

Criminal Law

SOMEONE COMMITS a crime. He is arrested and punished. That is a cost to him, hence a reason not to commit the crime. Seen from this perspective, the explanation of criminal law is simple: It is a way of enforcing property rules. If your car is worth more to me than to you, I can buy it from you. If the law prevents me from stealing it from you, that changes the outcome only when it is worth less to me than to you, in which case the efficient outcome is for you to keep the car.

This argument assumes that crimes always have a cheaper market substitute, purchase instead of theft. What about assault? I get enormous pleasure from responding to your unkind remarks about my beloved Macintosh by slugging you. The consensual equivalent, a boxing match, say, would be no substitute. It is at least possible that my pleasure from slugging you is greater than your displeasure from being slugged, in which case that particular assault may be an efficient crime. Perhaps what we should aim for, in criminal law as in tort law, is not enough punishment to always deter but a punishment—more precisely, since not all criminals are caught, a combination of punishment and probability of being punished—that imposes the damage done on the one doing it. I will then assault you only when doing so is efficient.

Judge Posner has offered an example of an efficient crime that may seem more plausible than mine, especially to readers running Windows. A hunter, lost in the woods and starving, stumbles across a locked cabin containing food and a telephone, breaks in, feeds himself, and calls for help. His gain is more than the owner's loss, so his crime is efficient. It cannot be replaced by a market transaction because the owner of the cabin is not there to transact.

One way the legal system might permit such an offense is by setting the expected punishment roughly equal to damage done. The hunter then demonstrates that the crime is efficient by his decision to commit it. The value to him of breaking into the cabin must be more than the cost, which he pays in his punishment, or he wouldn't do it. Arguably, that is how we enforce parking laws. Illegal parking imposes costs on others. If it is sufficiently important to me, I demonstrate that by being willing to pay the price of occasional parking tickets.

Another way in which we might permit efficient crimes is by modifying the law so that they are no longer criminal. That is how we actually

1000

Urban
CA
S. H.

by
parking

handle the lost hunter problem; he is excused from criminal liability under the doctrine of necessity. Similarly, if the reason you are driving seventy-five is that your wife is in the back seat going into labor, the traffic cop may escort you to the hospital instead of writing you a ticket. This method works only when the special circumstances that make your act efficient are ones the court, or cop, can observe.

To test just how strong your support for efficiency is, consider a case of efficient murder. A wealthy sportsman concludes that the only game dangerous enough to be really worth hunting is man. He offers ten adventurers one hundred thousand dollars each in exchange for their agreeing that he may choose one of the ten at random and attempt to kill him. They agree. Should the law recognize the agreement and, if his hunt is successful, hold him innocent of murder on grounds of consent, or is that pushing freedom of contract a little too far?

The killing is efficient. It produces a net gain *ex ante*, as demonstrated by the assent of all concerned. Yet most of us would find it objectionable. Some might defend that conclusion along the lines of the commodification argument of chapter 13. Others might argue that the hunt is too likely to impose costs on innocent third parties caught in the line of fire. But one can always modify the scenario to eliminate such problems; perhaps the sportsman owns his own private hunting preserve. Such modifications are not likely to eliminate our reservations. Perhaps we have finally found a limit to efficiency as a source of law.

Coming to Terms with the Poverty of Our Circumstances

I have offered two models for criminal law: enough punishment always to deter and enough punishment to eliminate all and only inefficient offenses. Neither describes what we have. Crimes occur, so we do not have enough punishment always to deter. And almost all of the crimes that occur are inefficient ones, crimes in which the damage to the victim is more than the gain to the criminal.

If the punishment lottery for murder, the probability of being caught and convicted along with the resulting punishment, is at its optimal level, we ought to have only efficient murders, murders where the killer's gain is more than the victim's loss. If that is how our system works, the reason we don't try to catch more murderers or punish them more severely is that if we did there would not be enough murders. That does not sound like any world I have lived in recently. Nor does our present law fit the Pigouvian prescription: Expected punishment equals damage done. To make that work for murder, where the damage done is a life, we would have to

combine the less than unit probability of conviction with some punishment worse than death, perhaps lengthy torture.

The reason we do not catch all, or even almost all, murderers is that doing so would cost more than it is worth. Part of the cost would be in extra police and courts. Part of it would be in the punishment of innocent defendants. A standard of proof low enough to convict everyone who is guilty will also convict some who are not.

Taking Enforcement Costs Seriously

Making sense of criminal enforcement requires us to think seriously about its costs. To deter crime we must catch offenders and punish them. Both activities are costly, so we should take those costs into account in deciding what punishment to impose. The cost per offense usually increases with both probability of apprehension and severity of punishment, so a higher expected punishment costs more per offense to impose. It may sometimes, however, cost less in total, since a higher punishment will deter some offenses and offenses that are deterred do not have to be punished.

It is obvious why the cost per offense increases with probability of apprehension: It takes more police to catch fifty murderers out of a hundred than to catch twenty-five, and it takes more prosecutors and court time to convict them. To see why it also increases with the severity of the punishment, it is worth thinking a little more about punishment cost.

When a convicted criminal pays a thousand dollar fine to the state, the cost to him, which is what gives the punishment its deterrent effect, is a thousand dollars, but the net cost is zero. Every dollar the criminal loses the state collects. Punishment cost, defined as the difference between the cost the punishment imposes on the criminal and the benefit it provides to others, is zero.

