1-1-1988

An Economic Explanation of Punitive Damages

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

Follow this and additional works at: http://digitalcommons.law.scu.edu/facpubs

Recommended Citation
40 Ala. L. Rev. 1125

This Article is brought to you for free and open access by the Faculty Scholarship at Santa Clara Law Digital Commons. It has been accepted for inclusion in Faculty Publications by an authorized administrator of Santa Clara Law Digital Commons. For more information, please contact sculawlibrarian@gmail.com.
AN ECONOMIC EXPLANATION OF PUNITIVE DAMAGES

David Friedman*

I. INTRODUCTION—THE PROBLEM

I start with an argument and a puzzle. The argument is the familiar economic analysis from which both the Hand rule1 and the usual economic justification for compensatory damages are derived. The puzzle is the existence of punitive damages.

The argument goes as follows. I take some action that imposes costs on you. For our present purposes, it does not matter whether the act is a tort or a crime, intentional (in the lawyer's sense) or accidental. From the standpoint of economic efficiency, the problem is not that I am imposing costs, but that I will decide whether to take the action on the basis of costs and benefits to only myself. If the action provides me with a benefit of $50 but imposes a cost of $70 on you, I will take it; conversely, if it costs me $10 and benefits you by $50, I will not. The efficient policy is to take those actions and only those for which net benefits are larger than their net costs, and in this example, I am not doing so.

There are two obvious solutions. One is to charge me for the cost I inflict on you. In the context of the criminal law, we set the expected punishment (penalty times probability of being caught and convicted)2 equal to the damage done by the crime. I then commit the crime if and only if the benefit to me is greater than the cost to you,3 which is the efficient outcome. In the context of

* John M. Olin Fellow in Law and Economics, University of Chicago Law School.
** I would like to thank the participants in, and organizers of, the Conference on Puni
tive Damages at the University of Alabama Law School for their contributions to the birth of this article, my colleagues at the University of Chicago Law School, especially Judge Richard Posner, for their comments and suggestions, and the Olin Foundation and the Center for the Study of the Economy and the State for financial support.
2. For purposes of simplicity, I ignore problems of risk aversion here and throughout; all actors are assumed to be risk neutral.
3. Judge Posner's famous example of such an "efficient crime" is the lost man who breaks into a locked cabin in the woods in order to get food and summon help. See R.

1125
the civil law (under strict liability), we require the tortfeasor to compensate fully his victim, thus transferring the entire cost to the tortfeasor. He then minimizes the sum of his costs (of accident and prevention) plus the victim's cost, which is the efficient policy.

The alternative solution is the Hand rule. Instead of transferring the cost my action creates back to me and letting me figure out how to minimize total costs, the court decides what I should have done in order to minimize total costs and then punishes me (by holding me liable for damages) if I do not do it. In the context of civil damages, where the Hand rule is normally applied, this means that I am deemed negligent if and only if the cost of precautions that would prevent the accident is less than the cost imposed by the accident times the probability that the accident will occur if I do not take the precautions. In the criminal context, this approach leads us to treat as crimes only those actions for which there is some presumption that total costs are larger than benefits.

We now have two solutions to the problem of using the legal system to prevent inefficient acts. From here on, I will limit myself to the first and consider it only in the context of civil law, since that is the context in which punitive damages appear. We then have a simple argument for the efficiency of compensatory damages. If the tortfeasor pays compensatory damages equal to the full cost imposed on the victim, he bears all the costs of his act and so makes the efficient decision.

We now have a puzzle. The argument as given says nothing at all about whether the tort was intentional or reckless, the two labels of culpability commonly used to justify punitive damages. So far as economic efficiency is concerned, that issue seems irrelevant.

Posner, Economic Analysis of Law 206 (3d ed. 1986). A more mundane example is the driver who exceeds the speed limit because he is in such a hurry that it is worth doing so, even after allowing for the costs that his speed imposes on both himself and others.

4. This can be generalized from the binary case of take/don't take precaution to situations where level of precaution is a continuous variable. The rule then requires that one take that level of precaution for which marginal cost of additional precaution equals marginal expected benefit.

