attempts are punished could be explained in part by saying that the commission of an attempt constitutes evidence that the party is of a type who needs to be incapacitated. And the fact that attempts are punished less than acts resulting in harm could be explained by saying that the evidence of the need for incapacitation is less where harm is not done, or that the retributive motive does not come much into play unless harm is done.

APPENDIX

1. A simple formal version of the model of deterrence discussed in the text is described here, and the arguments of section III.A. on the optimal use of nonmonetary sanctions are sketched.¹¹⁶

2. According to the model, parties choose whether to commit harmful acts from which they would derive private benefits. Let b be the private benefit a party would derive from committing a harmful act, where $0 \leq b \leq \bar{b}$,¹¹⁷ and let $h \geq 0$ be the harm that the act would cause; thus, a party is identified by the pair (b, h) . If a party commits an act, he will be apprehended with probability p and suffer a sanction s , where $0 \leq s \leq \bar{s}$;¹¹⁸ s will be determined by a sanctioning function, that is, a function of the variables that the courts can observe (as discussed below). The utility of a party who commits an act will be b if he is not sanctioned and $b - s$ if he is; if he does not commit an act, his utility will be 0. Therefore, he will commit an act if $b > ps$. Otherwise he will be deterred.¹¹⁹ Because s is bounded by \bar{s} , it will be impossible to deter a party if $b > p\bar{s}$.

3. Social welfare is comprised of the following elements for each party who commits a harmful act: the social value of the private benefits, namely βb , where $\beta \geq 0$; the harm h ; and the expected social cost of imposing sanctions, σps , where $\sigma > 0$. Social welfare then equals $\beta b - h - \sigma ps$ integrated over all individuals who commit acts, less the expense $c(p) > 0$ of maintaining the probability at p , where $c'(p) > 0$. The social problem is to choose a sanctioning function and p to maximize social welfare.

4. The first-best behavior of parties—that which would maximize social welfare if the parties' behavior could be commanded—is for parties to commit acts if and only if $\beta b > h$.

5. Suppose that the courts have perfect information about b and h . In this case the sanctioning function can depend directly on b and h , and the optimal sanctioning function is easy to describe: (i) $s = 0$ if $\beta b > h$, as then parties will commit desirable acts, and imposing sanctions would lower social welfare; (ii) $s = 0$ if $\beta b \leq h$ but $p\bar{s} < b$, for where parties cannot be deterred from committing undesirable acts, imposing sanctions would lower social welfare but not affect behavior; and (iii) $s \geq b/p$ if $\beta b \leq h$ and $p\bar{s} \geq b$, for where parties can be deterred from

116. The arguments are made in detail in Shavell, *The Optimal Use of Nonmonetary Sanctions as a Deterrent* (Sept. 1985) (unpublished manuscript) (on file at the offices of the Columbia Law Review).

117. The benefits and other components of a party's utility are assumed to be bounded because, as is well known, this follows from the axioms of expected utility theory. See *supra* note 37.

118. The probability of apprehension p is the same for all parties.

119. For simplicity, I adopt the convention that when the party is indifferent, he will be deterred, and I will adopt similar conventions below.

committing undesirable acts, this will be optimal to do since it will prevent harm, yet not involve the socially costly imposition of sanctions.

It is apparent from this description that sanctions are never actually imposed, and that the minimum sanction necessary to deter in case (iii) (namely b/p) depends on b but not on h .

The optimal probability of imposition of sanctions is found by maximizing social welfare, taking into account the optimal sanctioning system; and the first-order condition determining p equates $c'(p)$ with the losses due to the commission of harmful acts by parties who are just at the margin of deterrability.¹²⁰

6. Suppose now that although the courts can observe h , they have only imperfect information about b : rather than observing b , they observe $r = r(b, \theta)$, where θ is a random element and where in general there are a variety of b and θ such that $r = r(b, \theta)$. The optimal sanction given r and h is determined by maximizing

$$(1) \quad \int (\beta b - h - \sigma p s) f(b|r, h) db,$$

where f is the probability density of b conditional on r and h , and where the integration is performed over the set of parties of type (b, h) for whom $r = r(b, \theta)$ and who commit the act, that is, for whom $b > ps(r, h)$. The optimal s might be zero if the conditional distribution of βb is much higher than h (that is, if usually the act is desirable) or if the distribution of b is so high that parties are usually undeterrable. If the optimal s is positive and interior to $[0, \bar{s}]$, then it is determined by the first-order condition

$$(2) \quad \sigma p \text{Prob}[b \geq ps|r, h] = (h + \sigma ps - \beta ps) f(ps|r, h).$$

The left-hand side is the marginal cost of increasing s —the expected marginal social cost σp of increasing the sanction per party who commits a harmful act times its conditional likelihood—and the right-hand side is the marginal benefit of increasing s —the expected social savings per party just deterred, namely, $(h + \sigma ps - \beta ps)$, times the density of these individuals.

Notice that in contrast to the situation where the courts had perfect information, sanctions are here sometimes imposed at the optimal s (as reflected by the left-hand side of (2)), that the optimal s depends implicitly on the distribution of b conditional on r , and that the optimal s depends on h and increases with it, other things being equal. To see the latter, observe that equation (2) is of the form $g(s, h) = 0$. Implicitly differentiating this with respect to h and solving for $s'(h)$, we obtain $s'(h) = -g_h(s, h)/g_s(s, h)$. Now $g_s(s, h) < 0$ (the second-order condition for optimality of s), and assuming that the conditional density f does not change with h (to isolate the effect on s of a change in h), we have $g_h(s, h) = f(ps|r, h) > 0$, so that $s'(h) > 0$; the optimal sanction rises with the level of harm.

120. See Shavell, *supra* note 116, equation (7).

The optimal probability is determined in a way similar to that in the previous case.

7. Of course, a more general version of the model would incorporate the various possibilities mentioned in the text: an act would cause a probability distribution of harm instead of a single level of harm; the courts might not be able to observe this distribution even though they could observe the harm that actually occurs; private benefits might depend on the occurrence of harm; parties might choose among a variety of acts; and so forth.