Suppose the criminal cannot pay enough to provide the amount of deterrence we want to impose. Instead of (or in addition to) fining him, we imprison him for a year, which is equivalent, from his standpoint, to, say, a ten thousand dollar fine. The punishment costs him ten thousand dollars, but the enforcement system receives nothing. Instead the rest of us must spend money, say another ten thousand dollars, to run the prison. The net cost of the punishment, the criminal's loss plus the enforcement system's loss, is twenty thousand dollars. It is as if he had paid a fine of ten thousand dollars but we had collected a fine of minus ten thousand.

As we increase the size of the punishment we wish to impose, the number of offenders who can pay it as a fine decreases, so we tend to shift to

more costly punishments such as imprisonment. Hence increasing the severity of the punishment typically increases the punishment cost per offense punished.



This observation solves a puzzle first raised by Gary Becker in the article that started the modern economic literature on crime. Suppose we currently deal with some offense by imposing a 20 percent probability of a ten thousand dollar punishment. Why not switch to a 10 percent probability of a twenty thousand dollar punishment? The effect on the criminal will be the same, assuming he is risk neutral, so the deterrence will be the same. We will only have to catch and try half as many criminals, so we can save money by firing some police, judges, and prosecutors.

It works, so we repeat the process: 5 percent and forty thousand. And repeat it again. The conclusion seems to be that the efficient punishment lottery is a corner solution, an infinitely severe punishment imposed with an infinitesimal probability.

The reason this does not work is that we cannot simply double the fine forever; the criminal who can pay ten thousand dollars may be unable to pay twenty. As the punishment increases we are forced to shift to less and less efficient punishments, raising the punishment cost. An efficient system will accordingly choose, among combinations of punishment and probability that are equivalent from the standpoint of the offender and so have the same deterrent effect, the one that minimizes the sum of punishment cost and apprehension cost. Whatever level of deterrence we provide will then be provided with the least-cost combination of punishment and probability.

The next question is how much deterrence that should be. How many offenses should we deter, and how many should we fail to deter? It is inefficient for me to steal a television set that is worth five hundred dollars to you and only four hundred dollars to me. But it is still more inefficient to prevent me from stealing the set if the cost of doing so is two hundred dollars spent on police, courts, and prisons. The rule "prevent all inefficient offenses and only inefficient offenses" is correct only if doing so is costless. The correct rule in the more general case is to prevent an offense if and only if the net cost from the offense occurring is greater than the cost of preventing it. The reason we do not increase the punishment for murder may be, and probably is, that although we would like to prevent more murders than we do prevent (indeed, we might like to prevent all murders), the cost of doing so is more than we are willing to pay.

While the cost per offense increases with increases in expected punishment, the number of offenses decreases, since a higher expected punishment deters some offenses. The fewer offenses occur, the less must be spent to apprehend and punish offenders. If the decrease in offenses out-

weighs the increase in cost per offense, increasing the expected punishment reduces total enforcement and punishment cost. A system with higher punishments and thus fewer offenses then costs less than a system with lower punishments and more offenses. The additional cost of deterring one more offense is negative, making it efficient to prevent not only all inefficient offenses but some efficient ones as well—in order to save the cost of punishing them. In the extreme one could imagine a society where the penalty for shoplifting was death, with the result that there were no shoplifters and nobody ever had to be executed.

The Theory of Optimal Punishment

The next step is to make the argument more precise in order to see how, in principle, we would calculate the efficient punishment for any crime.

Consider a crime that does a thousand dollars worth of damage each time it is committed. We start by setting the expected punishment for the crime at \$900; each time the crime is committed, the offender is subject to a punishment lottery, a combination of probability and punishment, that is equivalent, from his standpoint, to a \$900 fine collected with certainty.

By raising the expected punishment to, say, \$901 we can deter one more offense. Unfortunately, imposing a slightly higher punishment on a slightly smaller number of offenders will increase our total cost of catching and punishing criminals by \$50. Should we do it? The answer depends on how large the net benefit is from deterring the crime.

Deterring the crime saves the victim a thousand dollars, but we must also take account of the effect on the criminal. If he got nothing from committing the crime, he wouldn't commit it. To calculate the net benefit from deterrence, we must subtract the gain the criminal would have gotten from the crime from the loss the crime would have imposed on the victim. How do we measure the criminal's gain? Not by asking him but by watching him. He will commit an offense if and only if its value to him is more than the punishment he expects for committing it.

We are interested, not in criminals in general or offenses in general, but in the gain that the criminal committing the particular offense we are deterring would have received by committing it. If that gain had been less than \$900, he would have been deterred even before we increased the punishment. If it was more than \$901, he would not have been deterred even by the higher punishment. It follows that the offense we deterred was worth between \$900 and \$901 to the offender. It did \$1,000 in damage. So the net damage done by the offense was about \$100. By deterring the offense we eliminated \$100 of net damage at a cost of only \$50, so it was worth doing.

Following out this argument, we keep raising the expected punishment as long as the gain is more than the cost. The process stops when the last offense deterred costs just as much to deter as the net damage it does. Putting it a little more formally:

Net damage = damage to victim – gain to criminal.

For the marginal offense, the one that will be deterred if we raise the punishment just a little more,

Gain to criminal = expected punishment,

hence

Net damage = damage to victim – expected punishment.

For the optimal punishment,

Cost of deterring one more offense = Net damage = damage to victim – expected punishment.