5. Under this approach (and, I believe, under existing law), the starving hunter who breaks into the cabin has not committed a crime. Similarly, the driver who is speeding to get his wife to the hospital before she gives birth is unlikely to receive a ticket. In the one case the offense is excused under the doctrine of necessity, and in the other it is permitted by an exercise of discretion on the part of a police officer; but the effect is in both cases the same.
Economic Explanation

Consider two torts, one accidental and one intentional. The first is an auto accident—my car hits your car. The accident is not intentional in the sense in which the term is used in either the law or ordinary language, but it is intentional in a sense useful for economic analysis. The accident is the probabilistic result of decisions I make: how often to have my brakes checked, how fast to drive, whether to drive when it is raining, whether to drive after having a drink, how much attention to pay to the road and how much to the radio . . . . In making those decisions, I am, explicitly or implicitly, balancing costs and benefits. Compensatory damages force me to bear all of the costs of my action. Insofar as I am rational, I then choose the efficient level of precautions.

The second tort is an assault—my fist hits your chin. By hitting you, I impose a cost, in pain and humiliation, of say $50. Under a system of compensatory damages, I will have to pay you $50. If I am rational, I choose to hit you only if the pleasure I get from doing so is more than $50—say $70.

In that situation, compensatory damages will not deter the tort. That fact is not, however, a justification for punitive damages—at least, not from the standpoint of economic efficiency. In the situation as described, my hitting you produces a net gain of $20. Just as in the previous case, compensatory damages produce an efficient outcome. They fail to deter this particular assault because, from the standpoint of efficiency, it should not be deterred. It is an efficient assault.

If the analysis I have presented is correct, punitive damages should not exist. Yet they do exist, and they have existed in the common law for some centuries. One solution to this puzzle is to conclude that the common law is, in this respect, wrong—or at least inefficient. A piece of evidence that should perhaps make us hesitant to accept that solution is that it is the common law, not the conclusion of the economic argument, that fits the moral intuition of most observers. Whether or not they can find any flaw in the argument, few people find the idea that efficient assaults should be permitted an intuitively appealing one.

Another solution is to argue that compensatory damages as defined by the courts are not really compensatory—that for one of
several reasons they understate the costs actually imposed by tortfeasors, and that what are called punitive damages are merely the adjustment necessary to compensate fully the victims. The problem with this solution is that it fails to explain why punitive damages are limited to a small subset of cases—traditionally to those involving deliberate or reckless acts by the tortfeasor.

A. The Ellis-Tullock-Stigler Out

So far, I have been restating an argument most of which is made by Dean Ellis in a well-known article. The solution that he offered in that paper was to classify some benefits as "illegitimate satisfactions." The time and money that I save by having my brakes checked only once instead of twice a year is a legitimate benefit; I am entitled to weigh it against the cost imposed on others (and myself) by a slightly higher chance of an accident due to brake failure. The pleasure I get from slugging someone who has just expressed a wholly indefensible opinion about my favorite baseball team, on the other hand, is an illegitimate satisfaction. I am not entitled to count it on the positive side in deciding whether to slug him. If that pleasure is enough to more than balance the cost to me of fully compensating him, punitive damages are necessary and appropriate to deter me.

6. One argument is that ordinary damages fail to compensate the victim for the cost to him of litigating the case. Another is that some victims will not sue or will lose their cases due to court error, and that the tortfeasor therefore pays damages with a probability substantially less than one. A third is that ordinary damages are, or at least traditionally have been, limited to those damages easily measured, and they therefore omit important nonpecuniary costs. A number of these points are discussed in other articles in this issue and in Ellis, Fairness and Efficiency in the Law of Punitive Damages, 56 S. Cal. L. Rev. 1 (1982).

7. One possible answer to this question, suggested by Landes and Posner, is that deliberate torts are more likely to be concealed, hence they require a larger punishment when detected. One would think, however, that just the opposite would be the case for reckless torts. Another answer suggested by the same authors is that, if occurrences of a particular sort of tort are very unlikely to be efficient, punishment above damage done has no disadvantages, and it protects against accidental underdeterrence due to court error (underestimating damage done, for instance). Both of these arguments are made in V Landes & R. Posner, The Economic Structure of Tort Law 160-62 (1987).

8. Ellis, supra note 6.

9. As Dean Ellis pointed out in the article, the idea of illegitimate satisfactions had been used before (in other contexts), by Gordon Tullock, The Logic of the Law 243 (1977) and George Stigler, The Optimum Enforcement of Laws, 78 J. Pol. Econ. 526, 529 (1970). See Ellis, supra note 6, at 32.
Economic Explanation

I believe that what is illegitimate here is not the satisfaction of the intentional tortfeasor, but the argument of Ellis, Tullock and Stigler. To justify that claim, I must first digress onto a brief explanation of why the economic analysis of law, in particular the analysis of what legal rules are economically efficient, is an interesting enterprise.

