Integrating Punishment and Efficiency Concerns in Punitive Damages for Reckless Disregard of Risks to Others

Peter Diamond
Massachusetts Institute of Technology

Justifications for the use of punitive damages refer to deterrence and punishment. After formulating a social welfare function that incorporates both economic efficiency and a desire for retribution, optimal punitive damages are considered to balance concerns for economic efficiency and for retribution. This optimal balancing is considered where compensatory damages alone provide the correct level of deterrence, allowing the ideal retribution to vary with the level of wealth and with the level of precaution. The analysis is extended to situations where some accidents do not result in liability.

1. Introduction

Both judicial and academic justifications for the use of punitive damages refer to deterrence and punishment as the two bases for determining their size (e.g., Ellis, 1982). The social interest in deterrence is linked to a concern for economic efficiency. Behavior potentially subjected to punitive damages may be economically inefficient otherwise, for example, when decisions are based on inadequate financial incentives, despite the presence of both compensatory damages and (possibly) civil and criminal punishments.[1] The social interest in punishment comes from a view that a balancing of outrageous behavior with punishment makes the outcome more socially desirable.[2] A desire to punish per se, to make the punishment fit the crime, a just-desert theory of damages has punishment as an end, not as a means to deterrence. Indeed, we can define outrageous behavior in terms of a social desire to have the defendant pay more than the costs that fell on the plaintiff. We will refer to this motivation as retribution.

These two sides of punitive damages are not separable. The assessment of punitive damages to have more deterrence is a form of punishment; the assessment of punitive damages as retribution is a further deterrent. This

I am indebted to Mark Liffman and Steve Shavell for comments on an earlier draft. This research was supported by a grant from Exxon Company, USA. The views expressed are those of the author and not necessarily those of Exxon.

1. On punitive damages and deterrence, see Diamond (1997) and Polinsky and Shavell (1998).

2. Punitive damages are only supposed to be assessed in response to outrageous behavior.

article explores an evaluation of the effects of punitive damages incorporating both efficiency and retribution concerns.[3]

One can dispute the appropriateness of using punitive damages for punishment (retribution) purposes rather than relying only on fines and criminal law; one can question the appropriateness of a punishment orientation without the procedural protections of criminal law; and one can question the setting of punishment by a jury process with nothing in the way of guidelines for suitable punishments for different examples of outrageous behavior.[4] Nevertheless, with continued use of punitive damages to punish, it seems worthwhile to explore the implications of accepting a jury's desire for retribution along with an interest in economic efficiency.

This article explores situations where reckless disregard,[5] without malicious intent, might be viewed as outrageous.[6] First, behavior involving risk might be considered outrageous even though the defendant was making a rational decision that reflected all social costs. While such a judgment seems unfair and is conducive to inefficiency, it might reflect public attitudes toward risks that are simply not consistent with efficiency considerations (Breyer, 1993). In this case, catering to such a public desire for retribution creates a tension between retribution and efficiency; any level of punitive damages in response to such an attitude lowers efficiency. The article derives the first-order condition for balancing retribution and efficiency in this setting.

Second, if there are inadequacies of compensatory damages, rational risk taking based on the costs borne by the defendant might reflect costs that are less than the full social costs. If some examples of such behavior are judged outrageous, then we can consider retribution along with the need for additional deterrence when defendants bear only part of accident costs. In this setting, punitive damages needed for economic efficiency may be larger or smaller than those desired for retribution. To analyze this issue the article uses two simpler concepts of levels of punitive damages. When considering only retribution, there is some level of punitive damages, referred to as the ideal retribution level, that would be the best level considering only issues of retribution. Similarly, when considering only deterrence, there is some level of punitive damages, referred to as the ideal deterrence level, that would

3. Polinsky and Shavell (2000) use the approach developed here to analyze further issues.

4. On the difficulty juries have expressing outrage in monetary terms, see Kahneman, Schkade, and Sunstein (1998) and Sunstein, Kahneman, and Schkade (1998).

5. Outrageous behavior without malicious intent is referred to by a variety of (not fully interchangeable) terms such as reckless or callous disregard or reckless indifference to the rights of others, gross negligence, and legal malice.

6. This article formally models situations involving individual defendants and accidents among strangers. The issues analyzed here are relevant for other situations, but additional issues may arise. For example, analysis of accidents involving people who have a contractual relationship (e.g., products liability) also needs to consider how punitive damages affect the prices charged. Insofar as a corporation can be considered as a unitary decision maker, the analysis in this article is applicable. For example, the analysis holds for a corporation with a general policy of polluting. But the analysis is missing elements if some workers pollute without the knowledge of their supervisors.

be the best level considering only issues of economic efficiency. These two levels might coincide, or they might be different, with either one larger than the other. Optimal punitive damages reflecting both a desire for retribution and a concern for economic efficiency might be intermediate or might be higher than either of the levels that consider one issue alone. The article does not formally consider the possibility of wrongly assessing punitive damages, a possibility that would lower the optimum level.

Third, reckless disregard might be based on nonrational thinking. Some drunk driving is an example, where some risks are simply ignored, even though their existence is known.[7] However, this article only considers rational behavior by defendants.

Section 2 presents the notation of defendant choice. Section 3 formulates a social welfare function that incorporates both economic efficiency and a desire for retribution. Section 4 considers the optimal balancing of deterrence and retribution in a situation where compensatory damages alone provide the correct level of deterrence, with mathematical analysis in Appendix A. Sections 5 and 6 extend the analysis to consider the assumptions that the ideal retribution varies with the level of wealth and with the level of precaution (with derivation in Appendix B). The analysis in Section 4 is then extended to situations where some accidents do not result in liability (Section 7, a derivation in Appendix C). Section 8 concludes.