Rearranging gives us

Expected punishment = damage to victim – cost of deterring one more offense,

or, in the more compact notation of mathematics,

$$\langle P \rangle = D - MC.$$

If the cost of deterring one more offense is positive, as it will be if the supply of offenses is very inelastic so that it takes a lot of increased punishment to deter an offense, we should set expected punishment below damage done. We want to deter only very inefficient offenses, ones that will be deterred even by a low punishment, because only they do enough net damage to be worth the cost of deterring. If the cost of deterring one more offense is negative, as it will be if the supply of offenses is very elastic so that a small increase in expected punishment deters a lot of offenses, we should set expected punishment above damage done. We are deterring not only all inefficient offenses but some mildly efficient ones as well, a few hungry hunters and outraged Macintosh users, in order to save the cost of punishing them. We are still not deterring the very efficient offense, the lost hunter who is not just cold and wet but literally starving. He is willing to buy the use of someone else's cabin even at our inflated price.

This solution to the problem of setting the level of punishment combines elements of two different intuitions: punishment equal to damage done ("an eye for an eye," "make the punishment fit the crime") and

enough punishment to deter. If catching and punishing criminals is easy and inexpensive, the optimum is about equal to damage done; enforcement and punishment costs are unimportant, so we simply design our system to deter all inefficient and only inefficient crimes. If the supply of offenses is highly elastic at some particular level of punishment, so that below that level there are many offenses and above it very few, and if we expect few offenses to be efficient, then we set the punishment at the point where any further increase would have very little deterrent effect to balance its cost—just enough punishment to deter most offenses.



So far I have been describing offenses as efficient if the gain to the criminal is more than the cost to the victim, with the lost hunter as my standard example. We can now see that there is a different sense in which some offenses can be described as efficient. There are some offenses that we would like to deter if we could do so costlessly, but not given the actual cost of doing so. These are offenses that it is inefficient for the criminal to commit but also inefficient for the rest of us to deter, given the cost of deterring them.

The analysis of optimal punishment brings us back, in a somewhat more sophisticated form, to our earlier discussion of property rules versus liability rules. Where the marginal cost of deterring offenses is low—where, as we put it in chapter 5, allocating resources via the court is inexpensive—we want a liability rule, setting damages roughly equal to damage done. Where the low cost of voluntary transactions makes efficient involuntary transactions rare and the cost of using the court is high, we want a property rule, setting damages high enough to deter almost all involuntary transactions and thus eliminate the cost of litigating them.

Why Count Benefits to Criminals?

You may have found something very odd about the last few pages. In considering the net benefit from deterring a murder I subtracted out the gain to the murderer. In analyzing punishment costs I treated the fine collected by the state as a benefit but the fine paid, or life or liberty lost, by the criminal as a cost. Throughout this discussion I have taken it for granted that costs and benefits to criminals go into the calculations that generate efficient law in the same way as costs and benefits to anyone else.

The obvious response is that costs and benefits ought to have a moral as well as an economic dimension. I have a right to my life and property, so my loss as the victim of murder or theft counts. You do not have a right to my life and property, so the loss to you when we prevent you from killing or robbing me does not count.

This is a persuasive argument, but, for the purposes of this book, it is a mistake. It is a mistake because it assumes its conclusions instead of proving them.

One of the attractions of the economic analysis of law is that it provides a way of answering questions about what the law ought to be, what rights we ought to have. It starts with what looks like a very weak premise—that one should design legal rules to maximize the size of the pie. It assumes nothing at all about the sorts of things we expect legal and ethical rules to be based on: desert, rights, justice, fairness.

Starting with this premise, economic theory, and very little else, one produces—by the end of this book we will have produced—a long list of prescriptions. They include:

Theft and murder should be punished. Contracts should be enforced. The imposition of criminal penalties should require higher standards of proof than the imposition of civil penalties.

We start with economic efficiency and end with conclusions that fit reasonably well both existing legal rules and our ethical intuitions. Somehow we get out quite a lot more than we put in. That is one of the reasons the project is interesting. If instead of treating all benefits to everyone equally we first sort people into the deserving and the undeserving, the just and the unjust, the criminals and the victims, we are simply assuming our conclusions. Benefits to bad people don't count, so rules against bad people are automatically efficient. We cannot deduce moral conclusions from economics if we start the economics by assuming the moral conclusions.

What acts are or are not crimes is one of the things our theory is supposed to tell us. Murder may be a simple case, but what about speeding? What about breaking into an empty house when you are lost and starving? What about slugging someone who has the presumption to suggest that Windows is a better operating system than the Mac OS? Economic analysis provides tools for answering these questions. If we treat it, instead, as an elaborate machinery for justifying the answers we already have, we will learn little that we do not already know.

A second reason to include costs and benefits to criminals in our calculations, and one that links back to our moral intuitions, is that even if we agree that it is good to prevent crimes, we must still decide how good it is and hence how hard it is worth trying to do it. Compare a poor man who shoplifts with an arsonist who burns down buildings for the fun of it. Over the course of a year each happens to impose the same costs. Both are committing crimes we would like to prevent. But we are willing to go to a lot more trouble to stop the arsonist than to stop the shoplifter—because his crime is such a waste. Similarly, most of us are far more will-

ing to excuse a murderer when his murder was committed "out of necessity"—when, as in a lifeboat situation, the only alternative was his own death. Not only does economic theory imply that gains to criminals count, but a more careful examination of our own moral intuitions, motivated by economic theory, suggests that gains to offenders are not entirely irrelevant to our moral judgments.