B. Legal Rabbits from Efficient Hats

One starts with a single premise—that legal rules should be economically efficient. This may be viewed as a policy designed to approximate the utilitarian prescription of maximizing total happiness.10 The premise appears to have no other ethical content. It says nothing at all about desert, rights, justice, or fairness—the sorts of things on which we expect legal and ethical rules to be based.

Starting with this premise, economic theory, and very little else, one produces a long list of prescriptions. They include the following:

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. People who injure others should be required to pay damages sufficient to make the victims whole.

We have started with the goal of economic efficiency and ended up with conclusions that fit quite closely both existing legal rules and our ethical intuitions. Somehow we have gotten out quite a lot more than we put in.

Some of the congruence between economics, ethics, and law may be bogus—evidence only that if you present a sufficiently clever economist with an existing legal rule, he usually can find some plausible economic argument to show that the rule is efficient. But I believe, and I think most scholars in the field would agree, that most of it is real. Surprising though it may seem, a sub-

stantial fraction of both our law and our ethical judgements can be deduced from economic efficiency.

If this is true, it is surely interesting. From the standpoint of the philosopher, it suggests that our ethical judgments may contain considerably more disguised utilitarianism than we might have thought. From the standpoint of the economist and the legal scholar, it suggests various conjectures about how common law (and perhaps also our ethical intuitions) is formed, and why various common-law rules exist.

It is interesting for another reason as well. If the common law, the prescriptions of economic efficiency, and our ethical intuitions have some deep relation, then situations where they do not agree may be instructive. If the common law does not follow the rule we think is economically efficient, that may be evidence that our economic analysis is wrong. It also may be evidence that something has gone wrong with the common law, or that whatever forces push it towards economic efficiency apply in only some areas and not others. If the prescription of economic efficiency differs from our ethical intuitions (or the conclusions of ethical theories that we find convincing), that is a reason to rethink both the economic analysis and the ethics.

It should now be clear what is wrong with the notion of "ille-gitimate satisfactions." What Dean Ellis (and his precursors) have done is to deal with an apparent conflict between economic efficiency on the one side and legal and ethical rules on the other not by rethinking either, but by simply defining the problem out of existence. They have created a special exception to the normal definition of economic efficiency (which defines values in terms of observed behavior, not moral judgments) designed to force the economic analysis to give the right answer. They have chosen to put their ethical assumptions in at the beginning of the process—after which it is hardly surprising to find them coming out at the end. By appropriate expansions of their technique, one can get "economic efficiency" to imply anything one likes—and, in the process, entirely empty it of content. This is, I presume, one reason why Dean Ellis proposed his solution in such a tentative way.

11. An alternative explanation is that there is some deep connection between the prescriptions of utilitarianism and those of some other ethical principle upon which our judgments are in large part based.
II. THE CASE FOR PUNITIVE DAMAGES

If Dean Ellis's solution is not satisfactory, is there another that is? Can punitive damages be explained in terms of economic efficiency? I believe that the answer is yes. Most of the remainder of this Article will be devoted to doing so.

Before I begin, however, I must make one important caveat. While I am prepared to argue that economic efficiency explains, or is at least consistent with, substantial elements in the common law, I am not claiming that the common law is always and everywhere economically efficient. If one wishes to argue for the efficiency of the common law, the first step is to identify the past twenty years as experimental error. What I will be explaining, therefore, is not the doctrine of punitive damages as it is applied by courts in the United States at the present time, but as it traditionally was applied prior to its recent expansion. My view of that doctrine is based mostly on the other articles in this issue and on Dean Ellis's 1982 article.

The traditional doctrine of punitive damages, as I understand it, had the following characteristics. First, punitive damages were limited to torts that were deliberate or reckless. Second, punitive damages were not typically a large multiple of compensatory damages. Third, compensatory damages were narrowly defined.

A. Getting Damages Right

Earlier in this Article, I sketched the argument that implies compensatory damages lead to the efficient result. The first step in explaining punitive damages is to point out that that argument, and that conclusion, are wrong, save in a very restricted set of circumstances. The reason the argument is wrong is that it ignores the cost of the legal system itself. Imposing punishments, whether civil or criminal, is not a costless activity. In the case of criminal law, there is the cost of catching and punishing the criminal. In the case of the civil law, there is the cost of litigating cases. In both civil and criminal law, the cost depends very much on the size of the punishment being imposed.