2. Precaution and Accidents

It is convenient to consider the probability of an accident as a control variable of the defendant (within limits).[8] Let π equal the probability of *avoiding* an accident, where π defines the level of *precaution* being taken by the defendant. Let $u[\pi]$ be the utility of the activity, net of costs of the activity, including the cost of precaution, but gross of any legal liability of the defendant.[9] Recognizing both the costs to the defendant of any accident and the costs of avoiding accidents, we assume that u is strictly concave, $u''[\pi] < 0$, first increasing, then decreasing in π. We also assume that the cost of avoiding accidents rises without limit as the probability of avoiding all accidents increases toward its upper limit, which is less than one. In Figure 1 we show a typical pattern of utility gross of legal costs relative to precaution. We assume that the defendant is risk neutral and so maximizes $u[\pi]$ less expected liability for compensatory damages plus fines plus punitive damages. This gives the chosen level of precaution as a function of these expected liability costs. We assume that precaution is increasing and concave

7. For an analysis of deterrence with nonrational thinking, see Diamond (1997).

8. If an explicit care or precaution variable is used, both the cost of care and probability of an accident are usually assumed to be monotonic in this variable. Using a probability as a control variable simplifies the notation, since there is a functional link between the cost of care and the probability of an accident.

9. Arguments of functions will be denoted by [], as in $u[\pi]$. Parentheses will be denoted by () and {}.

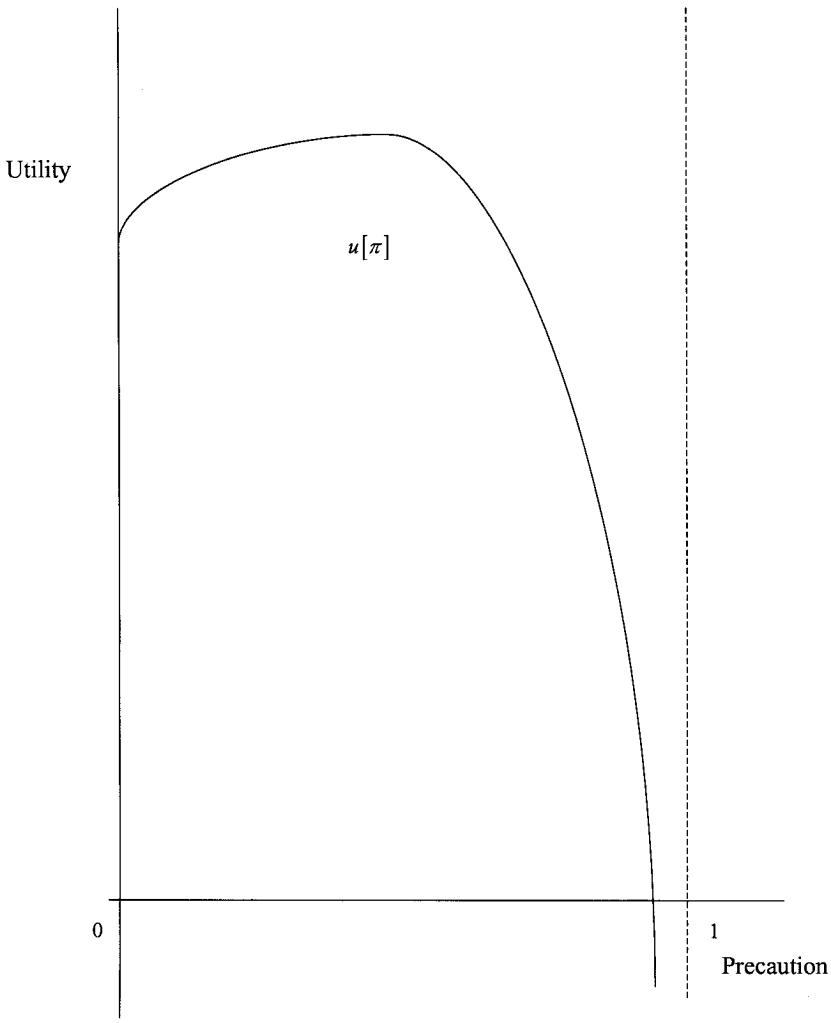

Figure 1. Utility as a function of precaution.

in expected costs. That is, optimal precaution is increasing in expected liability, but at a decreasing rate. This follows from the further assumption that $u'''[\pi] < 0$ (see Appendix A).[10]

10. For example, there might be a level of utility from the activity if there is no accident, an expected cost of accidents that is proportional to the probability of an accident, and a cost of avoiding accidents that is unbounded as the probability of an accident goes to zero. Then, $u[\pi]$ might have the form $u[\pi] = k_0 - k_1(1-\pi) - k_2(1-\pi)^{-1}$ for some positive constants k_i. With this utility function, we have the derivatives $u'[\pi] = k_1 - k_2(1-\pi)^{-2}$, $u''[\pi] = -2k_2(1-\pi)^{-3}$, and $u'''[\pi] = -6k_2(1-\pi)^{-4}$.

Denote by A the cost to the plaintiff in the event of an accident, including noneconomic costs.[11] We assume that the defendant is also risk neutral. If the social evaluation of the level of precaution were strictly utilitarian, denoted in monetary terms, the social value, W, of equilibrium with a level of precaution of π would satisfy

$$W = u[\pi] - (1 - \pi)A. \tag{1}$$

Thus the social evaluation recognizes the utility of the defendant, $u[\pi]$, and the expected accident costs of the plaintiff, $(1-\pi)A$.[12] The payment of damages by the defendant to the plaintiff is viewed as a transfer without direct social significance. For convenience, legal costs are taken to be zero.[13] If expected liability costs equal expected accident costs, then (in the absence of a need for incentives for victims) a rational decision maker will make an efficient choice [e.g., Cooter and Ulen (1997:272–6)]. For analytical convenience, we assume strict liability, not negligence, since any defendant at serious risk for being found liable for punitive damages because of reckless disregard is very likely to be held negligent and thus liable for compensatory damages if there is a negligence standard.

