1-1-1991

Impossibility, Subjective Probability and Punishment for Attempts

David D. Friedman
Santa Clara University School of Law, ddfr@daviddfriedman.com

Recommended Citation
20 J. Legal Stud. 179

IMPOSSIBILITY, SUBJECTIVE PROBABILITY, AND PUNISHMENT FOR ATTEMPTS

*DAVID D. FRIEDMAN**

STEVEN SHAVELL, in his recent article "Deterrence and the Punishment of Attempts,"[1] raises the question of whether people should be punished for attempted crimes that are impossible—that have zero probability of succeeding. Two examples are the person who attempts to commit murder by voodoo and the pickpocket who attempts to pick an empty pocket. His conclusion is that "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 punished because there is no need to deter them."

In his analysis, Shavell considers the objective probability that an act will succeed and the (subjective) probabilities held by the court. He does not consider probabilities from the standpoint of the offender. The purpose of this note is to argue that it is crucial, in analyzing the question of punishment for impossible attempts, to consider the subjective probabilities of the offender—his beliefs as to what methods of committing the crime work.

Obviously, if the offender knows that the method he is using will not work, he will not use it—or, if he does, he is not in fact attempting to commit the crime. The important case is the one where the court knows that the method is impossible but the offender does not. If the offender is aware of his own ignorance and rationally allows for it in his decision, then, as we will see, punishing impossible attempts does in fact deter offenses. If, however, the offender believes that he has perfect knowledge

* John M. Olin Visiting Fellow in Law and Economics, University of Chicago Law School.

[1] Steven Shavell, Deterrence and the Punishment of Attempts, 19 J. Legal Stud. 435 (1990).

[*Journal of Legal Studies*, vol. XX (January 1991)]

about what methods work but is wrong, then punishing impossible attempts serves no function. We will start with the former case.

RATIONAL VOODOO

Consider the following simple example. A criminal is considering committing a murder in one of two ways—poison or voodoo. The poison he is considering is invariably lethal, whereas sticking pins in a voodoo doll will have no effect at all on the prospective victim's life expectancy.

If he were aware of these facts, he would either choose poison or not attempt the murder at all. The problem arises because he is not aware of them. He knows that the methods he is contemplating may not work, but he does not know which is more likely to work. Assume, for simplicity, that he knows one of the methods works with p_s and the other has no effect but has no idea which is which.

Shavell's analysis is based on the deterrent effect of punishment, although he also discusses other functions, such as putting a criminal in jail in order to make future crimes impossible. My analysis will also be based on deterrence. To entirely eliminate the issue of incapacitation, I assume that the criminal will make at most one attempt; having spent his savings on either poison or a voodoo doll and failed, he will give up.

Like Shavell, I assume that, whichever method is used, there is some probability that the offender will be detected before the crime is complete. The question is whether, if he is detected, he should be punished only if he is using a method that might work. Shavell's answer is yes, mine is no.

The crucial point is that, from the criminal's standpoint, the legal rule is: *Attempts by impossible means are (or are not) punishable.* Since the criminal does not know which method is impossible (if he did, he would not bother using it), this does not translate, for him, into *attempts by voodoo are (or are not) punishable.* Since he knows that the method he is planning to use might turn out to be impossible, a policy of punishing attempts that (turn out to be) impossible raises the ex ante cost to him of trying to commit murder.

The argument can be made more formal as follows:

p_a = the probability that an attempt (of either sort) is detected,
p_c = the probability that an offense (in this case murder) is detected,
F_i = the punishment for an impossible attempt,
F_p = the punishment for a possible attempt,
F_c = the punishment for committing an offense, and
p_s = the probability that a possible attempt will succeed.

In the example given, the expected punishment (as perceived by the offender) for attempting to commit murder is

$$p_a(F_i + F_p)/2 + (1/2)p_s p_c F_c. \quad (1)$$

Here the first term is the probability that the attempt will be detected times the resulting expected punishment—a 50 percent chance that the method turns out to be impossible (punishment F_i) plus a 50 percent chance that it turns out to be possible (F_p). The second term is a 50 percent chance that the method chosen is correct, multiplied by the probability that the correct method will work, multiplied by the expected punishment for an offense. As you can see from equation (1), a punishment for an impossible attempt is, in this case at least, just as effective a deterrent as a punishment for a possible attempt. If, as Shavell assumes, punishment is costless but has an upper bound, we are indifferent among any combination of punishments for possible and impossible attempts having a given sum and will punish both kinds of offenses if the desired total is too high to be achieved by punishing only one. If, more realistically, marginal punishment cost increases with the size of punishment,[2] we produce a given level of deterrence at the lowest cost by imposing equal punishments for both the possible and the impossible attempt.

