

# DETERRENCE AND THE PUNISHMENT OF ATTEMPTS

STEVEN SHAVELL\*

## I. INTRODUCTION

By an "attempt" I shall simply mean a potentially harmful act that does not happen to result in harm, such as shooting at someone but missing.<sup>1</sup> The chief object of this article is to examine the punishment of attempts from the perspective of the theory of optimal deterrence of undesirable acts. Two basic questions are to be addressed. First, how may punishment of attempts be justified? And second, supposing that their punishment is justified, what should be the magnitude of the sanctions for attempts? In particular, should an attempt be punished by a lower sanction than a similar act that causes harm?<sup>2</sup>

In Section II of this article, I consider these questions employing a standard model of deterrence (the Appendix contains a formal analysis of

\* Professor of Law and Economics, Harvard Law School, and Research Associate, National Bureau of Economic Research. I wish to thank Lucian Bebchuk, Alan Dershowitz, Louis Kaplow, William Landes, Daniel Meltzer, and A. Mitchell Polinsky for comments, Eric Bilsky for research assistance, and the National Science Foundation (grant SES-8821400) for support.

<sup>1</sup> This definition will suffice for now even though it makes no reference to intent and is broader than the definition of attempt in criminal law; see Section III *infra*.

<sup>2</sup> Prior writing on deterrence and attempt is limited. Of note, however, are Jeremy Bentham, *The Theory of Legislation, Principles of the Penal Code*, Part IV, at ch. 15 (1st ed. 1864), dealing with both questions; Oliver Wendell Holmes, *The Common Law*, at lecture 2 (1st ed. 1881), discussing the second question; and Stephen J. Schulhofer, *Harm and Punishment: A Critique of Emphasis on the Results of Conduct in the Criminal Law*, 122 U. Pa. L. Rev. 1497, 1517-87 (1974), also emphasizing the second question. An earlier paper of mine, Steven Shavell, *Criminal Law and the Optimal Use of Nonmonetary Sanctions as a Deterrent*, 85 Colum. L. Rev. 1232, 1249-53 (1985), sketches the arguments to be elaborated here; Richard A. Posner, *An Economic Theory of the Criminal Law*, 85 Colum. L. Rev. 1193, 1217-19 (1985), contains remarks along similar lines; and Donald Wittman, *Prior Regulation versus Post Liability*, 6 J. Legal Stud. 193 (1977) analyzes closely related issues.

[*Journal of Legal Studies*, vol. XIX (June 1990)]

© 1990 by The University of Chicago. All rights reserved. 0047-2530/90/1902-0007\$01.50

the model). In the model, actors decide against committing potentially harmful acts if the sanctions they might face, including possible sanctions for attempts, outweigh in importance the benefits they hope to obtain. The main ideas developed in the model may be sketched as follows.

A partial answer to the initial question, Why punish attempts? is that punishment of attempts increases deterrence by raising the *probability* of imposing sanctions. Those who would shoot at another person will realize that they may be penalized if they miss as well as if they succeed; those contemplating engaging in theft will understand that they will suffer sanctions if they are apprehended trying to carry out theft as well as if they are caught red-handed with stolen goods. Punishing attempts thus increases deterrence by expanding the set of circumstances in which sanctions are imposed.

This point, however, is an incomplete justification for punishing attempts because, if society wants to augment deterrence, it apparently has an alternative to penalizing in broader circumstances. Society may elevate the *magnitude* of sanctions. Instead of punishing attempted theft, society may impose more severe sanctions for successful theft. Why can this method of increasing deterrence not always be employed? The reason given in the model is that there is an upper bound to punishment, such as life imprisonment. Once that bound has been set as the punishment for doing harm, the only avenue open for enhancing deterrence is raising the probability of punishment. Also, general reasons, not included in the model, may make it unwise to utilize the highest sanctions for causing any given type of harm.<sup>3</sup> One such reason is that it is desirable for sanctions for lesser harms to be low, and for the schedule of sanctions to rise with the magnitude of harm, so that people who commit bad acts will have an inducement to refrain from causing greater harms (to only beat a person, not kill him). Additionally, society may consider it unfair, and thus be unwilling, to impose sanctions that seem out of proportion to harm done (more than several years of imprisonment for car theft). If the size of sanctions for doing harm is effectively constrained for these reasons, then society must raise the probability of imposing sanctions in order to increase deterrence.

The argument is still not complete, though, because it must be asked why society should increase the probability of imposing sanctions by punishing attempts, rather than by punishing acts that do harm with greater frequency. Why punish attempted theft instead of catching and punishing successful thieves more often? The answer supplied in the

<sup>3</sup> These are mentioned in Section III.B *infra*.

model flows from the assumption that additional enforcement resources would be required to raise the probability of apprehension of those who do harm, whereas no resource increase is needed to apprehend at least a portion of those who commit attempts. This assumption is motivated by the presumption that people who commit attempts may be reported by witnesses or observed by enforcement agents in the course of their efforts to detect those who do harm. Apprehension of some persons who commit attempts is cheap, relative to the costs of comparably raising the probability of apprehending those who do harm.

My argument for punishing attempts is, in an important respect, not original. Since Jeremy Bentham, and probably before, punishing attempts has been regarded as a means of supplementing deterrence of undesirable behavior. What has not been investigated, however, is the matter of why raising either the magnitude or the probability of sanctions for completed harms is not a good substitute for punishing attempts.

Turning to the second question about the magnitude of punishment for attempts, the preliminary answer furnished in the model is that the sanction for an attempt need not be less than that for an identical act that results in harm. This conclusion holds only where courts can obtain perfect information about the dangerousness of acts. If courts' information about attempts is imperfect, the appropriate sanction for an attempt is typically less than that for an act with similar outward appearance but that results in harm. The explanation, in essence, is that the occurrence of harm may provide implicit evidence of the dangerousness of an act.

After developing these arguments, I proceed in Section III to an evaluation of the model and to issues going beyond it. Specifically, I comment on the generality of the analysis, pointing to its applicability to inchoate offenses and to acts that do positive but smaller harm than might have been anticipated. I then discuss the deterrence rationale for punishment of attempts: Does punishment of attempts create meaningful deterrence? What limitations are there on the severity of punishment? What is the role of intent? Should "impossible" attempts be punished? I next amplify the answer to the question of whether the sanction for attempts should be less than that for harm, examining inferences about dangerousness that may be drawn for incomplete and for complete attempts and inquiring about incentives to alter behavior during the course of attempts. I consider also the differences and the similarities between the legal treatment of attempts in criminal law and the treatment of the analogue of attempts in civil law, namely, of negligent acts that do not result in harm. Finally, in the concluding Section IV, I remark on the punishment of attempts implied by social goals different from deterrence.

## II. A MODEL OF ATTEMPTS AND OPTIMAL DETERRENCE

The primary features of the model are illustrated in Figure 1. An individual may commit a potentially harmful act from which he will obtain a benefit. If he commits such an act, harm will occur with some probability  $q$ , and harm will not occur with the complementary probability  $1 - q$ . In the latter case, we will say that the individual committed only an attempt. The benefit from committing an act, the probability of harm, and the magnitude of the potential harm are assumed to vary among individuals.

If an individual commits an act, he may be apprehended and suffer a sanction. The probability of sanctions, as well as their magnitude, may depend on whether he commits only an attempt or causes harm. It is assumed that there is a maximum possible sanction that can be imposed;<sup>4</sup> for purposes of numerical illustration, this will be taken to be 1,000.

The probability-discounted or *expected* sanction<sup>5</sup> facing an individual who commits an act equals the sum of two products: the probability  $1 - q$  of committing only an attempt, multiplied by the probability  $p_1$  of being punished for an attempt, multiplied by the sanction  $s_1$  for an attempt; plus the probability  $q$  of causing harm, multiplied by the probability  $p_2$  of being punished for causing harm, multiplied by the sanction  $s_2$  for harm. Suppose, for example, that an act will fail to cause harm—will result only in an attempt—with probability .60; that the likelihood of being punished for an attempt is .30; that the sanction for an attempt is 200; that the probab-

<sup>4</sup> The assumption that the disutility of sanctions is bounded is in fact a conclusion derivable from the postulates of expected utility theory. See Kenneth J. Arrow, *Essays in the Theory of Risk-bearing* 63–69 (1971); Leonard J. Savage, *The Foundations of Statistics* 81–82 (2d rev. ed. 1972). For discussions of practical reasons for thinking that sanctions are limited, see, for example, Frank H. Easterbrook, *Criminal Procedure as a Market System*, 12 J. Legal Stud. 289, 293–95 (1983); and Posner, *supra* note 2, at 1205–8.

The significance of the assumption is that it implies that the ability to deter is limited when sanctions are imposed with a probability. If the reader is troubled by the assumption, he may find two remarks helpful. The first I take from Shavell, *supra* note 2, at 1242, n.37: "It is likely that [the reader] 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."

Second, the reader may say that, whatever sanction one contemplates, a worse one can always be imagined. Execution by electrocution is not as bad as execution by hanging, and that is not as bad as being boiled in oil, and so forth. The point may well be true, but it does not contradict the assumption that the disutility of sanctions is bounded. The infinite sequence of ever-worse sanctions corresponds to an infinite increasing sequence of disutilities that approaches a bound asymptotically, rather than grows unboundedly.

<sup>5</sup> The expected sanction may be interpreted as the sanction that would be suffered on average were the act committed repeatedly. The definition of the expected value of an uncertain variable, here of a sanction, is conventional. See, for example, Robert S. Pindyck & Daniel L. Rubinfeld, *Microeconomics* ch. 5 (1989), for an introductory discussion of this and other concepts of expected utility theory.

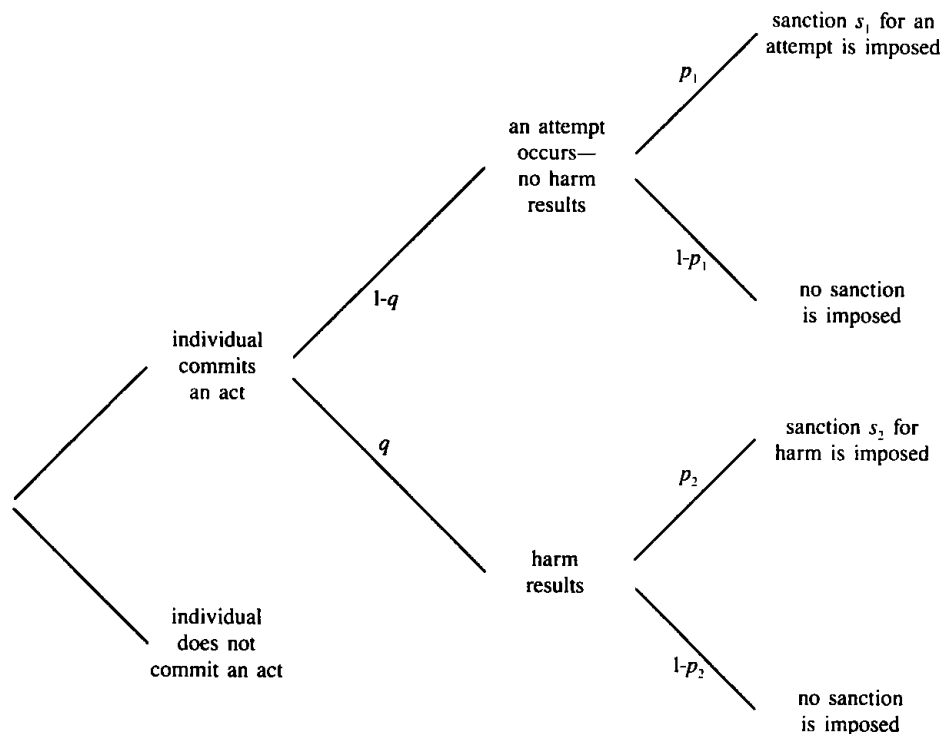


FIGURE 1.—The model

ity that the act will cause harm is .40; that the probability of being punished for causing harm is .50; and that the sanction for this is 300. Then the expected sanction for committing the act would be  $(.60)(.30)(200) + (.40)(.50)(300)$ , which equals  $36 + 60$ , or 96.