Implicit in this formulation of social welfare is an absence of externalities other than the possibility of an accident. If the activity of the defendant has social values that are different from the gain to the defendant less the accident costs, then this further deviation of individual and social values must be considered. For example, innovations in product and in technology are generally viewed as having externalities, since they affect the opportunities of others in ways that are not captured by market transactions.[14] In addition, the undertaking of large projects can involve consumer surplus, which again is not captured by market transactions.[15] To the extent that the activities deterred involve positive externalities, then the social concern for overdeterrence becomes larger. "Reputation costs" raise a similar issue. Conversely, when deterring accident-generating behavior also deters other negative externalities, then there is less concern about overdeterrence.

11. More generally, we could also allow accident costs, A, to vary with the level of precaution, but this would not alter the shape of the conclusions, given suitable assumptions on this interaction.

12. This approach ignores all issues of income distribution by adding up individual utilities in dollar terms. For a defense of this approach, see Shavell (1981).

13. This assumption makes the analysis easier to follow. Moreover, the effects of legal costs have been studied in the literature. The costs of the defendant are an additional deterrent to accident-generating behavior, although one that is a social cost of the accident. The legal costs of the plaintiff are an additional social cost of the accident. The effect of such costs on the analysis would pay attention particularly to the frequency of litigation with and without punitive damages and the advantages and disadvantages of different amounts of litigation.

14. On the impact of liability on innovation, see Huber and Litan (1991).

15. On the measurement of consumer surplus, see Diamond and McFadden (1974).

3. Social Evaluation of Punishment

Court decisions involving punitive damages refer to a desire to punish per se, to use punitive damages to inflict a cost on the defendant, to have the punishment fit the crime. In this section we present an approach to incorporating the desire for punishment along with concern for economic efficiency.

In recognizing both deterrence and a desire to punish, the motivation for punitive damages parallels that of the criminal law. For example, consider this statement by Rawls (1955:4–5):

> For our purposes we may say that there are two justifications of punishment. What we may call the retributive view is that punishment is justified on the grounds that the wrongdoing merits punishment. It is morally fitting that a person who does wrong should suffer in proportion to his wrongdoing. That a criminal should be punished follows from his guilt, and the severity of the appropriate punishment depends on the depravity of the act. The state of affairs where a wrongdoer suffers punishment is morally better than the state of affairs where he does not; and it is better irrespective of any of the consequences of punishing him.
>
> What we may call the utilitarian view holds that on the principle that bygones are bygones and that only future consequences are material to present decisions, punishment is justifiable only by reference to the probable consequences of maintaining it as one of the devices of the social order. Wrongs committed in the past are, as such, not relevant considerations for deciding what to do. If punishment can be shown to promote effectively the interest of society it is justifiable, otherwise it is not.

That there is a tension between these two justifications for punishment is well recognized in the literature. For example, chapter 1 in Packer (1968) is entitled "The Dilemma of Punishment;" Goldman (1979) writes of the "paradox of punishment." Some of the philosophical debate is about which of these two justifications is appropriate; some (such as Packer) argue for the relevance of both justifications in deciding when to punish. In contrast with the qualitative relevance of both justifications, Goldman identifies the dilemma in quantitative terms: that the level of punishment for criminal activity that he considers just (and so justified) is sufficiently low as a deterrent to crime as to be unpalatable; conversely, that the level he thinks needed for deterrence is too large to be just. The purpose of this section is to integrate a concern for retribution into a formulation of social welfare that also has a concern for efficiency in a form that leads to a quantitative outcome.[16] In other words, if

16. There is a need to consider punishment along with utilitarian concerns to make sense of having punishment vary monotonically with the seriousness of the crime. Optimal tax theory, which shows the importance for taxes of elasticities of demand, implies that monotonicity will

one considers only deterrence or only retribution, it is likely that the levels of punitive damages seen as appropriate for the two single purposes would be different from each other. When this is the case, the question might be how to choose between the two bases for setting punitive damages, or as analyzed here, how to compromise between the two levels suggested by the two bases. While it would appear that Goldman would not approve of grasping both horns of this dilemma and smoothly trading off the costs of improper deterrence and unjust retribution, that is the approach taken here. The focus here is not on the appropriateness of making this shotgun marriage of concerns, but on exploring the implications of having such a marriage. Nor do we explore the appropriateness of accepting a desire to punish; the starting place is to incorporate both retribution and deterrence concerns. As stated by Wertheimer (1975:420): "It seems reasonable to want a punishment to 'fit the crime,' and it seems reasonable to use that punishment that will maximize utility." Or, as stated by Posner (1975:778): "since no rational society can ignore the costs of its public policies, they are issues to which economics has great relevance. The demand for justice is not independent of its price."

In defining social welfare, we assume that the payment of punitive damages is a transfer that has no direct welfare implications as a transfer, but does as part of retribution (as well as affecting deterrence). That is, we assume that if an accident has occurred which is judged to have been caused by outrageous behavior, then there is a social desire to inflict a punishment. We refer to the level of punishment desired for retribution as ideal retribution and denote it by P^R. We assume that P^R is at least as large as the accident costs caused by the defendant, A. Indeed, we can view this as a definition of outrageous behavior. With an accident that does not follow from outrageous behavior, payment of compensatory damages is viewed as sufficient punishment. With an accident caused by outrageous behavior, we have the condition $P^R > A$. We do not analyze issues that might enter into the determination of the ideal retribution level; we merely assume an ideal retribution level for this class of defendants, one that does not vary with the precaution taken. Below we consider having ideal retribution vary with the precaution of a typical defendant and with the wealth of the defendant.