I conclude that I have been doing it right. Economic analysis should give equal weight to the costs and benefits of murderer and victim. If doing so produces the conclusion we want—that murder is a bad thing—that is interesting. If it does not, that too is interesting.

The Most Efficient Punishment of All

You are convicted of embezzling money from your employer. One consequence is that you spend the next year in jail. A second consequence is that, when you get out and go on the job market, your superb qualifications for the job of corporate treasurer get you no job offers. Your punishment consisted of both a jail sentence and stigma, the cost to you of other people's knowledge of your crime.

Stigma is a very real punishment. Economist John Lott has done two empirical studies of its size, one dealing with white-collar criminals and one with corporations charged with cheating their customers. The first found that the loss of income after conviction made up a substantial part of the total punishment. The second found that the loss in corporate value, measured by stock value, due to stigma was many times larger than the nominal punishment.

How do you measure the loss due to stigma? Lott looked at the stock price of companies accused of offenses against their customers. Using a multiple regression, he calculated how in the past the stock price of each company had been related to other variables, such as the stock price of other companies in the same industry. He then predicted from the regression what the company should have been worth just after the charges, compared that with what its stock actually was worth, and interpreted the difference as the loss in company value due to the charges. The loss he found was, on average, many times higher than the highest fine that would be imposed if the charges turned out to be true. He concluded that most of the loss was due not to the prospect of paying a fine but to the loss of reputation. The real function of the court process was not to provide punishment directly but to generate information.

Stigma differs from other punishments in at least two interesting ways. The first is that, unlike all other punishments, it can, and usually does, have a negative net cost. To prove that, try a simple experiment. First



get convicted of embezzlement and serve out your term. Then go to a prospective employer and tell him that you realize he is reluctant to hire a convicted embezzler as company treasurer, but you will make it worth his while by agreeing to work for a lower salary than other candidates for the job. If you can agree on a salary, you have demonstrated that the value to you of getting the job was more than the cost to him of giving it to you. If, as I think more likely, there is no salary he will offer and you accept, you have demonstrated that the cost to him of having an embezzler as treasurer is more than the value to you of getting the job.

Stigma is information, and information is, with rare exceptions, valuable, since it allows people to make more nearly correct choices. The knowledge that you are a crook is valuable to potential employers. If you can still persuade them to hire you by offering to work for less, the stigma has simply transferred money from you to them since, without that information, they would have hired you at the normal wage. If you cannot persuade them to do so, the information must have been worth more to them than it cost you. So stigma can be, and often is, a form of punishment with net negative cost, one that benefits other people by more than it hurts the person being punished.

Stigma has another special characteristic: How efficient a punishment it is depends on whether the person who was convicted was guilty. If we convict someone who is innocent, stigma becomes a very inefficient punishment, since the information we are creating is false. So the fact that criminal conviction usually does, and civil conviction usually does not, impose stigma is another reason why it makes sense to require a higher standard of proof for the former.

Should the Rich Pay Higher Fines?

A very long time ago, when I first started teaching economic analysis of law, one question routinely came up: If we punish offenses with fines, should the fine be the same for rich and poor? I thought it should; my students thought it should not.

My argument followed directly from the Pigouvian approach to law. A particular offense imposes some number of dollars worth of damage. If the benefit to the offender is more than that, his offense produces a net benefit; if less, a net loss. By setting punishment equal to damage done we get the efficient result, for both rich and poor. Rich people may end up committing more assaults, just as they buy more Cadillacs, but that is a feature, not a bug.

John Lott carried the argument one step further, in one of his more politically incorrect (which for John means very politically incorrect in-

deed) articles. He observed that rich people who commit crimes are, on average, less likely to get convicted than poor people, because they have better lawyers, and offered that fact as evidence of the efficiency of our legal system. Most crimes, after all, are punished by imprisonment, not fines. Two weeks in jail represents a larger dollar cost to a rich man than a poor man, so if we want both to face the same expected cost in dollars we must balance the higher cost with a lower probability of conviction.



The argument, mine and John's, is straightforward. It is also wrong, as I eventually realized when I stopped explaining to my students why they were wrong for long enough to start thinking about the problem and, in particular, the implication of enforcement costs. Equal dollar expected punishment is the right rule in a world where both catching and punishing criminals is costless, so all we have to worry about is deterring all and only inefficient offenses, but that is not the world we live in. The reason we do not raise the punishment for murder is not the fear that if we did we would have too few murders.

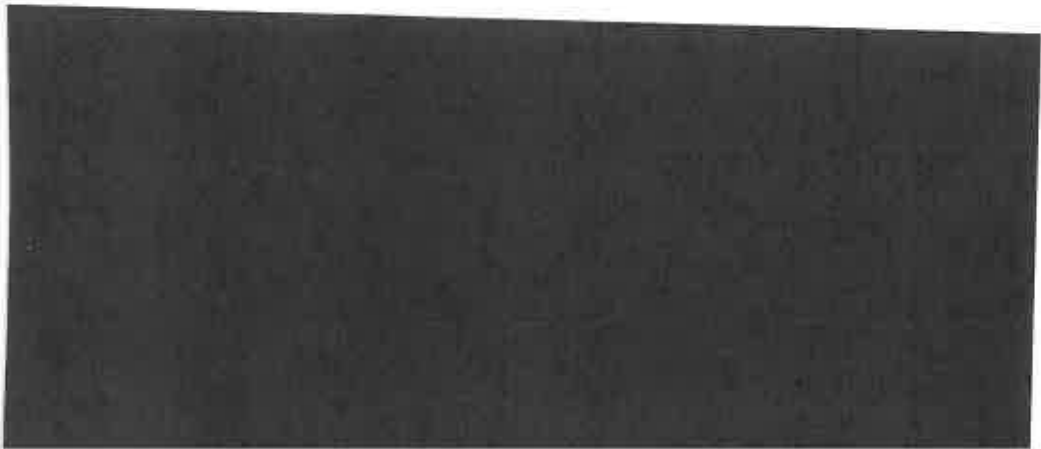
Once you take account of enforcement cost, the optimal punishment depends not only on damage done but also on how hard it is to deter offenses. In many cases the answer to that question is different for different people. As my students kept pointing out—to my deaf ears, since I already knew they were wrong—it takes a higher punishment to deter a rich criminal than to deter a poor one. If deterrence is expensive, it may make sense to impose on each sort of criminal just enough punishment to deter most offenses, which requires different punishments for different people.