An individual will decide not to commit an act, and will be said to be deterred, if the expected sanction would exceed his benefit. Otherwise, he will commit the act.

The state is envisioned as choosing a system of sanctions and a level of enforcement effort determining the probabilities of imposition of sanctions. The state makes its choice in an optimal way, so as to maximize the following measure of social welfare: the benefits individuals obtain from committing acts, less the sum of the harm done, the costs, if any, of imposing sanctions, and enforcement effort.<sup>6</sup> In what follows, I will explain the logic of the analysis only in the case of monetary sanctions that are costless to impose,<sup>7</sup> and I will defer discussion about the choice of

<sup>6</sup> The assumption that the benefits obtained from committing acts enter into the criterion is simplifying but is not necessary to the argument to be presented.

<sup>7</sup> The assumption that sanctions are costless to impose is, again, made to simplify the analysis. Were sanctions nonmonetary and costly to impose, the qualitative nature of the conclusions would not be changed, as is noted in the Appendix.

enforcement effort. Thus, I will focus until later on the selection of sanctions given some probability of imposing sanctions for an attempt and some probability of imposing sanctions for causing harm.

Observe first that it is desirable for the expected sanction for an act to equal the expected harm caused by the act, that is, the probability of harm multiplied by its magnitude. If the expected sanction equals the expected harm, an individual will be deterred from committing an act if and only if the expected harm exceeds his benefit, a socially ideal result within the model. Accordingly, note, the expected sanction should be higher the higher the probability of harm or its magnitude is.

Now let us examine the implications for the punishment of attempts of the desirability of equating the expected sanction to the expected harm. We shall study two situations: where courts can obtain perfect information about acts, and where they cannot.

#### *A. Courts' Information about Acts Is Perfect*

In this case, if a court has before it a person who committed an attempt, the court will be able to ascertain both the probability that his act would have resulted in harm and the magnitude of the potential harm. Likewise, if a court has before it a person who caused harm, the court will be able to determine what the probability of its occurrence was *ex ante*. In either event, therefore, the court will be able to determine the expected harm associated with the particular act and to tailor the sanction to it.

The problem for the state then becomes to design for each type of act—classified by its probability of causing harm and the size of the potential harm—sanctions for an attempt and for causing harm, with the property that the expected sanction equals the expected harm.

Can this problem be solved? The answer is often yes, and, in fact, in a multiplicity of ways. Consider an act that causes harm of 100 with probability .40, so that the expected harm is 40. For this act, the state wants to choose a sanction  $s_1$  for an attempt and a sanction  $s_2$  for doing harm such that the expected sanction equals the expected harm of 40. If, as before, the probabilities of being penalized for an attempt and for harm are .30 and .50, respectively, the state's problem is to fashion sanctions such that the expected sanction of  $(.60)(.30)s_1 + (.40)(.50)s_2$  is 40, which is to say, such that  $.18s_1 + .20s_2$  is 40. There is an extra degree of freedom, so many different combinations—indeed, a whole continuum—of sanctions will work, such as  $s_1$  equal to 100 and  $s_2$  equal to 110; or  $s_1$  and  $s_2$  each equal to 105.26; or  $s_1$  equal to 120 and  $s_2$  equal to 92. As the last pair of sanctions illustrates, the sanction for causing harm can be smaller than the sanction for an attempt.

Because it is the expected sanction that is important, not the individual sanctions for an attempt and for doing harm, there is no reason for the sanction for an attempt to be positive, provided that an expected sanction equal to the expected harm can be created solely with a sanction for harm. In the example, this can be done. An expected sanction equal to 40 can be created with a sanction  $s_2$  of 200 since  $(.20)(200)$  is 40. A positive sanction for attempts is not needed.

If, however, the expected harm is sufficiently high, it will be necessary to punish attempts as well as harm in order to create an expected sanction equal to the expected harm. In the example, the maximal expected sanction that can be created with a sanction for harm alone is  $(.20)(1,000)$ , or 200, since 1,000 is the highest possible sanction. Consequently, if the expected harm exceeds 200, it will be desirable to punish attempts. Suppose, for instance, that the potential harm in the example is 600 rather than 100. Then the expected harm would be  $(.40)(600)$ , or 240, so that to create an expected sanction of 240, attempts must be punished; if  $s_2$  equals 1,000,  $s_1$  would have to be set equal to 222.22 for the expected sanction to equal 240.<sup>8</sup>

We may therefore summarize the answer to the two questions, given the probabilities of imposing sanctions, as follows. First, *punishment of attempts is desirable for acts whose expected harm is sufficiently high to create an appropriately high expected sanction*. Second, *the optimal magnitude of the sanction for an attempt is not uniquely determined* and thus can be equal to (or in theory exceed) the sanction for causing harm. All that is uniquely determined is the magnitude of the expected sanction.<sup>9</sup>

It may be useful to amplify why only the combination of sanctions, as represented by the expected sanction, matters. The sole effect of sanctions in the model is on an individual's decision about committing an act. When an individual makes this decision, he does not know, if he acts,

<sup>8</sup> The threshold of expected harm above which it becomes optimal to punish attempts depends not only on the maximal sanction, here 1,000, but also on the probability that an individual will cause harm and be penalized, here  $(.40)(.50)$ , or .20. For these together determine the maximal expected sanction that can be created by use only of a sanction for causing harm. Thus, if we were considering an act that causes harm with probability .70 and a probability of imposition of sanctions for harm of .50, the maximal expected sanction that could be created would be  $(.35)(1,000)$ , or 350. Hence, 350 would be the threshold of expected harm above which it would become optimal to punish attempts.

<sup>9</sup> A slight qualification should be made. If the expected harm exceeds the maximal expected sanction, then not only is the expected sanction uniquely determined—it is equal to its maximal level (so are the sanctions for an attempt and for causing harm)—each is also equal to the maximal sanction. In the example, the expected sanction is  $.18s_1 + .20s_2$ , and therefore the maximal expected sanction is  $(.18)(1,000) + (.20)(1,000)$ , or 380. Thus, if the expected harm exceeds 380, the expected sanction equals 380, and the sanctions for an attempt and for causing harm are uniquely determined: each equals 1,000.

whether he will fail and be subject to the sanction for an attempt or will succeed and be subject to the sanction for doing harm. Hence, both of the sanctions will be of interest to him, in terms of their expected value. Moreover, after an individual decides to commit an act, the sanctions will have no effect on his behavior. Past the point of no return, he may be regarded as an automaton, who has set into motion forces that either will or will not lead to harm.<sup>10</sup>

Let us now address the issue of enforcement effort. We took enforcement effort and the probabilities of imposing sanctions above as fixed, even though enforcement effort is a variable that the state controls. The optimal degree of enforcement effort is determined by a balancing of the costs and benefits of increments in enforcement effort. The cost of additional enforcement effort is a resource cost borne by the state. The benefit is that raising the probabilities of imposing sanctions allows deterrence of some individuals who could not otherwise have been deterred, for when the probabilities of sanctions rise, the maximal expected sanction increases.<sup>11</sup>

Another point about enforcement effort should be made. Any given level of enforcement effort was presumed to be associated not only with a particular probability of imposing sanctions for causing harm but also with a particular probability of imposing sanctions for attempts. In other words, it was assumed that sanctions could be imposed for attempts with positive probability without further increasing enforcement effort. Hence, it was indeed the case, as stated in the Introduction, that society could increase deterrence for free by punishing attempts, whereas increasing deterrence comparably by raising the probability of sanctions for causing harm would involve incurring further resource costs.

### *B. Courts' Information about Acts Is Imperfect*

If a court is unable to obtain information about a characteristic of an act, then the court will be unable to make the level of the sanction depend on the unobserved characteristic. To understand the difficulties thus engendered, suppose that courts are unable to determine the probability of

<sup>10</sup> By contrast, if an individual makes decisions during the commission of an act, the sanction for an attempt will have importance in its own right. If the sanction for an attempt is low, the individual may be induced to abandon his attempt. This is discussed in Section IIIC *infra*.

<sup>11</sup> In the example, with probabilities of sanctions of .30 and .50, the maximal expected sanction for an act that causes harm with probability .40 is  $(.60)(30)(1,000) + (.40)(50)(1,000)$ , or 380. If the probabilities of sanctions are higher, for example, .35 and .55, the maximal expected sanction rises to  $(.60)(35)(1,000) + (.40)(55)(1,000)$ , or 430.



an act causing harm.<sup>12</sup> In this situation, courts will be unable to make sanctions depend on the probability of harm and, in particular, unable to make the sanction for attempts depend on the probability of success of an attempt. An attempt might not have been very dangerous because it involved only a low probability of harm, or it might have been quite dangerous; either way, the sanction for an attempt must be the same. A consequence is that the expected sanction for an act that is not very dangerous may be too high. Suppose that the sanction for an attempt is 100, that the probability of imposing sanctions for attempts is .30, as before, and that an act causes harm of 300 with a probability of just .01. Then the contribution to the expected sanction arising from the sanction for an attempt would be  $(.99)(.30)(100)$ , or 29.7 (since .99 is the probability of the act resulting in an attempt). But the appropriate expected sanction for the act is the expected harm,  $(.01)(300)$ , or 3, substantially less than 29.7.

The foregoing problem with punishing attempts—that those who commit acts that cause harm with a relatively low probability will face too high an expected sanction—can be remedied, or at least alleviated, by use of punishment for harm. The virtue of punishment for harm is that it causes the expected sanction to become linked in an *automatic* manner to the probability of harm. If punishment for harm is the only punishment, and an act leads to harm with a low probability, the expected sanction will be low because the probability of suffering the sanction for harm will be low; similarly, if an act causes harm with a high probability, the expected sanction will be high.

An illustration demonstrates that, if punishment for harm is the only punishment, it may be possible to create automatically the ideal expected sanctions for acts even though they vary in the probability of causing harm. Suppose that the potential harm due to acts is 100, that the probability of a sanction for harm is .50, and that the sanction for harm is 200. Then if an act causes harm with some probability  $q$ , the expected sanction will be  $q(.50)(200)$ , or  $100q$ , which is exactly the expected harm.<sup>13</sup> It is possible to create expected sanctions equal to expected harm here be-

<sup>12</sup> But, for simplicity, assume that a court is able to ascertain the magnitude of the potential harm (in the case of an attempt) or of the actual harm. In fact, of course, courts may lack information about the potential size of harm as well as about the probability of harm.