We assume that inflicting a level of punishment different from the ideal retribution involves a social cost. To integrate this social cost with economic efficiency concerns, we need to answer several questions. In considering the punishment of defendants, there are payments of punitive damages and of compensatory damages and often also of civil and criminal fines—do we consider the sum or only some of these payments to compare with the ideal retribution? We take the answer to be that the sum of all damages and fines

---

not necessarily follow from utilitarian concerns. This disagrees with assertions to the contrary, such as that of Rawls (1955). In note 14 (pp. 12–13), Rawls considers the proportionality of punishment to offense: "if utilitarian considerations are followed, penalties will be proportional to offenses in this sense: the order of offenses according to seriousness can be paired off with the order of penalties according to severity."

should be compared with P^R, although for the analytical purposes of this article, the distinction is not important, since we do not vary the cost of accidents or the level of fines.[17] If civil and criminal fines have been set to fully capture the desire for retribution, then P^R will be equal to the sum of accident costs and civil and criminal fines. We recognize the possibility that the desire for retribution exceeds the level of accident costs plus fines, leaving room for punitive damages to improve the satisfaction of the desire for retribution. We do not examine why legislated fines might be below such a level in some situations, nor the appropriateness of leaving the decision to punish by more than the legislated amount to a jury.[18]

We write the cost of incorrect punishment as a function of the difference between ideal retribution and the sum of compensatory damages, civil and criminal fines, and punitive damages, with a social cost from incorrect punishment whenever the sum of payments does not equal the ideal retribution. In units comparable to utility, we assume the existence of a smooth function giving this social cost.[19] There can be different causes of a deviation of actual punishment from ideal retribution—causes that include jury error in determining liability for punitive damages, jury error in setting a level of punitive damages, or a conscious social decision to deviate from ideal retribution because of possible economic inefficiency from levying ideal retribution. We do not consider having different functions for different causes, although one could extend the analysis in that way.

Denote by C the sum of compensatory damages and civil and criminal fines paid and by P the level of punitive damages paid. We write the social cost, measured in units of individual utility, from having levied a punishment different from the ideal retribution as $V[C+P-P^R]$, with $V[0]=0$. Since this is the smallest value of V, we also have $V'[0]=0$. We assume that V is concave, $V''>0$.

We define social welfare as the utility of the plaintiff plus the utility of the defendant less this social cost, V.[20] Thus when all accidents result in liability, we write social welfare as

$$W = u[\pi] - (1-\pi)A - (1-\pi)V\left[C+P-P^R\right]. \tag{2}$$

With this formulation, concern about accidents involves only actual accident costs, provided that the decision makers responsible for the accidents are

17. This view is not new. Morris (1931:1188) argued that the sum of compensatory and punitive damages should be viewed as punishment and applauded the fact that West Virginia juries are instructed to think in these terms (p. 1188). He also called for "reciprocal adjustment of penalties of the civil and criminal courts" (p. 1197).

18. This analysis takes the level of fines as given. An alternative interpretation of the model is an analysis of the optimal level of fines.

19. This issue is also relevant in the context of taxation. One can analyze how taxation should differ from ideal taxation to reflect the deadweight burdens of taxation as well as the administrative costs of trying to get taxes "just right."

20. This formulation does not recognize any social significance in the overcompensation of the plaintiff if the sum of punitive and compensatory damages exceeds accident costs.

properly punished. If they are not properly punished, then there is an added cost to any accident, a cost from inappropriate punishment, whether that punishment is too large or too small. With this structure we are only concerned with behavior that actually results in an accident, not behavior that might have resulted in an accident (no harm, no foul).[21]

While malicious intent is a natural source for outrage, this article considers only settings of reckless disregard that are considered outrageous. While subjecting others to risk is an everyday occurrence, sometimes it is viewed as outrageous. In the situation analyzed in Section 4, behavior involving risk is considered outrageous even though the defendant was making a rational decision that reflected all social costs. While such a judgment about risk taking seems unfair and is conducive to inefficiency, it might reflect hindsight bias or public attitudes toward risk that are simply not consistent with efficiency considerations.[22] For example, Breyer (1993) has argued that inconsistencies in public attitudes toward health and safety risks have contributed to inefficient regulation of risks. In this situation we have a tension between retribution and efficiency. That is, any level of punitive damages in response to such an attitude lowers efficiency. If the legal system chooses to go along with this desire to punish, then the analysis can be interpreted as the appropriate balancing between respecting that desire of juries and lowering economic efficiency.[23]

If expected compensatory damages plus fines are less than expected accident costs, with risk taking based on the costs borne by the defendant, the defendant's precaution decision is not efficient in the absence of punitive damages. If ignoring social costs that are not paid is judged outrageous, then we can combine retribution with efficiency concerns coming from only partial bearing of accident costs. Such a model is explored in Section 7 in a setting where ideal retribution might be larger or smaller than the liability level needed for efficiency. A different category of reckless disregard is

21. Implicit in Equation (2) is the idea that when punishment is at the ideal level for retribution and if that punishment induces economic efficiency, then there is no further concern with the occurrence of the behavior that was deemed to be in need of punishment. One might well argue for further deterrence to decrease the occurrence of outrageous behavior in response to a further dislike of such behavior, even when that behavior is ideally punished. Alternatively, one might consider economic efficiency and desired retribution as the only suitable bases for social evaluation. The formulation above is usable with either interpretation by interpreting the accident costs as either the costs that are suitable for economic efficiency or those costs plus any additional concern for the behavior that led to the occurrence of the accident. The text will be based on the interpretation that A is actual accident costs, without any additional outrage adjustment, consistent with linking deterrence to economic efficiency and punishment to retribution. The mathematics can be interpreted as fitting the alternative interpretation.

22. On hindsight bias, see Fischhoff (1982).

23. Another model of behavior that might be judged as reckless disregard is where the information used by the defendant in decision making was judged to be inadequate even though there was a rational decision as to how much information to gather. As argued in Shavell (1992), this behavior implies efficiency as long as decision makers are rational in seeking information and have appropriate prior beliefs. For this situation, the analysis in Section 4 again applies.

that of nonrational reckless disregard. Some drunk driving is an example of such behavior, where some risks are simply ignored, even though their existence is known. While deterrence of such behavior is examined in Diamond (1997), the combination of deterrence along with a desire for retribution is not analyzed in this article.