So far, I have assumed that the offender knows nothing at all about which method is impossible. Consider, at the opposite extreme, an offender who knows perfectly well which method works; if he chooses to use the method that does not, it is in order to achieve some objective other than committing the crime (perhaps he wants to impress a friend who believes in voodoo). In this case, punishment for the impossible attempt has no tendency to prevent the offense but does tend to prevent the (harmless) attempt, hence Shavell's conclusion is in this case correct.

Finally, consider an offender who is certain that the impossible method is the one that works. The knowledge that impossible attempts, if detected, will be punished has no deterrent effect on him since he is certain the method he is using is not the impossible one.

The argument can be expanded to cover the more general case where the offender has some, but imperfect, knowledge about the effectiveness of the alternative methods. This is done in the Appendix, where we analyze the case of a potential offender whose subjective probability that the method that is actually impossible (voodoo) is the possible one is p. The

[2] The simplest argument for this is that lower punishments are more likely to be payable as fines, which are simply transfers, rather than requiring costly punishments such as imprisonment or execution. This point is discussed at much greater length in David Friedman, Reflections on Optimal Punishment or Should the Rich Pay Higher Fines? 3 Res. L. & Econ. 185–205 (1981).

conclusion is that, if $p < .5$ (the offender considers the method that in fact is impossible less likely to work than the method that in fact is possible), then punishing impossible attempts deters offenses, although a given level of punishment of impossible attempts provides less deterrence than an equal punishment of possible attempts. If the offender's beliefs are perverse—he thinks voodoo more likely to work than poison—he will either not attempt to commit the offense or use the impossible method.

Irrational Voodoo and the Reasonable Man

Suppose we believe that there are two kinds of people. Reasonable people know which methods of committing crimes are impossible and do not employ them ($p = 0$). Unreasonable people sometimes employ impossible methods. Unreasonable people do not realize that they are unreasonable; they believe that they know which methods are impossible ($p = 1$). The knowledge that impossible attempts will be punished does not increase their (subjective) expected punishment and, hence, does not deter them. So there is no reason to punish impossible attempts. We are back with Shavell's result.

This argument may justify defining "impossibility" by a reasonable man standard: An attempt is impossible if a reasonable man would know that it would not work. An attempt that is in fact impossible but that a reasonable man might believe is possible (trying to pick a pocket that in fact is empty) should still be punished, as per our earlier analysis.

To put the argument differently, a method that a reasonable man would regard as impossible is one for which p has a sharply bipolar distribution. Almost everyone either is reasonable and is sure voodoo does not work ($p = 0$) or is unreasonable and thinks it probably does work ($p > .5$). Neither group will be deterred from murder by the knowledge that impossible attempts are punished, so there is no reason to punish them. A method that a reasonable man would regard as possible is one for which many potential offenders will have $0 < p < .5$, hence punishing attempts of that sort that turn out to be impossible provides useful deterrence.

The Many Person Case

Our argument so far involves deterring a single offense by a single offender. The next stage is to consider a population of potential offenders with varying values of p. Those for whom $p > .5$ will either not attempt the offense or pick the impossible method; they will produce no offenses but may produce some convictions for (impossible) attempts. Those for whom $p < .5$ will either not attempt the offense or choose the correct (possible) method. As can be seen from equation (A4) of the Appendix, the knowledge that impossible attempts, if detected, will be punished

increases the expected cost of the offense. The size of the effect is proportional to $p/(1 - p)$, hence increases with increasing p. So a policy of punishing impossible attempts selectively deters those for whom p is relatively large—potential offenders who are uncertain as to which method works.

The argument may be made verbally as well as mathematically. If you are unsure which method works, you must allow in your calculations for the chance you will choose the wrong one. If impossible attempts are not punished, then the wrong choice means that you will not succeed in your crime but will also not be punished. If impossible attempts are punished, you risk using an impossible method and being punished for your attempt. That possibility is one of the costs you must take into account in deciding whether or not to attempt the offense. The potential offender who is sure he knows which method works does not have to worry about that problem.