<sup>13</sup> Notice that *any* positive sanction for attempts will prevent this optimal result because the expected sanction for all acts with sufficiently small probabilities of causing harm will be too high. For example, a sanction of even 20 for attempts would lead to excessive sanctions for all acts that cause harm with a probability lower than about .10. (The expected sanction for an act that caused harm with probability .05, for instance, would be at least  $(.95)(.30)(20)$ , or 5.7, which exceeds the expected harm of 5.) Thus, the strategy of using a sanction for harm alone is the unique optimal strategy.

cause it is feasible to make the expected sanction conditional on the occurrence of harm fully equal to the harm (the expected sanction conditional on the occurrence of harm is  $(.50)(200)$ , or 100).

This result can be generalized. It can be proved that, if the magnitude of the potential harm is not too high, it is optimal to punish only for harm, with a sanction such that the expected sanction, conditional on the occurrence of harm, equals the harm. To elaborate, we have assumed that the maximal sanction is 1,000, so that the maximal expected sanction conditional on the occurrence of harm is  $(.50)(1,000)$ , or 500. Thus, if the magnitude of the potential harm does not exceed 500, it is optimal to punish only for harm, with a sanction equal to twice the harm.

If the size of the potential harm is too high, punishment for harm alone will not create an adequate expected sanction. Suppose, in the example, that the magnitude of the potential harm exceeds 500, say it is 800. Then even if the sanction for harm is maximal, 1,000, the expected sanction may be too low, only 500, if there is no sanction for attempts. That would be the case whenever the expected harm exceeds 500, which it does whenever the probability of harm exceeds .625 (for  $(.625)(800)$  is 500).

Hence, as in Section IIA above, *it is desirable to punish attempts for acts for which the potential harm is sufficiently high*. Unlike before, however, punishment of attempts has the disadvantage that those who cause harm with a low probability face excessive expected sanctions. Several implications follow. First, *optimal sanctions for attempts are positive only when sanctions for causing harm are maximal*. Second, *the optimal sanction for an attempt is uniquely determined* and represents a best compromise between the advantage of increasing deterrence of acts that cause harm with high probability and the disadvantage of creating excessive sanctions for acts that cause harm with low probability. Third, this best compromise leads to a higher optimal sanction for acts that are more likely to cause harm with high probability. Thus, *optimal sanctions for attempts rise with the location of the distribution of probabilities of causing harm*. Fourth, *optimal sanctions for attempts rise with the size of the potential harm*. Finally, it should be observed that were we to assume that a court could obtain some, rather than no, information about the probability that an attempt would have succeeded, then the optimal sanction for an attempt would reflect that information or, more generally, information about the degree of expected harm.

### III. DISCUSSION

I consider here a variety of issues that will help us to appraise, interpret, and extend the above analysis. The first subsection comments on the

generality of the analysis; the second, on the deterrence rationale for punishment of attempt; the third, on the question whether the sanction for an attempt should<sup>14</sup> be lower than that for harm; and the fourth, on the legal treatment of the counterpart of attempts in the civil law, risk-creating acts that do not result in harm.

### A. *Generality of the Analysis*

The analysis treated any potentially harmful act that does not do harm as an attempt. This class of acts is broader than the class of attempts in criminal law.<sup>15</sup> Carrying a concealed weapon, for example, does not constitute an attempt under criminal law; it is classified as a possession offense. Yet carrying a concealed weapon is a potentially dangerous act because it may lead to harm, such as robbery, but it is not harmful in itself. Thus, the analysis applies to the “attempt” of carrying a concealed weapon. In particular, if the threat of punishment for the harms to which the carrying of concealed weapons may lead is not an adequate deterrent, then punishing the carrying of concealed weapons will desirably augment deterrence. Similar observations may be made about other possession offenses, about conspiracy to commit criminal acts, about solicitation offenses, and about other inchoate offenses.<sup>16</sup> None are harmful in themselves, but each may lead to harm.

There is another regard in which the analysis is more general than it may seem. The analysis extends in a straightforward way to acts that result in positive but lower harm than might have occurred. A person may shoot at another and only graze the victim, even though death could well have resulted because he had aimed at a vital spot. In such a case, essentially the same argument advanced to explain why a shot that misses should be punished applies to the shot that grazes the victim: if we do not punish, and rather substantially, shots that graze, then we may fail to deter adequately shootings that cause death, and so forth.

<sup>14</sup> When I use the words “should” or “ought” in this section, they will refer to what appears to be best given the goal of deterrence, not to what may be best given other goals of criminal law as well.

<sup>15</sup> On criminal law and attempts, see generally George Fletcher, *Rethinking Criminal Law*, § 3.3 (1978); Wayne R. LaFare & Austin W. Scott, Jr., *Criminal Law*, §§ 59–60 (1972); and Stephen J. Schulhofer, *Attempt*, in 1 *Encyclopedia of Crime and Justice* 91 (1983). On the history of the legal treatment of attempt, see Jerome Hall, *Criminal Attempt—a Study of the Foundations of Criminal Liability*, 49 *Yale L. J.* 789 (1940).

<sup>16</sup> See, for example, LaFare & Scott, *supra* note 15, at 182, on crimes of possession, at § 58 on solicitation, and at §§ 61–62 on conspiracy.

### B. *The Deterrence Rationale for Punishment of Attempt*

In this subsection, I examine several issues bearing on the deterrence justification for the punishment of attempts.

1. *Degree of Deterrence Created by the Punishment of Attempt.* That punishment of attempts will augment deterrence is, I assume, undeniable as a logical proposition. But one may wonder about its significance. Is it reasonable to believe that the prospect of punishment of attempts creates meaningful deterrence? Before I offer an answer to this question, it will be helpful to distinguish among types of attempt.

Attempts fall into two basic categories: complete and incomplete.<sup>17</sup> A *complete attempt* does not do harm even though all steps necessary to the doing of harm are taken. The man who shoots a gun at another but misses has committed a complete attempt. If, however, the man only points his gun at another, he has committed an *incomplete attempt* because he has not taken all the steps necessary to the doing of harm.<sup>18</sup>

Incomplete attempts can themselves be subdivided into interrupted attempts and abandoned attempts. An *interrupted attempt* is one that is broken off by the intercession of another person, such as where an individual is prevented from firing his gun by someone who grabs it away. An *abandoned attempt*<sup>19</sup> is broken off by the person who is committing it, not by the intercession of another. The impetus to abandon an attempt may be guilty second thoughts, loss of nerve, and the like. On the other hand, the person beginning the attempt may notice that his chances of being identified and apprehended are higher than he had thought—perhaps he sees a policeman approaching. If he desists for this reason, the attempt will be said to be *strategically abandoned*.

Now let us address the question of whether the punishment of attempts

<sup>17</sup> For general discussions of types of attempt, see note 15 *supra*.

<sup>18</sup> It will be useful to employ the framework of decision theory (on which, see Arrow, *supra* note 4, and Savage, *supra* note 4) to make the distinction clear. In this framework, there are *acts*, designated by  $A_1, A_2, \dots$ ; *states of nature*, that is, possible circumstances, designated by  $S_1, S_2, \dots$ ; and *outcomes*, which in the context will either be  $N$ , meaning no harm, or  $H$ , meaning harm. Given an act  $A_j$ , the state of nature  $S_i$  determines the outcome; that is,  $f(A_j, S_i) = N$  or  $H$ , where the function  $f$  describes whether harm occurs given the act and the state of nature. An *incomplete attempt* is an act  $A_j$  such that  $f(A_j, S_i) = N$  for all states of nature  $S_i$ ; that is, harm does not occur under any circumstances because something necessary to the occurrence of harm was not done. A *complete attempt* is an act  $A_j$  such that  $f(A_j, S_i) = H$  for some state(s) of nature  $S_i$ ; that is, the act  $A_j$  will result in harm under some circumstances (perhaps many).

<sup>19</sup> On this subject generally, see Fletcher, *supra* note 15, at 184–97; and LaFave & Scott, *supra* note 15, at 446–51; and see also Daniel Rotenberg, Withdrawal as a Defense to Relational Crimes, Wis. L. Rev., 1962, at 596.

is likely to be able to deter. The answer depends on the probability that a person can be apprehended and punished for committing an attempt. If this probability is significant relative to the probability of punishment for actually doing harm, then it seems as sensible to suppose that punishment of attempts deters (or could do so given sufficiently high sanctions) as it is to believe that punishment for doing harm deters.

The probability of apprehension for an attempt equals the sum of the probabilities of apprehension for an incomplete attempt and for a complete attempt. For many types of acts, one supposes that the sum of these probabilities is in the range of (or exceeds) the probability of apprehension for doing harm. When crimes require preparation, as is frequently the case with robbery, theft, or fraud, for example, the probability of apprehension before the attempt is complete strikes one as often not small because the ability to conceal preparations may be limited. And when, as is sometimes true, complete attempts have a good chance of failing (the intended victim of a robbery may run away), there is another reason to believe that the likelihood of apprehension for an attempt may not be small compared to the likelihood of apprehension for doing harm. Punishment of attempts would seem to hold little promise of deterrence, relative to punishment for doing harm, mainly for the following set of acts: those committed with little or no preparation and yet for which the odds of success are high (shooting someone point blank in the heat of passion), and those for which preparations are easily concealed and for which the odds of success are high.

In taking this view, I reject the skeptical, if not dismissive, opinion that punishment of attempt cannot effectively deter. A good example is provided in the introduction to Article 5 (Inchoate Crimes) of the Model Penal Code (1985),<sup>20</sup> which reads in part, "Since these offenses [attempts] always presuppose a purpose to commit another crime, it is doubtful that the threat of punishment for their commission can significantly add to the deterrent efficacy of the sanction—which the actor by hypothesis ignores—that is threatened for the crime that is his objective." There are two responses to this statement. First, the statement is illogical. It is true that a person who has decided to commit a crime has decided to risk (not necessarily ignore) possible sanctions. But that hardly means that sanctions for attempts cannot deter; there may well be individuals who do not commit crimes just because of these sanctions. Second, as long as the probability of suffering a sanction for an attempt is not small relative to the probability of suffering a sanction for doing harm, then one cannot

<sup>20</sup> Model Penal Code, Article 5 (1985), at 293.

claim that punishment of attempts is inefficacious, whereas punishment for harm done may be worthwhile.<sup>21</sup>

Another point, concerning strategic abandonment of attempts, deserves emphasis. It is a good guess that if this type of attempt were not punished, there would be a peculiar and perverse effect enabling the successful commission of many crimes. Burglars could cruise the streets, not worrying particularly about concealing burglary tools or about being spotted. If police or potential witnesses were nearby, all burglars would need to do would be to go elsewhere, or wait until later. They could pick the perfect place and moment to complete their criminal acts. Thus, I suspect that the level of burglary—and of any other crime where picking the right occasion raises the probability of escaping detection or of success in the act itself—would be substantially higher than it now is.