4. Balancing Deterrence and Retribution

To see the basic interaction between deterrence and retribution, it is helpful to begin with a simple situation where the defendant pays all of the accident costs, denoted by A, as well as paying civil or criminal penalties, if any, C-A, and punitive damages, P. That is, we assume that the utility of the defendant is

$$U[\pi] = u[\pi] - (1-\pi)(C+P). \tag{3}$$

The level of precaution chosen to optimize $U[\pi]$ in Equation (3) can be written as a function of total payments in the event of an accident, $\pi[C+P]$. With $u'' < 0$ and $u''' < 0$, the level of precaution increases with damages, but at a decreasing rate, $\pi' > 0$, $\pi'' < 0$.

In this setting, maximizing the sum of the utilities, $u[\pi] - (1-\pi)A$, would call for both fines and punitive damages to equal zero. In contrast, to minimize $V[C+P-P^R]$ we would set punitive damages equal to the ideal retribution level less compensatory damages and fines paid, $P^R - C$, if this difference is positive. To characterize the optimal level of punitive damages we can calculate a first-order condition for the maximization of social welfare, Equation (2), taking as given that individuals set precaution to maximize utility, Equation (3). Rewriting the social objective function to incorporate defendant behavior, we have

$$W[P] = u[\pi[C+P]] - (1-\pi[C+P])\big(A+V\big[C+P-P^R\big]\big). \tag{4}$$

We assume that W is concave in P, avoiding issues of multiple solutions to first-order conditions.

In the situation where fines already accomplish retribution, $P^R \leq C$, W is decreased by imposing punitive damages and so punitive damages should be zero. In the situation where fines are not large enough for ideal retribution, $P^R > C$ and so punitive damages can increase social welfare, the first-order condition for optimal punitive damages is

$$(1-\pi)V' - V\pi' = (A-P-C)\pi'. \tag{5}$$

This formulation brings out the two elements that are balanced in achieving an interior optimum. When total damages plus fines are less than ideal retribution, an increase in punitive damages lowers the expected cost of inadequate punishment, $(1-\pi)V$. This marginal gain from increased retribution equals the probability of an accident, $1-\pi$, times the marginal

social disutility of inadequate punishment, V', plus the decrease in accident probability, $-\pi'$, times the social disutility of inadequate punishment after an accident, V. However, the decrease in accident probability induced by increased punitive damages lowers efficiency by increasing overdeterrence. The marginal efficiency cost of this overdeterrence equals the extent to which punitive damages plus fines exceed the level just necessary to induce efficient deterrence. Since, in the situation being analyzed, there is no need for punitive damages or fines to induce efficient deterrence, the marginal cost of overdeterrence is the full amount of the punitive damages plus fines. That is, punitive damages and fines work like distortionary taxes.[24] The optimum occurs where the gain from having a more appropriate punishment when accidents do occur is just balanced with the two effects from inducing more precaution—a lowering of the economic efficiency of precaution, and a lowering of the frequency of inadequately punished defendants.

If $C < P^R$, there is an optimum with positive punitive damages. The optimum, which is the solution to Equation (5), satisfies the following inequalities (see Appendix A):

$$A + V\left[C + P - P^R\right] < P + C < P^R. \tag{6}$$

That is, at the optimum the social cost of an accident, including both the resource costs, A, and the social costs of inadequate punishment after an accident, V, is less than the payment of the defendant, $P + C$, which, in turn, is less than the level of ideal retribution, P^R. In this situation, where punitive damages are not needed for deterrence, accepting a desire to punish implies an average of the level desired for retribution and the zero level appropriate for deterrence. From the second inequality in Equation (6), at the optimum we have $V'[C + P - P^R] < 0$. For example, consider a case where accident costs (A) are 100 and there is a desire to punish which doubles the size of accident costs, so that ideal retribution (P^R) is 200. Then punitive damages plus fines should be set below 100, the level that would achieve ideal retribution (given compensatory damages of 100). How much below this level would be optimal depends on the shape of the cost of inadequate retribution and the response of precaution to expected costs. If punishing at 150 when the ideal is 200 generates a social cost of inadequate punishment of less than 50, then we would satisfy the inequalities above. If the social cost of too little retribution is larger than the sum of punitive damages and fines, then the latter should be larger.

24. Marginal individual return to precaution is $u' + C + P$. The economic side of social welfare has a marginal return to precaution, $u' + A$. Thus individuals take too much precaution relative to the standard for economic efficiency, with $A - P - C$ being equal to the amount that the social return to precaution exceeds the private return.

5. Punishment and Wealth

Without exploring whether ideal retribution should vary with the wealth of the defendant, we explore how punitive damages should vary with wealth if one assumes that ideal retribution is proportional to wealth. This situation is chosen for analysis because plaintiffs sometimes call for punitive damages proportional to the wealth of defendants. Thus we assume that ideal punishment is written as $P^R = P^r Y$, where Y is the measure of wealth relevant for describing ideal retribution. For each value of Y, we have the first-order condition in the general form of Equation (5). The goal is to examine the solutions to these first-order conditions as a function of Y. While retribution has the assumed proportionality to wealth, precautionary behavior depends on the absolute level of liability costs, not a relative level (assuming that u is the same at all wealth levels). Thus one might expect that punishment, $C+P$, should increase with wealth, but less than in proportion to wealth. While this may be the result, additional assumptions are needed for the conclusion. Such complexity is common in second-best settings.

In addition to scaling ideal retribution for wealth, we need to scale the evaluation of deviations from the ideal. Two simple situations are to evaluate these deviations in absolute terms and in proportional terms. If we scale deviations in proportional terms, then we would write V as a function of $[(C+P-P^rY)/Y]$. On the other hand, if we scale deviations in absolute terms, then we would write V as a function of $[C+P-P^rY]$. Using both of these formulations, we see how punitive damages should vary with the wealth of the defendant if one assumes proportionality in ideal retribution.