Consider two sorts of offense—those that have a roughly equal payoff in utility for rich and poor and those that have a roughly equal payoff in money. Stealing a hundred dollars provides the same amount of money to a rich man as to a poor man, so the same fine should deter it. Indeed, since the time of the rich man is worth more dollars per hour than that of the poor man, if it takes each of them the same amount of time to steal a hundred dollars, the rich man should be deterred by a lower fine than the poor one.



Contrast that with an offense such as slugging someone you are mad at or saving ten minutes by speeding. The money value of the offense is higher to the richer offender, so it takes a higher (money) punishment to deter him. That may be an argument for imposing higher fines on richer defendants or equal jail sentences with equal probability.

As I discovered, my students' intuition was in part consistent with a more sophisticated application of economic analysis than I was offering them, but only in part. If you work through the analysis, the efficient rule sometimes involves charging the rich higher fines. But sometimes, with some assumptions about supply curves and cost of enforcement, the



efficient rule is to punish the poor and ignore the rich—because a punishment high enough to deter rich offenders costs more than it is worth.

There is a second difference between rich and poor with slightly less ambiguous implications. Fines are more efficient punishments than imprisonment, and richer offenders can pay higher fines. Even if neither offender can pay a sufficiently high fine, imposing a given dollar punishment via imprisonment requires fewer days in jail for a higher-income offender and is therefore cheaper. So punishment costs (per dollar of punishment) should decrease as income rises, which implies a higher efficient dollar level of punishment for richer offenders.

Or perhaps not. Punishing rich criminals is cheaper than punishing poor criminals. But convicting rich criminals is more expensive because, as John Lott pointed out, they have better lawyers. In designing an efficient legal system, we must take account of both sorts of costs. That might, depending on the numbers, bring us back to letting the rich get away with murder or, more plausibly, drunk and disorderly. If so, you will have to decide for yourself whether justice is worth the cost.

What I Have Left Out

Throughout this chapter I have taken it for granted that the purpose of the criminal law is deterrence, making it in the interest of potential criminals not to commit crimes. While that assumption greatly simplifies the analysis, it may also ignore some relevant considerations.

One is that some sorts of criminal punishment achieve a different objective—incapacitation. Someone in jail may still be able to commit crimes, but his opportunities, and his potential victims, are much more limited than if he were free, so jailing a criminal should reduce the damage he does. Hanging him reduces it even more.

In a simple Pigouvian world where convicting criminals is costless and all punishments are fines, criminals do no net damage; only efficient crimes occur, and we do not want to prevent those. In that world incapacitation is worthless. But in the more difficult world of positive enforcement and punishment costs, preventing a crime by incapacitating the potential criminal saves the cost of catching him, the cost of punishing him, and the net damage done by the crime, which is almost always positive. Hence incapacitation is a net benefit.

A more elaborate version of my analysis would include that benefit as a subtraction from punishment cost. Locking someone up costs his freedom and some of our taxes, but it also provides a benefit by reducing his opportunity to commit crimes, so imprisonment is a more efficient punishment than a simple calculation would suggest. Similarly for execution.

The logic of the argument remains the same, but the relative efficiency of different punishments might change.

A second argument offered for locking up criminals is that it gives us an opportunity to change them into good citizens, to rehabilitate them. This belief has been in and out of fashion several times over the past two centuries; the original penitentiaries, built in the early nineteenth century, were called that because they were places where people were supposed to learn to repent their crimes. Methods for teaching penitence included shaved heads, prison uniforms, hard work, solitary confinement, and Bible study. My own suspicion is that if we cannot teach people to be virtuous with twelve years of public schooling, a few more years of imprisonment are unlikely to do the job. I will therefore continue, for the most part, to ignore rehabilitation, while noting that insofar as it does exist it too can, in principle, be folded into the calculation of punishment cost as a subtraction, a negative cost.

Another way in which I have implicitly simplified the analysis is in considering only one crime at a time. As I pointed out back in in chapter 1, that is sometimes a mistake. If two crimes are substitutes for each other—armed robbery and armed robbery plus murder, in my earlier example—raising the punishment for one may deter the criminal into committing the other instead. An efficient legal system must take account of that possibility, a problem known in the literature as “marginal deterrence.” Doing so makes the analysis, and its explanation, considerably more complicated. Interested readers can check that claim by looking at an article that I co-authored with William Sjostrom, in which we show how the problem of marginal deterrence could be incorporated into the analysis of efficient punishment.