So a policy of punishing impossible attempts tends to deter offenses, especially by those offenders uncertain as to which method works. The cost of that deterrence is that some people caught making impossible attempts must be punished. These will, by the analysis of the Appendix, be people for whom $p > .5$—potential offenders who believe that voodoo is more effective than poison. The offenses that will be deterred will be by potential offenders who think poison is more effective than voodoo but are not sure. Whether punishing impossible attempts is a relatively efficient or inefficient way of deterring crime will depend on the relative sizes of the two groups.[3] Even if punishing impossible attempts has a relatively small deterrent effect (because most potential offenders have $p \ll .5$), it may be an efficient form of deterrence if there are very few impossible attempts that need be punished (because very few potential offenders have $p > .5$).[4]

Are All Failed Attempts Irrational?

One objection to both my analysis and Shavell's might be that the distinction we make between "possible" and "impossible" attempts is meaningless. All unsuccessful attempts are impossible ex post; they differ only in the degree to which the perpetrator knew that they were impossible. The only possible attempts are those that succeed.

[3] It will also depend on other characteristics of the two groups that affect how likely they are to commit the crime for a given level of expected punishment.

[4] If we made our model more complicated by allowing for the fact that different methods have different costs (making a voodoo doll might be cheaper than buying poison), the detailed quantitative conclusions would change. The division between those who chose the two methods would no longer be at $p = .5$; some people might use the method they thought less likely to work because it was cheaper. The logic of the argument, and the quantitative results, would remain the same.

It is useful, and reasonably straightforward, to reformulate my analysis in these terms. In the simplest case, we have two kinds of methods—those that work with certainty and those that fail. The question of whether to punish impossible attempts becomes the question of whether to punish unsuccessful attempts. The answer, as shown above, depends on the distribution of subjective probabilities in the population of potential offenders. The analysis is the same as before, with $p_s = 1$ and $F_p = 0$.

More generally, we have a large number of methods, some of which work with certainty and some of which fail. Among the failed methods, some are ones for which the perpetrator had a high subjective probability of success (putting cyanide in a glass of water and handing it to the victim—who drops the glass) and some are ones for which the perpetrator had a low subjective probability of success. Applying our analysis to this more general case, with a distribution of subjective probabilities over both methods and potential perpetrators, would tell us which impossible (that is, failed) attempts should be punished and which labeled "impossible" and excused.

IRRATIONALITY

So far I have assumed that potential offenders are rational and fully informed about everything affecting their crime except what method works—that even people who believe in voodoo know the legal doctrine with regard to impossible attempts and correctly allow for its effect on their expected punishment. If this is not true, the results of punishing impossible attempts are much less clear; they depend on precisely who is irrational or poorly informed, how much, and in what way.

Suppose, for instance, that anyone who thinks voodoo might work is so badly informed, or so irrational, that actual legal doctrines have no effect on him at all. Such people may make impossible attempts and be punished for them, but they will not be deterred by that possibility since they will not be aware of it. In that case, punishment of impossible attempts will have no deterrent effect.

How plausible we think this argument is depends partly on how obvious it is which methods are impossible; someone may be both rational and well informed about the law, yet still try to pick a pocket that happens to be empty. It also depends on how we believe that poorly informed and/or irrational people make decisions. If, for instance, irrationality takes the form of overestimating the probability of unlikely events, a claim sometimes used to explain the purchase of lottery tickets, then it may increase deterrence—if, as seems likely, the apprehension of someone using voodoo to try to commit murder is an unlikely event.

INCAPACITATION

Suppose we expand our analysis to allow for incapacitation as a purpose of the penal system. We might then argue that people who attempt to commit offenses in ways that cannot work are harmless; their knowledge of real-world causality is sufficiently bad that anything they attempt, including murder, is likely to fail. Hence there is no advantage to imprisoning then.

This argument can, however, be made the other way. One of the things we learn from a correct knowledge of causality (at least under an efficient legal system) is that killing people may well cause us to be punished. People who do not understand causality may fail to perceive that relation. If so, they will attempt more murders than would more rational people—and eventually they may give up on voodoo and use poison instead.

CONCLUSION

In the course of this note I have tried to establish two conclusions, one substantive and one methodological. The former is that punishing impossible attempts provides some deterrence against actual crimes and may even be an efficient way of deterring them. The latter is that, in analyzing issues of deterrence, we must take into account the beliefs of the people we are trying to deter, as embodied in their subjective probabilities, since it is on the basis of those beliefs that they will act.