Since wealth is observable, we have the same first-order condition for optimizing social welfare as above for each level of wealth. Thus Equations (5) and (6) continue to hold. To analyze the pattern, we differentiate the first-order condition to relate P to Y. With the evaluation of deviations in proportional terms, it does not necessarily follow that optimal punitive damages should increase with the wealth of the defendant, even though ideal retribution is assumed to be proportional to wealth. A sufficient condition for such an increase is that the elasticity of accident probability with respect to expected liability costs $(-(C+P)\pi'/(1-\pi))$ be less than one. With evaluation of deviations in absolute terms, no additional conditions are needed for the conclusion that punitive damages should increase with the wealth of the defendant.

With evaluation of deviations in proportional terms, when punitive damages should increase with wealth, a sufficient condition for total liability, $C+P$, to increase less than in proportion to wealth is that π'' not be too large. This can be expressed alternatively in terms of u''' being small. With evaluation of deviations in absolute terms, small π'' and total liability being close to ideal retribution are sufficient for a less than proportional increase

in total liability. However, a great curvature in the precaution being chosen by defendants of some wealth level might reverse this conclusion.[25]

6. Ideal Retribution Varying with Precaution and Individual Precaution Unobservable

In the analysis above, the ideal retribution was based on the occurrence of an accident. No adjustment was made either for the typical level of precaution of all defendants or the actual level of precaution of this particular defendant. We pursue these two ideas in this section. For simplicity, we continue to assume full bearing of accident costs and $P^R > C$.

First, we assume that the precaution of the particular defendant is not observable by the jury setting punitive damages. Thus the level of punitive damages, P, does not vary with the precaution of the particular defendant and the precaution of the defendant, $\pi[C+P]$, is the same function as above. But we assume that the jury does know the typical level of precaution and has an ideal retribution that varies with the typical level of precaution, denoted $P^R[\pi]$. Thus we rewrite social welfare as

$$W[P] = u[\pi[C+P]] - (1-\pi[C+P]) \times \{A+V[C+P-P^R[\pi[C+P]]]\}. \tag{7}$$

We continue to assume that social welfare is a concave function of punitive damages.

Maximizing social welfare, as given in Equation (7), assuming that precaution is chosen to maximize defendant utility, given in Equation (3), we obtain a first-order condition that has one more term than previously. In place of Equation (5), we now have

$$(1-\pi)V'(1-P^{R\prime}\pi') - V\pi' = (A-P-C)\pi'. \tag{8}$$

Previously the expression $(1-\pi)V'$ was multiplied by 1; now it is multiplied by $(1-P^{R\prime}\pi')$, which is more than 1 since $P^{R\prime}$ is negative, while π' is positive. That is, since a higher level of punitive damages increases precaution, it decreases the ideal retribution, thereby closing some of the gap between actual and ideal punishment levels. It remains true that compensatory damages plus fines plus the optimal level of punitive plus compensatory damages plus fines lies below the ideal retribution and above the social costs of an accident including inappropriate punishment, $A+V$.

To see the effect of this change in the social welfare function, we compare optimal punitive damages when ideal retribution does and does not vary with precaution. Let us denote the optimal level of punitive damages when ideal retribution does not vary by P_0. That is, P_0 is the solution to Equation (5)

25. There would also be an efficiency cost, presumably small, from implicitly taxing wealth, since the same activity results in higher damages for people who saved more or earn more.

when ideal retribution is P_0^R. With this level of punitive damages, the equilibrium level of precaution is $\pi[C + P_0]$. In order to make the comparison interesting, we want to select a function for ideal retribution that passes through the same value; that is, that has $P^R[\pi[C + P_0]] = P_0^R$. Now we can ask whether it is better to raise or lower the level of punitive damages when the ideal retribution has changed from a scalar to this function of the precaution level. In other words, we want to evaluate the derivative of W in Equation (7), evaluated at the point where Equation (5) holds. Evaluating the derivative of welfare at this level, the derivative is not equal to zero, but is equal to $(1 - \pi)V'P^{R\prime}\pi'$. This expression is positive, implying that it would be good to increase the level of punitive damages. In other words, when the ideal level of retribution decreases with the level of precaution, the optimal level of punitive damages is more than it would be if the ideal level did not vary, assuming that the ideal function passes through the previous equilibrium.

The analysis would become much more complicated if we assumed that the defendant's precaution were observable. In this situation, the juries collectively are not selecting a typical level of punitive damages, P. Rather the juries collectively are now selecting a level of punitive damages that varies with the actual precaution of the defendant, $P[\pi]$. When choosing precaution, the defendant would now pay attention to the fact that punitive damages will be lower in the event of an accident should a higher level of precaution be chosen. Thus the first-order condition for precaution depends not only on the level of punitive damages, but also on how they vary with precaution. If all defendants were the same, it would be easy to manipulate the threatened increase in punitive damages in order to encourage precaution cheaply (from an efficiency point of view). Thus to analyze this model in an interesting fashion, we would need to consider unobservable heterogeneous defendant characteristics that result in a distribution of actual choices of precaution. We would then be selecting punitive damages as a function of precaution in order to do this maximization. The formal analysis would use the techniques developed by Mirrlees (1971) in his analysis of income taxation. Such an analysis is sufficiently complex that it is not considered here.

7. Underassessment of Compensatory Damages[26]

We now extend the analysis of Section 4 to the setting where some punitive damages increase efficiency because compensatory damages plus fines are too small for economic efficiency. Assume that compensatory damages are assessed after a fraction a of accidents, with compensatory damages equal to the fraction b of accident costs. Assume that in all of the accidents that result in compensatory damages, fines, F, and punitive damages, P, are assessed

26. Some analyses of criminal law use a similar model to the one employed here. Thus a concern about appropriate punishment will move optimal punishment away from the solution where sanctions are maximized and enforcement is minimized.

as well. We now denote by C the sum of compensatory damages and fines paid, $C = bA + F$.