12. As other articles in this issue make clear, a variety of terms was used, but most of them fit into one or the other of those categories.
There are two ways in which the size of the punishment affects the total cost of imposing it. Increasing the amount at stake in litigation increases the amount spent on it; lawyers make more out of a million dollar case than out of a thousand dollar case. At the same time, increasing the punishment for an offense, whether in the form of damages, fine, or imprisonment, decreases the frequency with which the offense occurs. The fewer torts occur, the fewer must be litigated. Thus increasing the size of a punishment simultaneously increases litigation costs (or enforcement, litigation, and punishment costs in the criminal case) per offense, and decreases the number of offenses. Total cost (cost per offense times number of offenses) may go up or down.

The earlier argument for damages equal to harm done ignored the costs of litigation. It is correct only if litigation is costless, or if the cost is independent of the amount of damages. In all other cases, the efficient level of damages is not equal to the amount of harm done. As we will see shortly, it may be more or less.

In the Appendix, I present the mathematical derivation of the efficient level of damages. Here I simply will give a verbal analysis of two cases, one in which the efficient level of damages is lower than the actual damage done and one in which it is higher.

Assume there is some tort that imposes a cost on the victim of $1000. Further assume that if committing the tort results in paying $1000 worth of damages, it will be committed a thousand times a year. Since we assume that tortfeasors are rational, all thousand occurrences are "efficient torts"; the cost to the tortfeasor of precautions that would prevent them is higher than the damage.

13. The latter situation exists if the two effects happen to exactly cancel; while this is possible, there is no reason to expect it.

14. It is occasionally said that the efficient punishment is the damage done including the cost of litigation. This is correct if "including" means "modified to take account of." It is incorrect if, as a reader would normally assume, "including the cost of litigation" means "adding in the cost of litigation," since in that case the efficient punishment would always be at least as great as the damage done.


16. Throughout this discussion, I will assume that guilty tortfeasors are always sued and always lose. The analysis can be modified to include uncertainty by replacing the punishment with its expected value.
payment (equal to damage done by the tort), which is why the tortfeasor does not take the precautions. Is this the efficient outcome?

Probably not—if litigation costs are taken into account. Assume that the cost of litigating a claim is equal, on average, to half the amount paid out.\(^{17}\) We are deterring all inefficient occurrences of the tort and only inefficient occurrences, but we are doing so at a cost of $500,000 per year in litigation.

Suppose, to take an extreme case, that over some range of damages the supply of offenses is perfectly inelastic: changing the damages that the tortfeasor must pay has no effect on the number of offenses. There is no one at all who will commit the tort if the damage payment is $500 but will not commit it if the payment is $1000. In this situation, we clearly make the system more efficient by lowering the fine to $500. The number of offenses remains the same, and the litigation cost is cut in half. The same is true if the supply is highly, but not perfectly, elastic. If lowering the damages to $500 increases the number of offenses to 1100, the saving in litigation costs more than makes up for the cost of 100 inefficient torts.

How do we know that? The extra 100 occurrences impose a cost on the victims of $100,000 per year. Preventing each of them would cost the tortfeasor at least $500—that is why he no longer prevents them when the resulting damage payment is reduced to $500. So potential tortfeasors save at least $50,000 per year by not preventing those 100 occurrences of the tort. On net, there is a loss of at most $50,000 per year and a reduction in litigation costs of $250,000 per year. We are better off, on net, by at least $200,000 after the legal change that reduces damages.

This argument implies that the efficient damage payment is less than the actual harm done by the offense if the supply of offenses is sufficiently inelastic. To see why the efficient level of damages might be more than the actual damage done, we will next consider the same situation with one change—a highly elastic supply of offenses. This time, we assume that an increase in the

\(^{17}\) For simplicity, I am assuming here that the only cost of running the system is litigation cost, and that it is simply proportional to the amount paid out in damages. A fuller version of the analysis would take account of additional costs, such as those associated with incorrect court decisions, and would allow for the possibility that costs might be related to damages in a more complicated fashion, as discussed in the Appendix.
damage payment from $1000 per offense to $1100 per offense will reduce the number of occurrences to zero.

In the previous case, lowering the damage payment permitted some inefficient torts; this time, raising it deters some efficient ones. Potential tortfeasors who would have committed the tort if the cost to them of doing so was a damage payment of $1000 but no longer do so if it is $1100 have a cost of precaution somewhere between the two values. Since the injury done to the victim is $1000, there is a net cost (prevention cost minus the savings to the potential victim of not becoming an actual victim) of between zero and $100 for each such tort prevented. So eliminating 1000 efficient torts results in a net cost of between zero and $100,000.