2. *Need for Punishing Attempts and Limitation of the Magnitude of the Sanction for Harm.* The need for punishing attempts, as explained in the model, derived solely from the inadequacy as a deterrent of use of the maximal sanction for doing harm. Punishing an attempt was advantageous only when the threat of the maximal sanction for doing harm would not result in a sufficient level of deterrence. Indeed, in the case where courts' information was assumed to be imperfect, the conclusion was that punishment of attempts was appropriate only when the sanction for harm was maximal.

Yet, in fact, we often see punishment of attempts where sanctions for harm are far short of maximal and where one surmises that, were maximal sanctions for harm employed, deterrence might not be inadequate. The example from the Introduction is illustrative. A person may be punished for attempted theft even though the sanction for actual theft is unlikely to approach life imprisonment, and if the sanction for most theft were life imprisonment, one supposes that much of it would be deterred. These observations appear to raise questions about the descriptive worth of the theory advanced in the model.

But, on reflection, and as suggested in the Introduction, the theory leads to the conclusion that punishment of attempt may be desirable as long as there is any reason for limiting the sanction for doing harm. The reason need not be that an absolute bound, like life imprisonment, has been employed. Further, any reason for limiting the sanction for harm in fact serves to strengthen the deterrent argument for punishing attempts, for any constraint on the size of sanctions for harm reduces the deterrent value of punishing completed harms.

<sup>21</sup> For further argument against the view stated in the Model Penal Code, see Schulhofer, *supra* note 2, at 1533–54.

There are, as was noted, two reasons why sanctions for harm may be constrained in magnitude. The first is itself deterrence related; it is that "marginal deterrence" requires a gradation in sanctions. We want individuals who are not deterred from committing bad acts to do less rather than more harm when committing their acts. The classic case is that of a kidnapper; we would like him to spare the life of his victim. If the sanction for kidnapping is lower than that for killing a kidnap victim, the kidnapper will have an inducement not to kill his victim. Considerations of marginal deterrence, therefore, imply that the schedule of sanctions rises with the magnitude of harm. Overall, this acts as a limitation on the size of sanctions.<sup>22</sup>

The other general constraint on the size of sanctions is that society may desire that there be some proportionality maintained between the gravity of bad acts and the severity of punishment for them. If so, the size of sanctions for doing harm is limited by the proportionally fair magnitude.

3. *Intent and Punishment of Attempt.* In criminal law, the presence of "intent" is generally a prerequisite for punishment and, in particular, for punishment of attempt. This makes sense from the standpoint of deterrence. As I have argued elsewhere, intent is an indicator both of the expected harmfulness of acts and of the difficulty of deterring them.<sup>23</sup> In brief, intent is correlated with the expected harmfulness of acts because a person who intends a result is more likely to achieve it than someone who does not. Additionally, intent is correlated with the difficulty of deterring acts for two reasons. First, those who intend a result presumably obtain greater benefits from its occurrence than those who do not; thus, the former are more difficult to deter. Second, those who intend a result are more likely to plan to escape detection than those who do not. Since, then, intent is associated with greater difficulty in accomplishing deterrence and with greater expected harm, intent signals the need to augment deterrence and thus the desirability of punishment of attempt.

4. *Impossible Attempts.* The so-called impossible attempt has been of concern to commentators and the courts, and it is of interest to exam-

<sup>22</sup> On marginal deterrence, see George Stigler, *The Optimum Enforcement of Laws*, 78 J. Pol. Econ. 526 (1970); Steven Shavell, *A Note on Optimal Deterrence When Individuals Choose among Harmful Acts* (Discussion Paper No. 57, Harvard Law School, Program in Law and Economics, 1989); and Louis Wilde, *Criminal Choice, Nonmonetary Sanctions and Marginal Deterrence* (unpublished manuscript, Cal. Inst. Tech., Dept. of Economics, May 1989).

<sup>23</sup> See Shavell, *supra* note 2, at 1247-49. Holmes, *supra* note 2, makes a similar argument (emphasizing intent as an indicator of the probability of harm); as does Posner, *supra* note 2, at 1221-23.

ine this topic in light of the theory of deterrence.<sup>24</sup> Let me first mention several examples of impossible attempt and indicate a fundamental problem with the concept of impossibility. One oft-given illustration is sticking pins in a voodoo doll, the object being to kill the person in whose image the doll is made.<sup>25</sup> On the theory that one cannot kill in this manner, the attempt is termed impossible. Another example of an impossible attempt is shooting a dummy believed to be a person, or picking an empty pocket (it being impossible to steal money from an empty pocket).

Although labeling attempts like these impossible may seem appealing, it becomes clear after some thought that "impossible attempts" do not constitute a proper subcategory of attempts. Any attempt, no matter how ordinary it may appear, can be recognized as an impossible attempt. Consider shooting at someone and missing. Whatever the explanation is for a bullet's failure to hit its mark—perhaps a gust of wind, perhaps an improperly sighted gun—it can be said to have made the attempt impossible. More generally, whatever circumstance causes an attempt to fail can be said to have made it impossible. In formal terms,<sup>26</sup> an attempt that fails is an act  $A$  that did not result in harm because the state of nature  $S_i$  was such that no harm occurred. Whatever this state  $S_i$  is makes the attempt impossible; all attempts are impossible in this sense.

Still, suppose it is urged: "When a person shoots into a dummy, it is known by us, as onlookers, that the state  $S_d$ —that the "victim" is a dummy—makes the attempt doomed from the outset; it is therefore impossible for harm to come about. Yet, when a person shoots at a live target, the attempt is not doomed from the outset; it will fail only if the state  $S_w$  is such that the wind picks up; it is thus possible for harm to come about." This argument relies on a temporal distinction, that the attempt on the dummy is doomed from the outset since  $S_d$  is known by us at the outset. Yet the attempt on the live target is not doomed from the outset since the state  $S_w$  is not known by us at the outset. But this distinction is artificial. If we were intelligent enough to master meteorology we would know what the wind speed would be, and we could say, just as in the case of the dummy, that the attempt on the live target was doomed from the start.<sup>27</sup>

<sup>24</sup> On impossible attempts, see, for example, Fletcher, *supra* note 15, at § 3.3.3; and LaFave & Scott, *supra* note 15, at § 60.

<sup>25</sup> See, for example, LaFave & Scott, *supra* note 15, at 445.

<sup>26</sup> See note 18 *supra*.

<sup>27</sup> The opinion that impossible attempts cannot be distinguished from "ordinary" attempts has also been expressed by others. See, for instance, Alan M. Dershowitz, *Why Do Criminal Attempts Fail?* 70 Yale L. J. 160, at n.17 (1960); and H. L. A. Hart, *The House of Lords on Attempting the Impossible*, in *Crime, Proof, & Punishment*, 1, 12 (1981).



Since, then, all attempts can be regarded as impossible, any argument about punishment of attempts that rests only on the notion of impossibility devolves into an argument that must apply to all attempts. Thus, for instance, an argument that impossible attempts should not be punished would lead to the absurdity that all attempts should be free from punishment.

The subcategory of attempts for which I think the courts and commentators have been groping is simply the class of acts that did not do harm and whose *probability* of causing harm was zero. An act in this class does no harm no matter what the circumstances. It constitutes an impossible attempt in an *ex ante* sense. Most acts are not impossible in this *ex ante* sense because they will cause harm with positive probability, even though, if they do not cause harm, they can be said, *ex post*, to have been impossible.

Of course, an act whose probability of causing harm is zero should not be punished; there is no reason for society to bear any costs in punishing it since it creates no expected harm. Thus, sticking pins in a voodoo doll should not be punished, given the assumption that this act can never result in harm.

One should be cautious, however, about the interpretation of a person's act. Suppose that we have reason to believe that a person who fails to kill with a voodoo doll might decide to use a knife or a gun to achieve his objective. Then we may regard his attempt with the voodoo doll as an unsuccessful carrying out of the act "trying to kill by some means." This act causes harm with positive probability and may warrant punishment. A similar analysis applies to a case where someone attempts to smuggle a good that he believes to be subject to tariff but that is not in fact. If we think that this is the only good the person would ever want to bring into the country, then the correct interpretation of the act is that it is harmless, and he ought not be punished. If he might wish to bring in other goods that are subject to tariff, then his act should be conceived as involving a positive probability of harm, and therefore it may be appropriate to punish him.<sup>28</sup>

Whatever these complications, the case with shooting a dummy is quite different, for that act is clearly associated with a high probability of harm. Shooting a thing believed to be a person is likely to cause harm because usually that "thing" will in fact be a person. Likewise, picking an empty

<sup>28</sup> The distinction under discussion may be formally illustrated. Suppose that taking some action (smuggling a good) will be harmless given state  $S_1$  (the good is not subject to tariff) and harmful given state  $S_2$  (the good is subject to tariff). If a person will take the action only given the state  $S_1$ , his act will never cause harm. If a person will take the action whether the state is  $S_1$  or  $S_2$ , then his act will cause harm with positive probability.

pocket creates a not insubstantial probability of theft since picked pockets will often contain wallets. Hence, there is ample reason to penalize such attempts.

In sum, we can reformulate the idea of “impossible” attempts so that it corresponds to those acts that carry with them a negligible probability of harm. Such acts should not be penalized because there is no need for society to bear the costs of punishing them. By contrast, the traditional notion of impossibility is not helpful because it does not turn out to define a subcategory of attempts; any unsuccessful attempt, even if it is the outcome of an act highly likely to cause great harm, can be regarded as having been impossible after the fact.

### *C. The Sanction for an Attempt versus the Sanction for Harm*

I deal here with two topics that bear on the question whether, for reasons of deterrence, an attempt should be punished less than a completed harm. The first topic concerns the dangerousness of attempts, and the second, incentives to alter conduct during the commission of attempts.

1. *Dangerousness of Attempts.* It was stated in the analysis that where courts’ information is imperfect, the sanction for an attempt should reflect the likely probability of harm (more generally, the expected harm). In this Subsection I will ask when the probability of harm is lower for an attempt than for an act that turns out to cause harm. If that is so, the sanction for an attempt should be lower than the sanction for an act that causes harm, other things being equal.

By the “probability” of harm, I will not usually mean the probability of harm from the particular act a person actually committed, for the court will generally be supposed to have only imperfect information about the act and not to know the true probability. Instead, I will refer to the rational inference the court draws about the true probability of harm—in the form of an expected probability, using Bayes’s rule—after taking into account the possible acts that could have led to the observed attempt.

Consider, first, incomplete attempts. By definition, an incomplete attempt could not have resulted in harm because steps essential to the causing of harm were not taken. Accordingly, there is some chance that the missing steps would not have been taken. Hence, the probability that an incomplete attempt would result in harm has to be lower than that of a similar but complete attempt that does cause harm.