The behavior of defendants is described by the maximization of individual utility, which is given by a modification of Equation (3) to reflect the probability a of assessment of damages and fines:

$$U[\pi] = u[\pi] - a(1-\pi)(C+P). \tag{9}$$

Given our assumption of risk neutrality on the part of defendants, the precaution decision is being made in response to expected costs $a(C+P)$ and can be written as $\pi[a(C+P)]$.

We must modify the social welfare function [as given in Equation (4)] to reflect the assumption that after some accidents no defendant is held liable. There is no punishment after the fraction $(1-a)$ of accidents, accidents that warrant retribution at the level P^R, including the payment of compensatory damages and fines. We assume that the social concern for this failure to hold the responsible party liable is measured by the same function that assesses inadequate punishment with liability. Thus when some accidents are caused by defendants who escape liability, the expected cost of inappropriate punishment changes from $\{(1-\pi)V[C+P-P^R]\}$ to $a(1-\pi)V[C+P-P^R] + (1-a)(1-\pi)V[-P^R]$. With a desire for retribution $P^R > 0$, accidents without liability lower social welfare by more than the same accidents would if the defendant paid damages, unless punishment is well in excess of desired retribution. This additional concern increases the social gain from deterring accidents, thereby generating a payoff to deterrence beyond the level that induces economic efficiency. Note that the assumption that some accident generation is escaping punishment is not being viewed as a reason to lower the ideal retribution on those who do pay damages.

The social welfare function now becomes

$$W[P] = u[\pi[a(C+P)]] - (1-\pi[a(C+P)]) \times \left(A + aV\left[C+P-P^R\right] + (1-a)V\left[-P^R\right]\right). \tag{10}$$

Note that this formulation includes both situations where there is a desire for retribution beyond the needs of deterrence and situations where, because of the poor workings of compensatory damages, ideal deterrence needs would call for a level of punitive damages beyond what would be ideal for retribution.

As above, we want to maximize social welfare, given in Equation (10), subject to the constraint that defendants maximize utility, as given in Equation (9). The first-order condition for optimal punitive damages differs in several ways from that given above in Equation (5). As derived in Appendix C, the first-order condition is now written as

$$(1-\pi)V'\left[C+P-P^R\right] - \left\{aV\left[C+P-P^R\right] + (1-a)V\left[-P^R\right]\right\}\pi' = (A - a(P+C))\pi'. \tag{11}$$

This expression matches that in Equation (5) when a and b are both equal to one. Increasing punitive damages changes the level of inappropriate punishment and increases the level of precaution. Both of these effects work only through the fraction a of accidents that result in liability, a multiplicative factor that cancels in Equation (11). The marginal social value of increased punishment for those held liable is $(1-\pi)V'[C+P-P^R]$. The marginal social value of increased precaution depends on the efficiency effect, which is the excess of the private over the social return to precaution, $\{a(C+P)-A\} = \{a(F+bA+P)-A\} = \{a(F+P)-(1-ab)A\}$, which can be positive or negative. It also depends on the avoidance of accidents that lead to inappropriate punishment, from $C+P$ not equaling P^R for those who are held liable and from the lack of liability for the rest.

We have modeled two ways in which compensatory damages can be inadequate. Compensatory damages might be less than accident costs. However, fines are assumed to be assessed on all defendants who pay compensatory damages. Thus fines can perfectly offset the economic effects of compensatory damages that are too low.[27] Moreover, fines and compensatory damages enter the evaluation of the level of retribution in the same way. However, when some people who cause accidents escape all liability, the situation is more complicated. While increasing fines can offset the inadequate ex ante incentive for precaution, it does not enter the evaluation of retribution in a perfectly offsetting way. In particular, because of inadequate retribution for those escaping liability, there remains a social gain from additional deterrence even if fines are set at a level that produces economic efficiency and happens to just match the desired level of retribution of those who pay the fines. To see this point we examine the ideal deterrent and retributive levels of punishment and note that optimal punishment is larger than the smaller of these and larger than both of them when they are equal.

The ideal deterrence level of punitive damages is the level that would be set if there were no concern with punishment per se. To induce efficient precaution, compensatory and punitive damages and fines would equate expected payments with expected accident costs to others. Thus the ideal deterrent, denoted by P^D, satisfies

$$P^D = A/a. \tag{12}$$

The ideal level of punitive damages for deterrence (when positive) is equal to $P^D - C$.

A central question is whether the ideal retribution, P^R, is larger or smaller than the ideal deterrent, P^D. The probability of liability, a, is central to this comparison. As the probability of liability gets small, P^D rises without limit, eventually exceeding the ideal retribution. As the probability of liability goes to one, P^D goes to the level of accident costs caused, which we have assumed

27. If $a = 1$, then having $F = (1-b)A$ removes the difference between private and social economic incentives for precaution.

is less than the ideal retribution, on the grounds that people should at least pay for the accident costs caused. When there are no fines and compensatory damages are adequate, ideal punitive damages for deterrence tend to zero, becoming less than ideal punitive damages for retribution.

As shown in Appendix C, the optimal level of punitive damages is larger than the minimum of these two ideal punishments, less the other costs borne, $P^D - C$ and $P^R - C$, and may even be larger than both of them. Indeed, if P^D and P^R happen to coincide, then the optimal level of punitive damages is larger than both. The pressure for more deterrence comes from the social disutility of the lack of punishment after accidents that result in no payment of damages. Larger punitive damages deter such accidents as long as there is a risk of being held liable for punitive damages. Thus having potential defendants escape liability adds to the importance of deterrence. Also, optimal punitive damages might be zero if C is large enough.

The structure of results here is strongly influenced by the assumption that all potential defendants have the same probability of being held liable. This assumption permits a greater punishment when liable to substitute for a smaller probability of being held liable. With variation in the population in the probability of being held liable, the results would change. This issue, which is also important for the theory of criminal penalties, would be interesting to explore.

We have considered punishment in a setting where some people escape liability, but no people are wrongly assessed damages. In settings where the latter is a possibility, then the social welfare function needs to be adjusted to recognize the probability of such errors, leading to a change in the first-order condition for optimal damages.