But eliminating 1000 efficient torts also eliminates the associated litigation, for a saving of $500,000 per year. On net, the increase in damages reduces total costs (precaution, injury, and litigation) by something between $400,000 and $500,000 per year.

We have now seen why the efficient level of damages is not, in general, equal to the harm done by the tort, once litigation costs are taken into account. If the supply of offenses is highly inelastic, a large increase in damages and resulting litigation costs is required to deter one more offense. It is not worth paying that price to deter an offense that is only slightly inefficient—a thousand dollar accident that could be prevented only at a cost of nine hundred and ninety dollars. It is worth paying it to deter a very inefficient offense—a thousand dollar accident that would cost only a hundred dollars to prevent. The efficient damage payment is substantially less than the harm done, so as to deter only very inefficient offenses.

If the supply of offenses is elastic, on the other hand, increasing the level of damages actually decreases litigation costs, since the decrease in the number of offenses outweighs the increase in litigation cost per offense. In this case, it is worth accepting the cost of deterring offenses even when doing so produces a small net cost (a thousand dollar accident that costs eleven hundred dollars to prevent) in order to avoid the cost of litigating them. We do so by setting the level of damages above the amount of harm done. Offenses that are very costly to prevent (for example, a thousand
dollar accident that would cost ten thousand dollars worth of precautions to prevent) still occur—and should.16

Throughout this discussion, I have been assuming that torts are detected and tortfeasors convicted with certainty. If we drop that assumption for a moment, we can generate one more prediction. Assume, as seems plausible, that the production function for detecting and convicting tortfeasors has rising marginal cost—it costs more to raise the fraction of guilty tortfeasors convicted from 90% to 91% than from 10% to 11%—and similarly for the production function for collecting damages. It is then easy to show that the least expensive way of increasing the expected damage payment involves increasing both the amount of damages collected from convicted tortfeasors and the percentage of guilty tortfeasors convicted.

What does this imply for a system of civil damages? Under such a system, detecting and prosecuting torts is the job of the victim. The higher the damages he receives when successful, the greater his incentive to find and convict those who have committed torts against him. So, not only is it efficient for those who commit torts for which the supply is more elastic to pay more damages for a given amount of injury, it is also efficient for their victims to receive more damages in order that they will have a greater incentive to sue.

B. The Relation Between Elasticity and Accident

We are now almost finished. The final step is to link this argument to the case of punitive damages. We do so as follows:

Consider a typical accidental tort—say an automobile accident. In deciding how often to drive, what kind of a car to get, and all of the other things that affect the probability of accidents, potential damage payments are one of the costs that the driver considers, but only one. Most of those decisions involve other and, allowing for the low probability of an accident, much larger costs and benefits. Even if we limit ourselves to costs directly associated with accidents, damage payments, although important, are not the

18. As I show in the Appendix, the argument can be made more precise. If litigation costs are proportional to damage payments, as I assume in my examples, then the efficient level of damages is equal to the harm done if elasticity of supply (at the efficient level of damages) is one, greater if elasticity is greater than one, and less if it is less than one.
most important costs; most drivers are more concerned about killing themselves than about having to pay damages for killing someone else.

If the cost of paying damages is only a small part of the total cost of driving a car, we would expect changes in the damages rule to have only a small effect on how much people drive. If I expect to cause one accident for every hundred thousand miles, for instance, then reducing the damages I must pay in case of an accident from $1000 to $500, as in our earlier example, reduces the cost of driving by only half a cent per mile. If all costs of driving come to twenty-five cents a mile, then a fifty percent reduction in damages results in only a two percent reduction in driving costs. The effect would presumably be larger with regard to decisions that more directly effect the number of accidents, such as whether to drive after having a beer, or how fast to drive. But it still seems likely, in the case not only of driving, but of accidental torts in general, that the cost of paying damages for accidents is typically a small part of the total costs and benefits involved in the relevant decisions, and that the number of accidents would therefore typically, although not invariably, be only modestly affected by even quite substantial changes in the amount of damages awarded per accident. If so, the analysis of the previous section implies that the efficient level of damages for accidental torts is substantially less than the injury actually suffered by the victim.

Next, consider a deliberate tort—my slugging you. In contrast to the case of the automobile accident, the relevant decision—to throw a punch—is almost certain to result in the tort occurring and my paying damages.19 While that is not the only cost I will consider in deciding whether to start a fight, it is a very substantial one. It therefore seems likely that increases in the amount of damages awarded will have a substantial effect on my action. If so, that suggests a relatively elastic supply of offenses, and hence an argument for setting damages above the actual harm done.