To illustrate, suppose that two steps,  $A_1$  and  $A_2$ , need to be taken for an attempt to be complete. Assume too that there are two, equally numerous, groups of people who take  $A_1$ : type 1 individuals who will definitely

abandon their attempt before taking  $A_2$ , and type 2 individuals, who will definitely go on to take  $A_2$ , in which case there will be a .70 chance of harm. Suppose that courts cannot tell whether a person who has taken just  $A_1$  is a type 1 or a type 2. Then all the courts can infer is that there is a .50 chance that the person is a type 1 and a .50 chance the person is a type 2. Now if the person is a type 1, the true probability of harm is zero since a type 1 will not complete his attempt; if a person is a type 2, the true probability of harm is .70 since a type 2 will complete his attempt. Thus, the expected probability of harm is  $(.50)(0) + (.50)(.70)$ , or .35. If a person is brought to the court having completed an attempt, that is, after taking  $A_2$ , then the court will infer that he is certainly a type 2, so that the probability of harm for his act must have been .70.<sup>29</sup> Thus, the sanction for an incomplete attempt here ought to be less than that for a similar complete attempt that causes harm.

How small the probability is that an incomplete attempt would cause harm depends on two general factors: the stage at which an incomplete attempt is discovered, and whether an incomplete attempt is abandoned or is interrupted. The earlier an attempt is discovered, the greater the chance it would not have been completed, and thus the lower the probability it would have caused harm. If a person about whom there was some evidence of intent to commit murder by administering poison to his victim is found approaching a drugstore where the poison is sold, there may be doubt whether he would have purchased the poison (perhaps he was going to make innocent purchases) and doubt, too, whether he would have used the poison had he purchased it. If he is discovered after purchasing the poison, doubt about his carrying through his plans will be less, as one uncertainty has been resolved. But some uncertainty remains because his attempt is still incomplete.

The relevance of whether an attempt is abandoned or interrupted is plain. An abandoned attempt is presumably less likely to be completed and cause harm than an interrupted attempt, for an abandoned attempt might not be resumed. The person who turns away from the drugstore, or who buys the poison but then throws it out, must begin his act afresh and thus should be less likely to complete it than the person who is interrupted in the course of his attempt. Also, abandonment may indicate lack of true intention to do harm. It seems clear that many abandoned attempts and attempts interrupted early result in harm with a lower probability than

<sup>29</sup> An interesting contrasting example is this: there is only one type of individual—so that there can be no issue of uncertainty over his type—and there is a probability of .50 that he will abandon his act after  $A_1$ . Then the probability that he will do harm is .35, and this probability is what should influence the sanction whether the single type of person is brought before the court having taken  $A_1$  or  $A_1$  and  $A_2$ .

most complete attempts, making lower sanctions appropriate for the former—if any sanctions are required for them.

As an aside, it may be worth observing why punishment for *evil thoughts* is arguably inappropriate from the perspective under discussion. An evil thought may be regarded as an incomplete attempt about the completion of which there is so much doubt that no sanction is called for. There is only a small likelihood that evil thoughts will be followed by evil deeds, so the expected harmfulness of the “act” of having an evil thought is very small. (If, contrary to human nature, evil thoughts were usually followed by evil deeds, then we might well want to punish those who had evil thoughts. Or if courts could divine which evil thinkers would really carry out evil acts, we might well want to punish for evil thoughts alone.) When, however, a person with an evil thought commits an overt act, he distinguishes himself as someone who is more likely to complete the steps necessary to do harm.

From what has been said above, it is evident also that the situation for complete attempts is quite different from the situation for incomplete attempts. When a complete attempt is made, there are by definition no untaken steps necessary to the doing of harm. Thus, the only reason that the probability of harm may be less than that for an act that does do harm is that the unsuccessful attempt is different in character from the successful attempt. For example, if act  $A_1$  is that I give a low dose of poison to my intended victim and act  $A_2$  is that I give a high dose to him, then the two acts differ in character, and act  $A_1$  has a lower chance of killing the victim than does act  $A_2$ . Will the occurrence of harm cast light on the character of the completed attempt? If the intended victim dies of poisoning, then we may infer that the dose was more likely high than low.<sup>30</sup> Notice, however, that the occurrence of harm will not allow a useful inference to be drawn if we possess sufficient direct evidence about the quality of the act. We might, for instance, know the dosage of poison from forensic tests (suppose that fluids are drawn from the victim’s stomach); in this case, the victim’s death, or failure to die, would tell us nothing about the act, as the dosage is independently established.<sup>31</sup> In practice, such direct evi-

<sup>30</sup> Suppose that, if  $A_1$  is the act, the likelihood of death is .25, and that, if  $A_2$  is the act, the likelihood of death is .75. Suppose that our prior belief is that each act,  $A_1$  or  $A_2$ , is as likely to be committed as the other. If the intended victim dies, however, the posterior probability that  $A_2$  was the act is higher than .50; by Bayes’s rule, the posterior probability is  $(.50)(.75)/[(.50)(.75) + (.50)(.25)] = .375/.50 = .75$ .

<sup>31</sup> In formal terms, if  $A$  is an act and the court possesses complete knowledge of the act, the court will know whether the act will cause harm under every conceivable state of nature  $S_i$  and the probabilities of the  $S_i$ . Thus, the court will know the expected harm and be able to determine the optimal sanction. Whether harm actually occurs, and what the actual state of nature  $S_i$  turns out to be, will thus not affect the sanction the court should impose.

dence will often be lacking, but, even so, the occurrence or nonoccurrence of harm may not allow us to make a very useful inference about the character of the completed attempt. If the intended victim is saved because he spills his drink or because, after ingesting it, he is quickly given an antidote, then we would not be able to learn much about the dosage from the particular failure of harm to occur. For most complete, unsuccessful attempts that come to mind, the occurrence of harm would not lead to an inference that the probability of harm was substantially lower than the probability of harm for an otherwise identical-appearing act that did do harm. Thus, a substantially lower sanction would not be warranted for the attempt. In any event, the reasons for imposing lesser sanctions when a complete attempt fails are weaker than the reasons for imposing lesser sanctions for incomplete attempts.

2. *Incentives to Alter Conduct during Commission of Acts.* A factor that was not taken into account in the analysis is that sanctions for attempts may have an influence on individuals' behavior during the commission of acts.

It is often said that, if the sanction for an attempt is lower than the sanction for harm, individuals will have a reason to abandon an attempt before it is complete,<sup>32</sup> and they will have an incentive not to repeat a complete but failed attempt. However, this argument for setting the sanction for attempt below the sanction for harm is mistaken or at least not clearly stated. The argument does not recognize that different sanctions may be imposed for different types of attempt. For example, suppose that the sanction for a single, complete failed attempt and the sanction for doing harm on a first attempt are equal—say, to 10 years of imprisonment. But suppose that the sanction for a second attempt after a failed first attempt is different,<sup>33</sup> and higher, equal to 15 years of imprisonment (whether or not harm results on a second attempt). Then a person will have an incentive—namely, avoiding an additional 5 years of imprisonment—not to repeat his attempt if he at first fails, and he will have this incentive even though the sanction for a first attempt equals that for doing harm. Similarly, suppose that the sanction for an abandoned attempt is, say, 5 years of imprisonment and is thus lower than the 10-year sanction for a complete attempt or for doing harm. Then a person who has begun

<sup>32</sup> This argument is attributable to Cesare Bonesaria Beccaria, *An Essay on Crimes and Punishments* ch. 37 (1st ed. 1767), which reads at 153, "as there may be an interval of time between the attempt and the execution, it is proper to reserve the greater punishment for the actual commission, that even after the attempt there may be a motive for desisting."

<sup>33</sup> Notice that this requires that the court be able to distinguish between a single attempt and a repeated attempt.

an attempt will have an incentive to abandon it even though the sanction for a complete attempt equals that for doing harm.

The conclusion, therefore, is that the argument for making the sanction for attempts lower than that for doing harm, in order to induce early abandonment or abandonment after a failed attempt, reduces to an argument for a particular relationship among sanctions for different types of attempt: the sanction for an abandoned attempt should be lower than that for a complete attempt, and the sanction for repeated attempts should rise with the number of repetitions. There is no need, at any stage, for the sanction for doing harm to exceed that for an attempt.

It should be added that, if sanctions for abandoned attempts are lower than sanctions for complete attempts, or if sanctions rise with the number of repetitions, individuals will have an incentive to feign abandonment. Problems of strategic abandonment will emerge.

Sanctions for attempts may also affect the *effort* that individuals devote toward success. For example, an individual may try more or less hard to aim his gun accurately. If the sanction for causing harm is higher than the sanction for an attempt, the individual will have an incentive to aim the gun less carefully. It might be objected that, although this point is in strict logic valid, a person who has decided to do something like shoot at another person would be unlikely *then* to be influenced to fail in the attempt by the structure of sanctions. While there is real force to this objection, two answers can be given to it. First, sometimes the goal of a person may be mixed. A person who shoots may, in part, only be trying to put a scare into another. If so, he might aim wide of his target if the sanction for an attempt is lower than the sanction for success. Second, a similar objection could be made with regard to abandonment—that a person who sets out on a course of conduct would be unlikely to abandon it. Yet this objection is not emphasized in the abandonment context. The reason, I conjecture, is that many people do believe that the motives of those who engage in potentially harmful conduct are complex and subject to change.

*D. Attempts and Dangerous Acts That Do Not Do Harm:  
Criminal Law versus Civil Law*

It is interesting to compare the treatment of attempts under criminal law with the treatment of the analogue of attempts—dangerous acts that do not do harm—under civil law. Acts of the latter type—for instance, failing to place a cover over a manhole—do not usually create tort liability. Harm is normally a prerequisite. On the other hand, safety regulation may result in imposition of civil sanctions for certain dangerous practices whether or not they result in harm. Let us set aside for the moment any

sanctions for violating safety regulations and focus on the critical difference between tort law and criminal law: in the former, but not the latter, there must ordinarily be harm before sanctions are imposed.

The difference makes basic sense from the standpoint of the theory of deterrence. For acts in the core area of criminal law, the likelihood of apprehending those who do harm is low, often substantially less than one half.<sup>34</sup> If punishment were restricted to those who actually did harm, deterrence would be inadequate. Society really does need to take advantage of the opportunity to punish those who commit attempts. For acts in the core area of unintentional torts, the identity of those who do harm is far easier to establish. The likelihood of a tort action (or the threat of it and payment in settlement) is generally significantly greater than criminal prosecution. Consequently, the level of deterrence in the tort context should often be tolerably good even though liability is restricted to those occasions when harm actually occurs. It does not seem that there is a need in the area of tort comparable to that in the area of criminal law to increase deterrence by punishing acts that do not do harm.

Safety regulation provides further confirmation of the deterrence theory, precisely because this regulation seems to be employed where tort liability probably would not deter adequately.<sup>35</sup> It is used to overcome judgment-proof problems that dilute the threat of liability (consider harms due to fires or explosions). Similarly, regulation is often used when difficulty in establishing causation frustrates tort recovery or when the small size of separate claims precludes private damage actions (consider pollution and many environmental and health-related harms). Hence, safety regulation appears to be employed where it is needed to buttress deterrence and, thus, for the same reason that it is desirable to punish attempts under criminal law.

#### IV. COMMENT ON ATTEMPT AND OBJECTIVES OTHER THAN DETERRENCE

In this article, I have been concerned with the goal of deterrence. Of course, other social objectives also bear on the punishment of attempts, and, in closing, I would like to comment briefly on several of them: incapacitation, rehabilitation, and adherence to certain notions of fairness.