A final objection that may be offered by some is that my central assumption, that punishment deters, is false. One answer is that it is not an assumption but a conclusion; the assumption is rationality. A second answer is that, popular myth to the contrary, the evidence is overwhelmingly in favor of deterrence in the economist's sense. There have been many statistical studies measuring the effect on crime rates of changes in either the probability of apprehension, the punishment, or both; with few exceptions, they show that increasing expected punishment reduces crime rates.

Such studies also consistently show that crime rates are more sensitive to probability than to punishment. Increasing the chance of being caught and convicted from 10 to 20 percent reduces crime more than increasing the penalty from one year to two. This is sometimes interpreted as showing that criminals are risk preferrers, that they prefer the riskier 10 percent chance of a two-year sentence to the less risky 20 percent chance of one year, even though the expected cost in years is the same.

There is a simpler explanation. The cost of being tried for an offense is not limited to the punishment imposed by the court after conviction; it also includes time spent awaiting trial or money spent raising bond by offenders who have the good luck eventually to be acquitted, litigation costs in time and money, and stigma. Suppose the combined effect of all of those costs is equivalent to an extra year in jail. We then have

- 10 percent chance of 2 years in jail = .1(2 years in jail + 1 year in other costs) = .3 years,
 20 percent chance of 1 year in jail = .2(1 years in jail + 1 year in other costs) = .4 years.

The example generalizes. If the number of years in jail represents only part of the punishment and the other part is fixed, a given percentage increase in the jail term raises the expected cost to the criminal by less than the same increase in the probability of conviction and therefore has less deterrent effect even if criminals are risk neutral with regard to punishment.

Why Not Hang Them All?

Suppose armed robbery is currently punished by a ten-year prison term imposed on offenders with a probability of .6. Research on the tastes of potential criminals demonstrates that they are indifferent between a 60 percent chance of ten years in jail and a 10 percent chance of death. Obviously a reform is in order. We shut down the prison. Every time we convict a criminal, we roll a die. 1-5 we let him go, 6 we hang him.

On the face of it this is an unambiguous improvement. The criminal is as well off *ex ante*, we are getting the same deterrence so victims are no worse off, and we no longer have to spend money on prisons. There is, however, room for still more improvement. Why waste time and energy catching and convicting criminals only to turn them loose? We reduce our expenditure on courts and police until we are convicting only one criminal in ten, thus saving money on apprehension as well as on imprisonment.

The argument applies to any crime currently punished by imprisonment. If the sentence is only a month, we use a 720-sided die instead of a six-sided one. It follows that an efficient legal system will make no use of imprisonment. Defendants who can pay fines will be fined, since fines are more efficient than execution. Defendants who cannot pay fines will be executed, with probability scaled to the seriousness of the offense. No penalty will ever be used if there exists another penalty that is both more

severe and more efficient, since we can always make the system more efficient by substituting the higher penalty at a lower probability.

If we are going to be serious about constructing an efficient set of punishments, we should consider a wider set of options. In many historical societies, prisoners were required to work off their sentence. One could imagine a modern society where a criminal had the option of X years or Y dollars. The X years could be served out in a conventional prison. Alternatively, a private prison factory could obtain possession of the prisoner, with his consent, by posting a bond for Y dollars. The agreement between prison and prisoner would specify the terms of employment under which he would work off this fine. A sufficiently trusted prisoner could take an ordinary job, live an ordinary life, and pay his fine from his income. A more dangerous prisoner would require more guarding, would work inside the prison, and would pay off his fine more slowly. Prisoners too dangerous or unproductive to be employed at a profit could simply serve out their terms.

Whether a convicted criminal is judgment proof depends in part on whether it is in his interest to be; the same person might be "unable" to pay a fine but willing and able to pay a gambling debt of the same amount—given the alternatives. A legal system that routinely allowed convicted criminals to choose between fines and less attractive punishments, whether imprisonment or probabilistic execution, would make it in the interest of criminals to make sure that if they were caught they could pay.

Another alternative is to make execution more efficient. In eighteenth-century England corpses were routinely sold to surgeons for dissection. In twentieth-century America their value would be much higher. Execution presents the ideal environment for harvesting organs for transplant, giving the body of a criminal in good health a market value in the tens of thousands of dollars.

We can now see what an efficient system of criminal punishment might look like. It would be designed to squeeze the largest possible fines out of convicted criminals, using the threat of more unpleasant alternatives for those who failed to pay. If the fines that victims can pay, even under such threats, are inadequate, they are supplemented by penal slavery for criminals who can produce more than it costs to guard and feed them, execution with the organs forfeiting to the state for those who cannot. Any prisons that do exist and do not pay for themselves are as unpleasant as possible, so as to produce as much punishment as possible per dollar of imprisonment cost. It is a consistent picture, and considerable parts of it can be found in the not very distant past, but not a pretty one. Nor does it much resemble the system of punishment actually existing in modern

societies. It looks as though someone is making a mistake. Either there is something wrong with my analysis of what an efficient system would look like or our system is far from efficient.

Some Inadequate Responses

One argument offered against execution is that it is irreversible and courts sometimes make mistakes. But imprisonment is also irreversible. We cannot get back the years spent in prison any more than we can get back a life. The loss from one mistake is less, of course, but since imprisonment must be applied with higher probability than execution to get the same deterrent effect, there will be proportionally more mistakes. Imprisonment does have the advantage that we can partly correct errors if we discover them before the sentence has been served out, and we can to some extent compensate a prisoner after the fact for his false imprisonment. But both happen so rarely as to be an implausible explanation for the choice of imprisonment over execution.