One can make a similar argument in the case of a reckless tort. Describing a tort as reckless presumably means either that any reasonable man could see that what he was doing would result in

19. At this point, I am again assuming that all occurrences of torts result in successful suits. Dropping that assumption would lower expected cost in the case of both the accidental and the intentional tort.
injuring someone or that the cost of preventing the injury would be very low. In either case, one would expect the occurrence of the tort to be relatively sensitive to the amount of damages, so we again get an argument for setting damages above actual harm done.

If these arguments are correct, we would expect an efficient legal system to set expected damages substantially below harm done in the case of accidental torts and above harm done in the case of deliberate or reckless torts. Even if, in the latter case, the elasticity of supply of offenses is not high enough to justify damages above harm done, it still would be efficient for expected damages to be higher relative to harm than for accidental torts, as long as the corresponding elasticity was higher.

How well do these predictions fit the traditional common law of damages, including punitive damages? Under the traditional common law, damages for accidental torts were narrowly defined, excluding elements of injury that were difficult to measure, such as pain and suffering and the value of the victim’s life to himself. In addition, the damage payment was calculated without any adjustment for the probability that an offense would result in a successful action. If that probability was below one, as it surely was, the result would be to lower expected damages—damages times probability—below actual damages. On net, these two factors would make damage payments substantially lower than those implied by the rule “expected damage payment should equal injury done”—the equivalent, in a probabilistic sense, of “the tortfeasor should make the victim whole.”

Under the traditional common law, punitive damages were available for reckless or deliberate torts. The result would be to raise the expected damages above those in other cases, which, if my previous analysis is correct, corresponds to the efficient rule. Whether expected damages would be higher than actual injury would depend on how large the punitive damages actually were; whether expected damages should (from the standpoint of efficiency) be higher than actual damages would depend on how elastic the supply curve for reckless and deliberate torts happened to be.

An alternative way in which the legal system might deal with torts of varying elasticity would be to vary the amount paid by the tortfeasor but not the amount received by the victim. This would
correspond to a system of ordinary damages supplemented by fines.\textsuperscript{20} The earlier discussion of probability of conviction provides an explanation for why that may be a less efficient solution. The torts for which we want high damage payments are also the torts for which we want a high probability of conviction. We get it by giving the high damage payment to the victim, thus giving him a strong incentive to sue.\textsuperscript{21}

III. Is This the Best We Can Do?

The reader may reasonably ask why the law, if efficient, chooses such a clumsy device as punitive damages. Why not have a legal rule that makes damages depend explicitly on the elasticity of supply of offenses instead of trying to fit all torts into the two polar categories of “deliberate or reckless” and “all other”?\textsuperscript{22}

The answer brings us to what is, I believe, one of the most important factors determining the form of efficient law—the fact that courts are a very poor way of finding the correct answer to a difficult question. If you wish to diagnose an illness, design a computer, or discover a new scientific law, you do not do it by picking a dozen people at random, forming them into a committee, and demanding that they give you an answer. You do not even do it by picking one general-purpose expert and asking him.

Given the limitations of courts, it is sensible to try to avoid, so far as possible, asking them to do difficult things.\textsuperscript{23} I do not think that including this Article in the instructions to a jury in a civil case would be likely to result in a better, or even a more efficient, decision. The law should therefore, wherever practical, be stated in terms of simple rules, even if they only imperfectly reproduce the outcomes of the much more complicated correct rules. That is the

\textsuperscript{20} Such a system also might involve payments by the government to victims of torts for which the efficient damage payment was low.

\textsuperscript{21} Note that this is a qualitative, not a quantitative, argument. I have not shown that the efficient solution requires the same amount paid by the tortfeasor and received by the victim, only that both should be larger the more elastic the supply of the offense.

\textsuperscript{22} This seems to imply that there should be a relation between how good the mechanisms of a particular legal system are at solving problems and how hard the problems are that they are asked to solve, with specialized courts allowing more complicated legal rules. This point was suggested to me by Judge Posner.
justification for distinctions such as the one this Article attempts to explain.\textsuperscript{23}

What I believe I have shown is that the traditional common law of damages, including punitive damages, is consistent with the predictions of economic efficiency. I do not wish to claim more than that. I have certainly not shown that dividing torts into the two classes of "deliberate or reckless" and "all other" is the best way of approximating the efficient level of damages for each tort. Nor have I shown that the size of the observed damages, both punitive and ordinary, corresponds to the predictions of economic efficiency.