<sup>34</sup> For example, according to the F.B.I. Uniform Crime Reports 152 (1982), the percentages of reported crimes resulting in arrest were as follows: for murder, 72 percent; for rape, 48 percent; for robbery, 24 percent; for burglary, 14 percent; for motor vehicle theft, 14 percent. Because many arrests do not result in conviction, and because many crimes are unreported, these percentages must be taken as upper bounds on the likelihood of punishment.

<sup>35</sup> I develop this theme in Steven Shavell, *Liability for Harm versus Regulation of Safety*, 13 J. Legal Stud. 357 (1984).

Under the goal of incapacitation, society seeks to remove a person from the population if the future harm he would cause surpasses some threshold (which, for simplicity, may be taken to be the cost of incarceration).<sup>36</sup> Incapacitation thus offers an obvious justification for punishing attempts: committing an attempt will often be evidence of a person's future propensity to do harm. The theory of incapacitation does not, however, imply that the sanction for an attempt should be less than that for the doing of harm, even if the future dangerousness of a person who commits only an attempt is less than the future dangerousness of a person who does harm. This conclusion holds under the theory because, peculiarly, the optimal magnitude of sanctions does not rise in a continuous fashion with the future dangerousness of a person. Rather, if the future dangerousness of a person exceeds the threshold at which imprisonment becomes advantageous, then the person should be imprisoned until the time at which he would, if freed, no longer constitute a threat to society. This point of time is not connected in a necessary way to the dangerousness of the person's original act. Hence, although incapacitation supplies, as does deterrence, a reason for punishing attempts, it does not offer clear reason for punishing attempts less severely than acts that result in harm.

The goal of rehabilitation, if achievable during incarceration, also justifies sanctions for attempts because commission of an attempt may signal future dangerousness warranting rehabilitation. The need for rehabilitation, unlike the need for incapacitation, is presumably continuously related to the dangerousness of an individual. Hence, under the goal of rehabilitation, it seems that a person should be subjected to a less severe sanction for an attempt if it furnishes less evidence of dangerousness than does an act that results in harm.

A conception of fairness often invoked in relation to attempts is that punishment should be determined entirely by the quality of a person's act. The consequences of the act, being a product of fortuitous factors, should not matter. Accordingly, an attempt ought to be punished if it were an act with the qualities making punishment appropriate—if it were a sufficiently "bad" act. Further, an attempt ought to be punished in the same way as an identical act that turned out to cause harm. If, however, the court lacks information about the nature of the attempt, perhaps because it was incomplete, then the attempt may be punished less severely than an act that does harm because the court does not know how deserving of punishment the act is. Thus, the prescriptions for punishment resemble, in broad form, those following from deterrence theory. Another

<sup>36</sup> An economic analysis of incapacitation is contained in Steven Shavell, *A Model of Optimal Incapacitation*, 77 *Am. Econ. Rev.* 107 (1987).



notion of fairness is associated with satisfying demands for retribution, which are typically strongest when harm has been done. This notion of fairness argues against punishing attempts as severely as acts that do harm.

It appears, then, that the various goals of criminal law lead to recommendations for the punishment of attempts that are generally, although only roughly, congruent with one another.

## APPENDIX

### FORMAL ANALYSIS OF MODEL OF ATTEMPTS

Risk-neutral individuals choose whether to commit acts that may result in harm with a probability. If an individual commits such an act, he will obtain a benefit, and he will suffer a sanction with a probability that depends on whether he causes harm or does not cause harm, that is, whether he commits only an attempt. These probabilities of the imposition of sanctions are a function of enforcement effort. Sanctions are assumed to be monetary, but the case of nonmonetary sanctions is sketched.

Specifically, I define the following notation:

- $b$  = benefit from committing an act;<sup>37</sup>  $b \geq 0$ ;
- $f(b)$  = probability density of  $b$ ;  $f$  is continuous and positive on  $[0, \infty)$ ;
- $q$  = probability that an act will result in harm;
- $g(q)$  = probability density of  $q$ ;  $g$  is continuous and positive on  $[0, 1]$ ;
- $h$  = magnitude of the single level of harm in which an act may result;  $h \geq 0$ ;
- $r(h)$  = probability density of  $h$ ;  $r$  is continuous and positive on  $[0, \infty)$ ;
- $c$  = enforcement effort;  $c \geq 0$ ;
- $p_1(c)$  = probability that an individual is penalized, conditional on his not causing harm—committing only an attempt;  $p_1(0) = 0$ ;  $p_1'(c) > 0$ ;
- $s_1$  = sanction for an attempt;  $s_1 \geq 0$ ;
- $p_2(c)$  = probability that an individual is penalized, conditional on his committing an act and causing harm;  $p_2(0) = 0$ ;  $p_2'(c) > 0$ ;
- $s_2$  = sanction for committing an act and causing harm  $h$ ;  $s_2 \geq 0$ ; and
- $w$  = an individual's wealth;  $w > 0$ .

Each individual is faced with the choice of committing or not committing one type of potentially harmful act; thus, an individual is identified by a benefit  $b$ , a probability of doing harm  $q$ , and a potential harm  $h$ . The distributions of  $b$ ,  $q$ , and  $h$  over the population of individuals are assumed for simplicity to be independent.

Note, the single level of enforcement effort,  $c$ , determines both  $p_1$  and  $p_2$ . The justification is that much activity devoted to apprehension of individuals who do harm will also enable apprehension of those who commit attempts (conversely as well); the importance of this assumption is discussed in remark 2 following proposition 1. In addition, the same probabilities  $p_i$  are assumed to apply with regard to commission of acts with different  $h$ . The justification for the assumption is similar

<sup>37</sup> The benefit could be interpreted as an expected benefit.

to that just mentioned—that much activity devoted to enforcement will result in detection of a whole range of violations; the significance of this assumption is discussed in remark 3 following proposition 1. For ease,  $p_i$  will often be written without argument  $c$ .

Because a sanction cannot be higher than an individual's wealth,

$$s_i \leq w. \quad (\text{A1})$$

The sanctions are functions of what courts are able to observe. We now consider two cases, where courts are able to observe  $g$  and  $h$  and where they are unable to observe  $q$ .

### I. COURTS' INFORMATION IS PERFECT

Suppose here that courts are able to observe  $q$  and  $h$  both for attempts and for acts that result in harm; hence,  $s_i = s_i(q, h)$ . Individuals are assumed to know  $q$  and  $h$ . The expected sanction faced by an individual who commits an act of type  $(q, h)$  is

$$z(q, h) = (1 - q)p_1s_1(q, h) + qp_2s_2(q, h). \quad (\text{A2})$$

An individual will commit an act if and only if<sup>38</sup>

$$b \geq z(q, h). \quad (\text{A3})$$

Note that there is a maximal expected sanction. Let  $\bar{z}$  be the maximal expected sanction; thus,

$$\bar{z} = (1 - q)p_1w + qp_2w. \quad (\text{A4})$$

Social welfare is the sum of benefits individuals obtain from committing acts less the harm done and enforcement effort,

$$\int_0^\infty \int_0^1 \int_{z(q,h)}^\infty (b - qh)f(b)dbg(q)dq r(h)dh - c. \quad (\text{A5})$$

The social problem is to choose schedules of sanctions  $s_i(q, h)$  and enforcement effort  $c$  to maximize (A5), where  $z(q, h)$  is determined by (A2). Denote by  $*$  the optimal levels of the variables, and assume that  $c^*$  and, thus,  $p_i^* = p_i(c^*)$  are positive (otherwise the social problem is without interest). Then we have proposition 1.

**PROPOSITION 1.** Assume that courts can obtain perfect information about the probability  $q$  and the potential harm  $h$  both for attempts and for acts that cause harm. Then the optimal system of sanctions has the following properties.

*a)* Attempts are penalized when the expected harm  $qh$  exceeds a threshold (equal to  $qp_2^*w$ ). If the expected harm exceeds a further threshold (equal to  $(1 - q)p_1^*w + qp_2^*w$ ), both attempts and acts resulting in harm are penalized, and with the maximal sanction.

*b)* The sanctions for attempts and for acts resulting in harm are not uniquely determined, except in the region where both are maximal; only the expected sanction is uniquely determined.

*c)* The expected sanction equals the expected harm when this is feasible.

**REMARKS.** 1. Figure 2 illustrates proposition 1. The explanation for the propo-

<sup>38</sup> I adopt the convention that, if  $b = z(q, h)$ , an individual will commit an act.

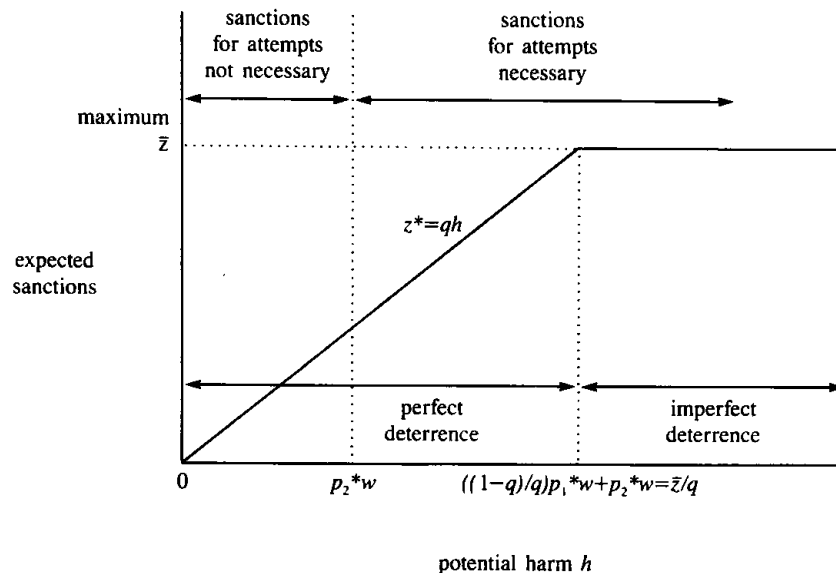


FIGURE 2.—Optimal sanctions when courts possess perfect information

sition is essentially as given in Section II: it is optimal for the expected sanction to equal the expected harm when possible. This implies that attempts must be penalized beyond some point; otherwise the expected sanction would be lower than the expected harm, and so forth.

2. The result that it is optimal to punish attempts for sufficiently high  $q$  and  $h$  would continue to hold under a wide range of assumptions about the determination of  $p_1$  and  $p_2$ . Clearly, under any assumption such that  $p_1$  is positive when  $p_2$  is positive, it will be optimal to punish attempts when  $q$  and  $h$  are sufficiently high, for then imposing sanctions only when harm is done would not result in a sanction of  $qh$ , and penalizing attempts is possible since  $p_1$  is positive. Thus, under the following assumptions, it would be optimal to punish attempts for high enough  $q$  and  $h$ : (a) For any  $c$ , the  $p_i$  are functions of  $h$  (rather than being the same for all  $h$ ). (b) The  $p_i$  are independent functions of two types of enforcement effort:  $p_1 = p_1(c_1)$ ,  $p_2 = p_2(c_2)$ , where  $p_1(0) > 0$ .<sup>39</sup> (Alternatively, even if  $p_1(0) = 0$ , it would still be optimal to punish attempts if  $p_1'(0)$  is sufficiently high.)