Incapacitation provides a reason why imprisonment might sometimes be a more efficient punishment than execution or a fine. So does rehabilitation, assuming that we actually know how to change the tastes or skills of criminals in a fashion that makes them unlikely to commit further crimes. While these complications might modify the implications of a straight deterrence model, it is hard to believe that they could justify the modern system of punishments, in which imprisonment is heavily used, execution and penal slavery rare, forfeiture of organs unknown.

Execution in America at present is very expensive, due to the protracted litigation it requires; that could be seen as an economic argument against it. But it seems more natural to interpret this as a consequence of our reluctance to employ execution than as a cause. The argument for the superiority of execution does not require any increase in how careful we are about whom we convict. Historical societies that made extensive use of execution, such as eighteenth-century England or nineteenth- and early-twentieth-century America, did so without extensive and expensive litigation.

A final defense of our present legal rules is that they are a consequence of the tastes of the populace. Modern people would find execution, penal slavery, penal torture, or the dismemberment of executed criminals to provide organs for transplant distasteful. Such punishments thus generate large negative externalities. Given the existence of those tastes, our present legal rules are efficient.

I reject that explanation for the same reason that I earlier rejected the idea that benefits to criminals should not count in defining efficiency. An

economic theory that is free to eliminate anomalies by explaining them away as due to someone's unexplained and peculiar tastes has very little in the way of testability or predictive power. Our objective is explanation, not description. It is possible that people do have those tastes, but if unexplained tastes are the sole reason for the institutions we wish to explain, we have failed to explain them. And our project is much more interesting if we can show that the institutions our tastes favor are in fact efficient, implying that our peculiar tastes are actually an efficient set of norms.

The Virtue of Inefficient Punishment

Consider the world of efficient punishment from a more symmetrical viewpoint than we have so far employed, one that takes account of the incentives of all relevant actors. It is a world where, by persuading a court that someone is guilty of a crime, I can expropriate large amounts of the defendant's financial, human, and perhaps biological capital. In a world of efficient punishments, somebody gets most of what the convicted defendant loses. It is in that somebody's interest to convict defendants, whether or not they are guilty.


The conventional analysis of optimal punishment on which my analysis up to this point has been built is based on a mistake that has been extensively criticized in other contexts: the philosopher-king view of government. The old textbook literature on regulation and related topics treated market participants as rational, self-interested actors, but the state as a proxy for the author—a wise, benevolent, wholly altruistic organization, doing whatever would best correct the failures of the market. It was by abandoning that model that we got public choice theory and the modern analysis of regulation.

The orthodox theory of optimal punishment makes the same mistake. It treats criminals as rational, self-interested actors. But it treats the enforcement apparatus of police, courts, prosecutors, and legislature as a philosopher-king, with imperfect knowledge but only the best of motives.


One cost of that approach is that it makes it harder to include tort law and criminal law in the same theory, despite their obvious similarity of means and end. Tort law is enforced by the actions of private parties, criminal law by the actions of the state. It seems obvious that private plaintiffs ought to be treated as rational self-interested actors, and that is how they generally have been treated in the law and economics literature. By treating state actors differently, we not only obscure the similarities, we also make it harder to think clearly about the choice between privately and publicly enforced law—the project of chapter 18.

Once we start treating all actors symmetrically it becomes obvious that a system of efficient punishments has a substantial cost as well as substantial benefits. The cost is rent seeking. The legal system becomes a mechanism to be used by some people to expropriate other people—who respond by taking expensive precautions to avoid being expropriated. The population as a whole might well be better off with less efficient punishments.


This is not a purely theoretical problem. Consider some real-world examples:



Modern tort law offers opportunities for plaintiffs, or at least plaintiffs' attorneys, to engage in extensive rent seeking via class actions and claims for punitive damages. It is widely alleged that the result has been a host of unjustified legal actions, with attorneys enriching themselves at the expense of defendant corporations and their insurers. One example, discussed in chapter 3, was the popularity of Fraud on the Market claims. Even a small probability of a very large damage judgment may give an innocent defendant good reason to accept an out-of-court settlement. The elimination of American production of small airplanes can be, and has been, plausibly explained as a defensive reaction to that sort of rent seeking: Firms found that they were better off exiting the industry than risking future litigation.



Civil forfeiture law creates a similar incentive for public law enforcement officials. Under current law property used in the commission of certain crimes can be seized by the state. Seizure does not require evidence that the owner of the property violated any law—the case is against the property, not the owner—nor does it require the criminal conviction of anyone, owner or user. It is up to the owner who wishes to get his property back to demonstrate, in a civil action, that it was not used to commit a crime.



Since forfeiture under federal law is to the federal government, one might expect it to provide an incentive only to federal law enforcement agencies, but that ignores the relevant market transactions. The Coase Theorem is alive and well in the law enforcement industry. Federal agencies divide the gains, and thus the incentives, with local law enforcement agencies. After federal law was altered to authorize such sharing there was a striking shift of local law enforcement effort toward the enforcement of drug laws, which provide lots of opportunities for forfeiture. Many states have their own forfeiture statutes as well, some of which provide for directing forfeited assets to the law enforcement agency responsible for obtaining them. The incentives implied by such institutions pose an obvious risk—that law enforcement will be directed toward seizing property rather than preventing crime.