\textsuperscript{23} This point is relevant to many issues other than the one I have been discussing. One example occurs in Professor David Owen’s contribution to this issue. He justifies the idea that all humans are entitled to equal respect, and thus have equal rights, in part on the grounds that all humans have the capacity for rational thought and possess free will. \textit{See} Owen, \textit{The Moral Foundations of Punitive Damages}, 40 \textit{Ala. L. Rev.} 705, 709 (1989). This strikes me as a singularly unconvincing argument, tenable only because most of us are reluctant to come out against equal rights. Rationality, after all, is something that we all believe is possessed in unequal amounts by different people, although we sometimes may disagree about which of us are the rational ones. If rights are founded on rationality, that would seem to imply \textit{unequal} rights. So far as free will is concerned, the evidence available to me that my cats have free will is the same as the evidence that my colleagues do, yet I am confident that neither group would concede equal rights to the other.

A more plausible way to justify equal (legal) rights, I think, is to start with the fact that although we differ in our degree of rationality, there is no simple, objective way of measuring that difference. If courts are instructed to give different rights to people according to how rational they are, the result will be a great deal of very expensive litigation, with everyone trying to prove that he belongs at the top of the hierarchy of rationality, and hence of rights. Whether or not someone is a human being, on the other hand, is fairly easy to determine. We therefore instruct our courts to give the same rights to all humans, with a few exceptions for differences that are fairly easy to measure, such as age and the more extreme forms of mental incapacity.
Appendix: The Efficient Damage Rule

Consider a potential tortfeasor. The number of offenses he supplies (auto accidents for which he is responsible, for example) is a function \( O(C_p) \) of his expenditure on precautions \( C_p \). For each offense he is charged damages \( D \). The tortfeasor's cost is damage cost plus precaution cost: \( DO(C_p) + C_p \). To minimize this, he chooses \( C_p \) such that:

\[
\frac{d}{dC_p}[DO(C_p)+C_p]=D \frac{dO(C_p)}{dC_p}+1=0 \tag{Equation 1}
\]

The solution to this defines an implicit function \( C_p^*(D) \), the level of precaution chosen by a potential tortfeasor as a function of the damages he must pay for committing a tort.

Each offense imposes a cost on the victim; for purposes of simplicity I will assume that the cost per offense is constant. Defining it as \( C_v \), the total cost imposed by the tort on the victims is \( C_vO(C_p) \). If there were no litigation costs, social cost would be the cost to the victim plus the cost of precautions: \( C_vO(C_p) + C_p \). Minimizing, we would get:

\[
\frac{d}{dC_p}[C_vO(C_p)+C_p]=C_v \frac{dO(C_p)}{dC_p} + 1=0 \tag{Equation 2}
\]

Equation 2 describes what we want the potential tortfeasor to do; equation 1 describes what he will choose to do. We make the two the same by setting damages equal to the cost imposed by the tort: \( D=C_v \). This is the mathematical equivalent of the verbal argument in Part I.

We now introduce litigation costs. For simplicity, we assume that the litigation cost \( (C_L) \) is a fixed fraction \( (f_L) \) of the damages paid by the tortfeasor; \( C_L=f_L DO(C_p) \). Since \( D \) in equation 1 represented the cost of the damages to the tortfeasor, it includes his legal costs. Litigation cost is the sum of the legal costs of the two parties.\(^4\) If, for example, each party pays $100 in attorneys' fees and the court awards a judgement of $500, then \( D=600 \), \( C_L=200=f_L D \), so \( f_L=1/3 \).

Our objective is now to pick the level of \( D \) that minimizes social costs:

\[
SC=\text{Social Cost}=\text{Cost of precaution}+\text{Cost of Injury}+\text{Cost of Litigation}
\]

\[
= C_p^*(D)+C_vO(C_p^*(D))+f_LDO(C_p^*(D))
\]

\[
\frac{dSC}{dD}=0=\frac{dC_p^*(D)}{dD}+C_v \left\{ \frac{dO(C_p^*(D))}{dD} \right\}+f_L \left\{ O(C_p^*(D))+D \frac{dO(C_p^*(D))}{dD} \right\} \tag{Equation 3}
\]

\(^4\)For simplicity, I ignore other litigation costs, such as the value of the judge's time; including them would not affect my results.
We know, from Equation 1, that the tortfeasor will choose a level of precaution \( C_p^* \) for which \( \frac{dO(C_p)}{dC_p} = \frac{-1}{D} \). Using that fact, we have:

\[
\left\{ \frac{dO(C_p^*(D))}{dD} \right\} = \frac{dO(C_p^*)}{dC_p^*} \frac{dC_p^*}{dD} = \frac{-1}{D} \frac{dC_p^*}{dD}
\]