3. If, contrary to the assumptions of the model, the level of enforcement effort may be independently set for acts of each level of potential harm—so that the  $p_i = p_i(c_h)$ , where  $c_h$  is enforcement effort for acts that may do harm of  $h$ —then the optimal sanctions for attempts and for causing harm are always maximal. The reason is that, if either  $s_1(h)$  or  $s_2(h)$  were less than maximal, that sanction could be raised and  $c_h$  lowered without altering the expected sanction; hence, social welfare could be increased. The assumption of the present model, that enforcement effort  $c$  is of a general nature, affecting the  $p_i$  for all  $h$ , results in different conclusions because the argument just noted cannot be applied independently for each  $h$ .<sup>40</sup>

<sup>39</sup> It is plausible that the  $p_i$  is positive when enforcement effort is zero since some attempts and some instances of harm will be observed by private citizens and reported.

<sup>40</sup> See Steven Shavell, *Specific versus General Enforcement of Law* (Working Paper No. 3062, National Bureau of Economic Research, August 1989).

*Proof.* The argument is made in two steps.

i)  $z^*(q, h) = qh$  when this is possible, that is, when  $qh \leq \bar{z} = (1 - q)p_1^*w + qp_2^*w$ —or equivalently, when  $h \leq \bar{z}/q = (1 - q)p_1^*w/q + p_2^*w$ ; otherwise  $z^*(q, h) = \bar{z}$ . Given the  $p_i^*$ , the social problem for any  $q$  and  $h$  is to maximize over the  $s_i(q, h)$ , and thus over  $z(q, h)$ , the quantity

$$\int_{z(q, h)}^{\infty} (b - qh)f(b)db. \quad (\text{A6})$$

It is convenient to treat  $z$  as the control variable. The derivative of (A6) with respect to  $z$  is

$$(qh - z)f(z), \quad (\text{A7})$$

which is positive when  $z < qh$ , zero at  $z = qh$ , and negative for larger  $z$ . Hence, (A6) is maximized at  $z = qh$ , so this must hold if possible, that is, if  $qh \leq \bar{z}$ ; otherwise, (A6) is maximized at  $z = \bar{z}$ .

ii) For  $h$  in  $[0, p_2^*w]$ ,  $s_1^*(q, h) = 0$  is possible (as is  $s_1^*(q, h) > 0$ ;  $s_1^*(q, h)$  is not uniquely determined). For  $h$  in  $(p_2^*w, \bar{z}/q)$ ,  $s_1^*(q, h) > 0$  (yet  $s_1^*(q, h)$  is not uniquely determined). For  $h$  exceeding  $\bar{z}/q$ ,  $s_1^*(q, h) = w$ : from step i we know that, for  $h$  in  $[0, p_2^*w]$ ,  $z^*(q, h) = qh$ , but this can be achieved with  $s_1 = 0$ , since  $qh < qp_2^*w$  in the interval. Similarly, we know that, for  $h$  in  $(p_2^*w, \bar{z}/q)$ ,  $z^*(q, h) = qh$ . But if  $s_1^*(q, h) = 0$ , then  $z^*(q, h) = qp_2^*s_2 < qh$  in the interval. Hence,  $s_1^*(q, h)$  must be positive in the interval. Finally, for larger  $h$ , we know from remark i that  $z^*(q, h) = \bar{z}$  and, hence, that  $s_1^*(q, h) = s_2^*(q, h) = w$ . Q.E.D.

*Optimal probabilities.* Using the proposition, the optimal choice of enforcement effort may be described. In view of the proposition, social welfare (A5) may be expressed as

$$\begin{aligned} & \int_0^1 \int_0^{\bar{z}/q} \int_{qh}^{\infty} (b - qh)f(b)db r(h)dh g(q)dq \\ & + \int_0^1 \int_{\bar{z}/q}^{\infty} \int_{\bar{z}}^{\infty} (b - qh)f(b)db r(h)dh g(q)dq - c, \end{aligned} \quad (\text{A8})$$

where, note,  $\bar{z} = (1 - q)p_1w + qp_2w$  depends on the  $p_i$  and, thus, on  $c$ . The first term in (A8) is associated with the region of  $h$  over which deterrence is perfect, and the second term is associated with the region of  $h$  over which deterrence is insufficient. Setting the derivative of (A8) with respect to  $c$  equal to zero, we obtain the first-order condition determining  $c^*$ ,

$$\int_0^1 \bar{z}'(c)f(\bar{z}(c)) \int_{\bar{z}/q}^{\infty} (qh - \bar{z})r(h)dh g(q)dq = 1, \quad (\text{A9})$$

which says that the marginal benefits of extra deterrence over the region where there is underdeterrence equal the marginal cost of raising enforcement effort, that is, unity. (Because the sanction is maximal in the region with underdeterrence, deterrence can be enhanced only by raising the likelihood of sanctions.)<sup>41</sup>

*Nonmonetary sanctions.* In the case where sanctions are nonmonetary,  $w$  will be interpreted as the maximum possible nonmonetary sanction, so that equation

<sup>41</sup> The first-order condition determining  $c^*$  in the case below, when  $q$  cannot be observed for attempts, is similar and will not be stated.

(A1) still applies. Also, it will be assumed that there is a social cost incurred when sanctions are imposed. Let  $\sigma$  = social cost of imposing a unit nonmonetary sanction, where  $\sigma > 0$ , so that  $\sigma s$  is the cost of imposing a sanction of magnitude  $s$ . Social welfare is then

$$\int_0^\infty \int_0^1 \int_{z(q,h)}^\infty (b - qh - \sigma z) f(b) db g(q) dq r(h) dh - c. \quad (\text{A5}')$$

The first two parts of the proposition can again be shown to hold using two steps analogous to the ones in the above proof: (i)  $z^*(q, h) = 0$  for all  $h$  sufficiently small;  $z^*(q, h)$  is positive for some  $h$ ; if  $z^*(q, h)$  is positive and less than  $\bar{z}$ , then  $z^*(q, h)$  is determined by the first-order condition that (A11) below is zero, and  $z^*(q, h)$  increases with  $h$ ;  $z^*(q, h) = \bar{z}$  for all  $h$  sufficiently high: given the  $p_i^*$ , the social problem for any  $q$  and  $h$  is to maximize over  $z(q, h)$  the quantity

$$\int_{z(q,h)}^\infty (b - qh - \sigma z) f(b) db, \quad (\text{A10})$$

the derivative with respect to  $z$  of which is

$$(qh + \sigma z - z)f(z) - \sigma[1 - F(z)], \quad (\text{A11})$$

where  $F$  is the cumulative distribution function of  $f$ . When  $z = 0$ , (A11) equals  $qh f(0) - \sigma$ . Thus, for all  $h$  sufficiently small, (A11) is negative, and  $z = 0$  is the optimal  $z$ .<sup>42</sup> Also, since  $qh f(0) - \sigma$  is positive for  $h$  sufficiently large, a positive  $z$  is optimal for all  $h$  sufficiently large. If  $z$  is an interior optimum, then it is determined by setting (A11) equal to zero. Since the partial derivative of (A11) with respect to  $h$  is  $qf(z) > 0$ ,  $z^{*'}(h) > 0$ .<sup>43</sup> It may be shown that  $z^*(q, h) = \bar{z}$  for all  $h$  sufficiently high. A sketch of the argument is that, since  $z^*(q, h)$  must rise with  $h$  once it becomes positive, it must tend to a limit, and this limit must be  $\bar{z}$ . If  $z^*(q, h)$  does not equal  $\bar{z}$  for all  $h$  beyond some point, the first-order condition that (A11) is zero must hold, but this cannot be.<sup>44</sup> (ii)  $s_1^*(q, h) = 0$  for all  $h$  sufficiently small; then over an interval of  $h$ ,  $s_1^*(q, h) > 0$  (but  $s_1^*(q, h)$  is not uniquely determined);  $s_1^*(q, h) = w$  for all larger  $h$ : By (i), we know that  $z^*(q, h)$  is at first zero and then rises monotonically to  $\bar{z}$ . Hence, we know that  $s_1^*(q, h) = 0$  for all  $h$  sufficiently low and that subsequently  $s_1^*(q, h)$  must be positive but is not uniquely determined. In the region of  $h$  where  $z^*(q, h) = \bar{z}$ ,  $s_1^*(q, h)$  and  $s_2^*(q, h)$  both equal  $w$ .

## II. COURTS' INFORMATION IS IMPERFECT

Assume now that courts cannot tell what the likelihood of harm  $q$  was when an attempt was committed or when harm was done, but that courts are able to determine the magnitude of the potential harm  $h$ . Thus, the assumption is that  $s_i =$

<sup>42</sup> What has just been shown is that  $z = 0$  is a local maximum of (A10);  $z = 0$  can also be shown to be a global maximum using essentially the argument given in Shavell, *supra* note 40.

<sup>43</sup> The first-order condition is of the form  $W(z, h) = 0$ . Implicitly differentiating with respect to  $h$ , we obtain  $W_z z' + W_h = 0$ , so that  $z' = -W_h/W_z$ . But  $W_z < 0$ , for this is the second-order condition for  $z$  to maximize  $W$ . Hence, the sign of  $z'$  equals that of  $W_h$ .

<sup>44</sup> The details may be obtained from a closely related argument given in the proof of proposition 2b in Shavell, *supra* note 40.

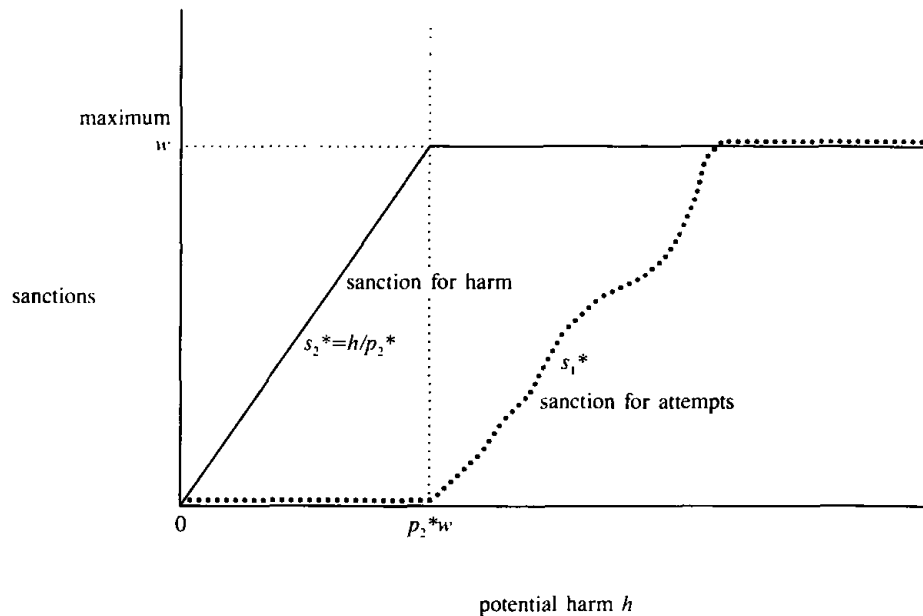


FIGURE 3.—Optimal sanctions when courts lack information about the probability of harm

$s_i(h)$ . Hence, the expected sanction is given by

$$z(h) = (1 - q)p_1s_1(h) + qp_2s_2(h), \quad (\text{A2}')$$

and social welfare is given by the expression (A5), with  $z(h)$  replacing  $z(q, h)$ . I will demonstrate proposition 2.