For the same problem in the context of criminal law, consider the Ruby Ridge killings. One point on which all parties seem to agree is that the

origin of the conflict was an attempt by the Bureau of Alcohol, Tobacco, and Firearms (BATF) to entrap Randy Weaver by persuading him to sell a BATF informer two shotguns whose barrels had been shortened to slightly below the legal length. The reason to do so was to force Weaver to spy for the BATF on other people who shared his (white separatist) political views. The intended gain to the enforcers was in services not cash, but the logic of the situation was the same. Since law enforcement agencies make extensive use of informers, it seems likely that similar actions have occurred in many less publicized cases.

Consider next a fictional example, now slightly dated:

In 1993 Vermont passed the first of the organ bank laws. Vermont had always had the death penalty. Now a condemned man could know that his death would save lives. It was no longer true that an execution served no good purpose. Not in Vermont.

Nor, later, in California. Or Washington, Georgia, Pakistan, England, Switzerland, Rhodesia.

Despite the new legislation transplant organs are still in short supply. The legal system responds to the resulting political pressures. At the end of the story the prosecution reads the capital charge for which the defendant is being tried:

The state will prove that the said Warren Lewis Knowles did, in the space of two years, willfully drive through a total of six red traffic lights. During that same period the same Warren Knowles exceeded local speed limits no less than ten times, once by as much as fifteen miles an hour. (From Larry Niven, "The Jigsaw Man," in *All the Myriad Ways*)

All of these examples demonstrate a common problem—the effect of efficient punishment on the incentives of enforcers, public (civil forfeiture and BATF entrapment) or private. The same institutions that, seen from the perspective of a philosopher-king model of law enforcement, produce an unambiguous improvement by lowering the cost of enforcing the criminal law have the potential, seen from a perspective of rational self-interest, to set off a costly rent-seeking struggle, a war of each against all, with each side trying to use legal institutions to expropriate others and avoid being expropriated.

Cannibalism, Rent Seeking, and Rape

For a more distant example of the same issue, consider the universal human prejudice against cannibalism. On the face of it the prohibition is irrational and inefficient. Once I am dead I have no more use for my body, so why not let someone else get some useful protein out of it?



The answer should now be obvious. In a society in which cannibalism is an accepted practice, everyone is at risk of becoming someone else's dinner. Each of us must go to a good deal of trouble to make sure that nobody has a safe opportunity to kill him. The ban on cannibalism means that most people, most of the time, have no good reason to kill each other, making life a good deal easier for all of us.

One of the Greek historians commented that the Persian empire was so well ruled that a virgin with a sack of gold could travel from one end of the Great Road to the other unharmed. The two values at risk in that case differ in an important respect relevant to this discussion. People who own sacks of gold rarely send them wandering about the countryside unprotected; they can, and usually do, keep them locked up safe. For a woman to lock up her virginity is, modern fantasies about medieval chastity belts to the contrary, a lot harder.

Rape may be the most common really serious crime. One reason is that most women have something—the use of their bodies—that many men want and that there is no inexpensive way of protecting. Some societies have solved that problem by literally locking up women's bodies, sequestering women, at least upper-class women, somewhere where only trusted men would have access to them. In a modern society that would be a very costly solution. Our society solves, or attempts to solve, the same problem by treating rape as a serious crime.

For another example of the same issue consider the question of whether there should be an entirely free market in organs for transplant. Organs are currently in short supply, and legalizing the market seems an obvious way of dealing with the problem. It would give individuals an incentive to sell the rights to their organs in advance, with the organs harvested only when the individual died, or to agree that organs could be sold at the time of death, with the money going to their heirs. More controversially, if the price were high enough, it would give people with two kidneys an incentive to sell one of them.

The best argument I know of against a completely free market in organs is the rent-seeking problem. Few of us carry more than a few hundred dollars in cash, but every healthy individual carries with him at all times tens of thousands of dollars worth of internal organs. With a completely free market, there would be a substantial incentive to kidnap people in order to dismember them, raising the risk of murder and the costs of defending against it. It has been alleged that this already is happening in some parts of the world, and organnapping features in a widely circulated urban legend, involving someone who wakes up, after a wild night, minus his kidneys.

One solution would be a market, but one with adequate precautions to establish chain of title to any organs sold. Another would be a legal re-

time in which organs could be sold by their owner, in advance or at time of death, but not resold.

Among nations of hunters, as there is scarce any property, or at least none that exceeds the value of two or three days' labour, so there is seldom any established magistrate or any regular administration of justice. Men who have no property can injure one another only in their persons or reputations. But when one man kills, wounds, beats, or defames another, though he to whom the injury is done suffers, he who does it receives no benefit. It is otherwise with the injuries to property. The benefit of the person who does the injury is often equal to the loss of him who suffers it. . . . The acquisition of valuable and extensive property, therefore, necessarily requires the establishment of civil government. Where there is no property, or at least none that exceeds the value of two or three days' labour, civil government is not so necessary. (Adam Smith, *The Wealth of Nations*, book 5, chapter 1, part 2)

Further Reading

David Friedman and William Sjoström, "Hanged for a Sheep: The Economics of Marginal Deterrence," *Journal of Legal Studies* (June 1993).

The title is lifted from an old English proverb: "As well hang for a sheep as for a lamb." At one time stealing either a sheep or a lamb was a capital offense, so you might as well steal the more valuable animal. The proverb nicely illustrates both the application of rationality to criminals and the fact that you do not have to be an economist to think like one.