MATHEMATICAL APPENDIX

We define:

method 1 = the method that in fact is impossible (voodoo),
method 2 = the method that in fact works with probability p_s (poison),
p = the offender's subjective probability that method 1 works (with probability p_s), and
V = the value to the offender of successfully committing the crime.

The expected cost to the offender of using method 1 (calculated using his subjective probabilities) is

$$C_1 = p_a\{pF_p + (1 - p)F_i\} + p_s p_c F_c p.$$

His expected benefit is

$$B_1 = pp_s V,$$

hence his expected gain from method 1 is[5]

$$B_1 - C_1 = pp_s(V - p_c F_c) - p_a\{pF_p + (1 - p)F_i\}. \quad \text{(A1)}$$

[5] For simplicity, I ignore the cost of using either method.

Similarly, for method 2, his expected gain is

$$B_2 - C_2 = (1 - p)p_s(V - p_cF_c) - p_a\{(1 - p)F_p + pF_i\}. \tag{A2}$$

How does the relative attractiveness of the two methods to the offender (calculated using his probabilities) depend on the punishments? We have

$$\begin{aligned} &\text{expected gain from method 2} - \text{expected gain from method 1} \\ &= (1 - p)p_s(V - p_cF_c) - p_a\{(1 - p)F_p + pF_i\} - pp_s(V - p_cF_c) \\ &\quad + P_a\{pF_p + (1 - p)F_i\} \\ &= (1 - 2p)p_s(V - p_cF_c) - p_a\{(1 - 2p)F_p + (1 - 2p)F_i\} \\ &\quad = (1 - 2p)p_s[(V - p_cF_c) - p_a(F_p + F_i)]. \end{aligned} \tag{A3}$$

Looking at equation (A3), we observe that the right-hand side (rhs) does not depend on F_p and F_i separately, but only on their sum. If our objective is to make method 2 more attractive than method 1, we can do so by making the punishments sufficiently large (if $p < .5$) or small (if $p > .5$), but we need not distinguish between possible and impossible attempts. Note that $p < .5$ corresponds to the (more plausible) assumption: the offender does not know which method works, but he considers the possible method more likely to work than the impossible method.

Under what circumstances will the offender choose method 2 over no attempt at all? From (A2) we have:

$$\begin{aligned} 0 < B_2 - C_2 &= (1 - p)p_s(V - p_cF_c) - p_a\{(1 - p)F_p + pF_i\} \\ &= (1 - p)[p_s(V - p_cF_c) - p_a\{F_p + F_ip/(1 - p)\}] \end{aligned} \tag{A4}$$

$$\begin{aligned} &= (1 - p)[p_s(V - p_cF_c) - p_a\{F_p + F_i\}] + [(1 - p)(p_aF_i) \\ &\quad - p_aF_ip] = (1 - p)[p_s(V - p_cF_c) - p_a\{F_p + F_i\}] + (p_aF_i)(1 - 2p). \end{aligned} \tag{A4'}$$

Comparing this with equation (A3), we observe that, if $p < .5$ and the rhs of equation (A3) is negative, then the rhs of equation (A4) is negative a fortiori, so the equation does not hold. Hence, if $p < .5$ and equation (A4) holds (method 2 is better than nothing), then the rhs of equation (A3) is positive and method 2 is also better than method 1. So if $p < .5$, equation (A4) is a necessary and sufficient condition for the offender committing the crime by method 2. Hence, if we wish to prevent the crime, we must do it by making the rhs of equation (A4) negative. Looking at the equation, we observe (in the final bracket on the right) that F_i enters with a weight of $p/(1 - p)$ relative to F_p. If $p < .5$, $p/(1 - p) < 1$, hence punishing an impossible attempt provides less deterrence than an equal punishment of a possible attempt, but still provides some deterrence.

Suppose, on the other hand, that $p > .5$: the offender believes voodoo is more likely to work than poison. Looking at equation (A3), we observe that method 2 is preferred if and only if

$$[p_s(V - p_cF_c) - p_a(F_p + F_i)] < 0. \tag{A3'}$$

Looking at equation (A4'), we observe that, if (A3') holds, then method 2 has a negative net return. Hence, if $p > .5$, the criminal either uses voodoo or nothing.