**PROPOSITION 2.** Assume that courts cannot determine the probability of harm  $q$  but do know the magnitude of the potential harm  $h$ . Then the optimal system of sanctions has the following properties.

a) The sanction for attempts is zero as long as the potential harm is below a threshold (equal to  $p_2^*w$ ). At this threshold, the sanction for causing harm first becomes maximal. Beyond the threshold, the sanction for attempts rises with the magnitude of potential harm, until it equals the maximal level.

b) The sanction for causing harm rises with the magnitude of potential harm (it equals  $h/p_2^*$ ) until it attains the maximal level, and then it remains at this level.

c) Sanctions for attempts and for causing harm are uniquely determined.

d) The expected sanction equals the expected harm in the region where attempts are not punished. In the region where attempts are punished, the expected sanction exceeds the expected harm when the probability of harm is low and is less than the expected harm when the probability of harm is high.

**REMARKS.** 1. The proposition is illustrated in Figure 3 and is explained as in the text. It is best for the expected sanction for an act of type  $(q, h)$  to be  $qh$ , but the court cannot observe  $q$ . Imposing a sanction for harm,  $s_2$ , helps to accomplish making the expected sanction equal to  $qh$  since the higher  $q$  is, the higher the likelihood is that harm will occur and that  $s_2$  will be suffered. This is why  $s_2$  alone is positive over a range of  $h$ . However, when  $h$  is sufficiently large, making the sanction for  $s_2$  positive is not enough to create adequate deterrence, so it becomes optimal to make  $s_1$  positive as well, even though this means that acts with low  $q$  will face an expected sanction greater than  $qh$ .

2. In a more general model than the one considered, courts may obtain some

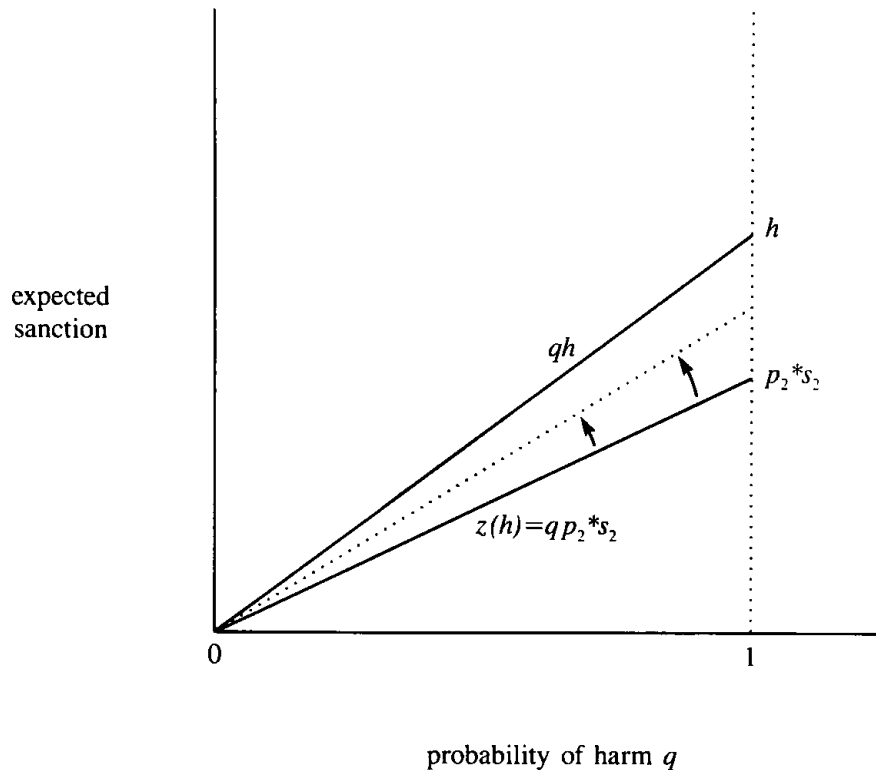


FIGURE 4.—Expected sanctions and the probability of harm—case 1

information about  $q$ . In such a model, one supposes that courts would use that information and would raise the sanction the higher  $q$  appeared to be, so that the expected sanction could more nearly approximate  $qh$ .

*Proof.* The argument is made in several steps.

i)  $s_1^*(h) = 0$  and  $s_2^*(h) = h/p_2^*$  for  $h \leq p_2^*w$ : for any  $h$ , the problem is to choose the  $s_i(h)$  to maximize

$$\int_0^1 \int_{z(h)}^{\infty} (b - qh)f(b)dbg(q)dq. \quad (\text{A12})$$

It is clear that the first-best outcome would be achieved if, for every  $q$ ,  $z(h) = qh$ . Now, by (A2'),  $z(h) = (1 - q)p_1^*s_1(h) + qp_2^*s_2(h)$ . Hence, if  $s_1(h) = 0$  and  $p_2^*s_2(h) = h$ , then  $z(h) = qh$  for every  $q$ . For  $p_2^*s_2(h)$  to equal  $h$ , we must have  $s_2(h) = h/p_2^*$ , which is feasible when and only when  $h/p_2^* \leq w$  or, equivalently, when and only when  $h \leq p_2^*w$ . Also, note that the only way that, in this region,  $z(h) = qh$  for all  $q$  is that  $s_i$  be as claimed. For  $(1 - q)p_1^*s_1(h) + qp_2^*s_2(h) = qh$  implies that  $p_1^*s_1(h) = [q/(1 - q)][h - p_2^*s_2(h)]$ . This can hold identically in  $q$  only if both  $s_1$  and  $h - p_2^*s_2$  are zero.

ii)  $s_2^*(h) = w$  for  $h > p_2^*w$ : consider Figures 4 and 5, which show the expected sanction  $z(h)$  as a function of  $q$  in two cases. In the first,  $s_1 = 0$ , so that the graph of  $z(h)$  is as in Figure 4. Note that  $p_2^*s_2$ , the intercept of  $z(h)$  with the line  $q = 1$ , must lie below  $h$  since  $p_2^*w < h$ . It is clearly social welfare enhancing for the graph to rotate upward about the origin since, then, for every  $q$  there will be less underdeterrence. Hence, the intercept with the line  $q = 1$  should be as high as possible, namely, equal to  $p_2^*s_2$ . But this means that  $s_2 = w$ .

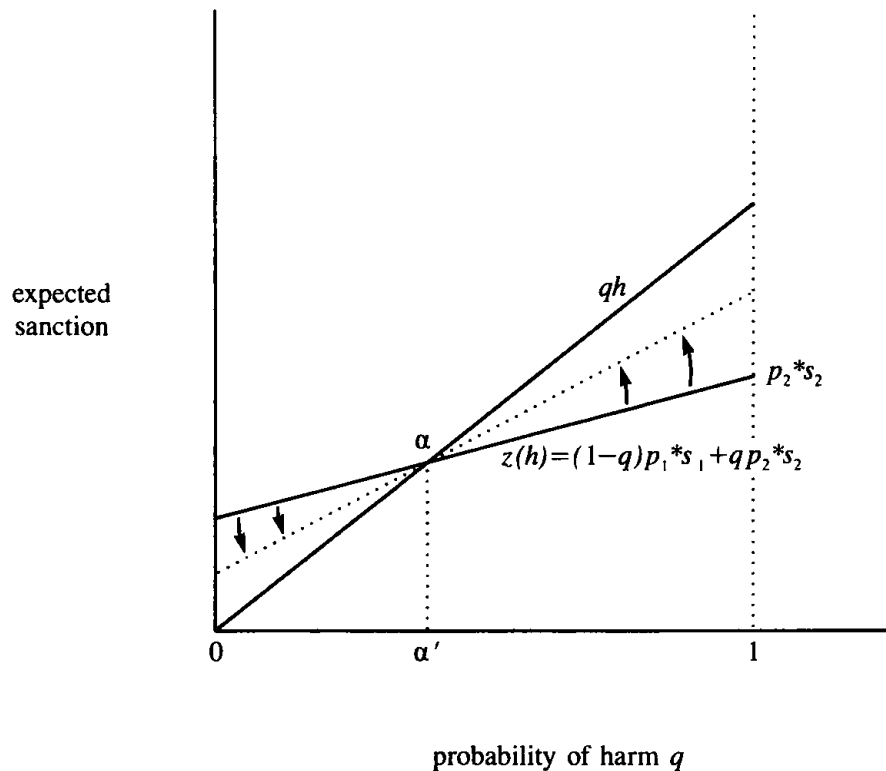


FIGURE 5.—Expected sanctions and the probability of harm—case 2

In the other case,  $s_1 > 0$ , and the graph of  $z(h)$  is as shown in Figure 5. Here it is desirable for the graph to rotate upward about the point  $\alpha$  of intersection with the line  $qh$ , for then there is less overdeterrence to the left of  $\alpha'$  and less underdeterrence to the right of  $\alpha'$ . But this corresponds to increasing  $s_2$  and should be done as long as possible, that is, until  $s_2 = w$ .

iii)  $s_1(h) > 0$  for  $h > p_2^*w$ , and  $s_1(h)$  increases with  $h$  in this region until it attains the maximal level  $w$ : since in the region of  $h$  under consideration,  $s_2$  is  $w$ ,  $s_1$  must maximize (A12), where  $z(h) = (1 - q)p_1^*s_1 + qp_2^*w$ . The derivative of (A12) with respect to  $s_1$  is

$$\int_0^1 (1 - q)p_1^* \{qh - [(1 - q)p_1^*s_1 + qp_2^*w]\} f((1 - q)p_1^*s_1 + qp_2^*w) g(q) dq. \quad (\text{A13})$$

Now (A13) is positive at  $s_1 = 0$ , for then the integrand equals  $(1 - q)p_1^*q(h - p_2^*w)f(qp_2^*w)g(q) > 0$  since  $h > p_2^*w$ . Hence,  $s_1^*$  is positive. Moreover, when  $s_1^*$  is an interior solution, determined by the condition that (A13) is zero, the sign of  $s_1^{*'}(h)$  is equal to the sign of the partial derivative of (A13) with respect to  $h$ , which is the integral of  $(1 - q)p_1^*qf((1 - q)p_1^*s_1 + qp_2^*w)g(q)$ , which is positive. Hence, as claimed,  $s_1^*(h)$  rises with  $h$  when it is less than  $w$ . Q.E.D.

*Nonmonetary sanctions.* If sanctions are nonmonetary, then one suspects that the optimal sanction for an attempt will often be less than that for doing harm, by essentially the reasoning given in the above proof. Namely, this will create an expected sanction that rises with  $q$ . However, I have been unable to find simple, sufficient conditions under which this can be proved.