Hence:

\[
\frac{dC_p^*}{dD} = -D \left\{ \frac{dO(C_p^*(D))}{dD} \right\}
\]

Plugging that into Equation 3, we have:

\[
0 = -D \left\{ \frac{dO(C_p^*(D))}{dD} \right\} + C_v \left\{ \frac{dO(C_p^*(D))}{dD} \right\} + f_L \left[ O(C_p^*(D)) + D \left\{ \frac{dO(C_p^*(D))}{dD} \right\} \right]
\]

\[
D \left\{ \frac{dO(C_p^*(D))}{dD} \right\} [1-f_L] = C_v \left\{ \frac{dO(C_p^*(D))}{dD} \right\} + f_L 0
\]

\[
D[1-f_L] = C_v + f_L O(C_p^*(D)) \left\{ \frac{dO(C_p^*(D))}{dD} \right\} = C_v - f_L D/\epsilon
\]

Where \( \epsilon \) is the elasticity of the supply of offenses:

\[
\epsilon = - \frac{D}{O} \left\{ \frac{dO(C_p^*(D))}{dD} \right\}
\]

Solving for \( D \) gives us:

\[
D = C_v / \left[ 1 - f_L \left( 1 - \frac{1}{\epsilon} \right) \right]
\]

(Equation 4)

Looking at Equation 4, we observe several things. First, if either \( f_L = 0 \), corresponding to costless litigation, or \( \epsilon = 1 \), the equation simplifies to \( D = C_v \). We are back at our earlier result: damages are equal to injury, so the tortfeasor must make the victim whole.

If \( \epsilon > 1 \) (the supply of offenses is elastic), Equation 4 implies that \( D > C_v \); the efficient level of damages is greater than the actual injury. If \( \epsilon < 1 \) (the supply of offenses is inelastic), Equation 4 implies that \( D < C_v \); the efficient level of damages is less than the actual injury. That is, in a more precise form, the result derived in the verbal analysis of Part II.
Some readers may be curious as to how far from $C_v$ we may expect the efficient value of $D$ to be. It is sometimes asserted that, in civil litigation, defendants gain about half what plaintiffs lose. If so, $f_L$ is about $1/2$. If, at the optimal level of damages, a one percent increase in $D$ results in a two percent decrease in the number of offenses (elasticity = 2), then Equation 4 implies that $D = (4/3)C_v$. If a two percent increase in $D$ results in a one percent decrease in $O$ (elasticity = .5), then Equation 4 implies that $D = C_v/2$. If we assume, with no evidence at all, that the former case corresponds to the typical deliberate tort and the latter to the typical accidental tort, we conclude that the efficient damage payment for a deliberate tort is $8/3$ as large as for an accidental tort, corresponding to punitive damages for the former equal to $5/3$ of ordinary damages.

Some Further Complications

I have assumed, for purposes of simplicity, that litigation costs are proportional to damages paid out. If one drops that assumption, the result is to change the quantitative but not the qualitative conclusions. The efficient level of damages increases with elasticity of supply of offenses, but the elasticity at which damages are equal to injury need not be one. If, for instance, we assume that marginal litigation cost falls with increasing damages, so that large cases cost proportionally less to litigate than small cases, it is easy to show that if the supply of offenses is unit elastic, the efficient damage payment is more than the injury done by the offense.

A second simplifying assumption I have used is that the court knows both the injury imposed by the tort and the expected damages paid by the tortfeasor. If the court cannot accurately measure injury, or if it does not know the expected damage because it does not know the probability that a tortfeasor will be convicted, additional difficulties arise. At one extreme, this leads us to an explanation for punitive damages proposed by Landes and Posner. They argue that intentional torts are virtually certain to be inefficient, since the same objective could be achieved less expensively by a market transaction. If so, the occurrence of such torts is simply evidence that the court system is underestimating injury or overestimating (for some tortfeasors) probability of conviction. Their solution is to use punitive damages to make it less likely that court errors will lead to such inefficient torts occurring. It would be possible, but complicated, to integrate that approach with mine. The result would be a model that took account of both court error and ex ante estimates of the probability distribution of net costs (injury to victim minus gain to tortfeasor) for any particular tort. Such distributions are unnecessary in my model, since injury is known with certainty and gain to the tortfeasor is revealed perfectly by the supply curve for offenses.

---