8. Concluding Remarks

Analyses like this one, which derive how to optimize the incentives coming from the legal system, have two goals in mind. First is to provide a framework for analysts. Without a derivation of what punitive damages would ideally accomplish, it is hard to see how analysts can evaluate the actual workings of punitive damages. That is, a sense of what punitive damages ought to be is critical in forming a judgement of whether they are too big or too little or, as some have argued, so far from the mark that we would be better off without them. Second is to suggest how legal proceedings might be altered in order to have better guidance for decision making by judges and juries.

Punitive damages are meant to deter and to punish. The presence of this double purpose complicates selection of a suitable level of damages. It seems implausible that juries, given little in the way of instructions for fulfilling this double purpose, could reason through to a satisfactory level of damages on a consistent basis. Indeed, calls for more guidance for juries have come from both judges (e.g., Justice O'Connor in *Pacific Mutual Life v. Haslip*) and academic writers [e.g., Owen (1994)]. Similarly, judicial oversight of damage levels also has little guidance about the level of damages (as opposed to the

process). This article (and my previous article) argue for the importance of distinguishing among different bases for levying punitive damages and giving different guidance to both juries and judges in different situations.[28] Without such distinctions, it seems implausible that punitive damages could serve their social role.

The analytical approach taken here could be a basis for jury instructions after a finding of liability for punitive damages in cases having this type of structure. Rather than simply asking juries to select a level of punitive damages to reflect both deterrence and punishment, the instructions could request that the jury consider each element separately. The instruction about deterrence would explain that the goal was efficient deterrence, not the deterrence of all accidents. The instruction might explain the sources of potential inefficiency by the defendant in guiding the jury's thinking. Thus the instruction would be different in different sorts of settings—malice rather than recklessness, different likelihoods of escaping liability altogether. After an instruction about punishment, the instruction would tell the jury not to simply add the levels suitable for deterrence and retribution, but to average them in some way. An alternative to a simple average might be adjusting upward if there is a strong need to deter because of similar (deterrable) behavior escaping liability. The judge or the jury would subtract compensatory damages and fines when moving from concerns about deterrence and punishment to levels of punitive damages. Such an approach does not suggest that the jury try to solve the first-order condition for an optimum. And such an approach might miss the mark, for example, possibly violating the conditions in Equation (6). But such an approach seems likely to produce a better overall outcome than does the current approach.

28. The first distinction to draw is between situations of reckless disregard and those of malice. In the presence of malicious intent, efficiency requires consideration of the preferences of the defendant while selecting a suitable punitive amount to deter behavior that generates gains that are not suitable for inclusion in social accounting. In contrast, reckless disregard of a risk is behavior that is viewed as outrageous as a result of inadequate attention to a cost falling on the plaintiff. Thus jury instructions in the case of reckless disregard should focus on the costs that are not adequately represented in the defendant's decision process.

Presumably, consideration of retribution would reinforce the advantages of drawing distinctions, since the basis for retribution is different with malice and reckless disregard. For example, Owen (1985): "If our accident law of torts should be abolished, as one day it probably should, a question arises whether there is any remaining place in the legal system for retribution or deterrence. I think there is. ... I believe that some form of punitive damages should be retained for intentional torts [footnote omitted]. The goals of retribution, corrective justice, and deterrence are most appropriate in redressing intentionally inflicted harm" (p. 670).

In my earlier article, it was argued that the role of punitive damages in improving efficiency differed between situations of rational disregard and situations of nonrational disregard. For the former, one wanted to consider the costs that did not enter into decisions, because the defendant did not expect to pay them. For the latter, the focus was on inducing the nonrational to behave more rationally. Moreover, it was argued that concerns about inappropriate deterrence were different, since recognizing when costs might not be faced was different from recognizing when rational decisions were absent.

It might be useful to research the thinking of juries with more ex post interviews. When juries use large punitive damage judgments to "send a message," is this a form of punishment or a strong deterrence motive? If, as I suspect, it is the latter, then there are two important questions to ask.[29] One is whether the punishment is excessive relative to retribution standards, and so, along the lines of the analysis above, there should be some compromise between deterrence and a lower ideal retribution level. Second is whether the jury version of deterrence by extreme punishment, while possibly suitable in some settings of malicious intent, is a manifestation of hindsight bias in a setting of reckless disregard. If the jury fails to recognize both the costs of overdeterrence and the risk of accident despite appropriate care, then these punishments are likely to be too large.

Appendix A: Derivation of Optimal Punitive Damages for Section 3

Precaution π is an increasing function of total damages plus fines assessed, $C+P$, which is given implicitly by the first-order condition for the maximization of Equation (3) above. Thus $\pi[C+P]$ satisfies

$$u'[\pi]+C+P=0. \tag{A1}$$

For later use we differentiate this expression twice with respect to P:

$$u''\pi'+1=0;\ u'''(\pi')^2+u''\pi''=0. \tag{A2}$$

We have assumed that in the relevant range, $u'''<0$, and so $\pi''<0$.

Calculating the derivative of W, Equation (4), with respect to P, we have

$$\begin{aligned} W'[P] &= (u'+A+V)\pi'-(1-\pi)V' \\ &= (-(C+P)+A+V)\pi'-(1-\pi)V', \end{aligned} \tag{A3}$$

where we have used Equation (A1) to simplify Equation (A3). In the situation where punitive damages should be positive, this gives the first-order condition in Equation (5).

If there are no fines, $C=A$, and a desire for retribution beyond accident costs, $P^R>A$, then $W[P]$ is increasing in P at $P=0$, since

$$W'[0]=V\left[A-P^R\right]\pi'-(1-\pi)V'\left[A-P^R\right]>0. \tag{A4}$$

With V positive, V' negative, and π' positive, we conclude that when $C=A$, $W'[0]>0$. When there are fines, $C>A$, then $W'[0]$ may be either positive or negative, even if $C<P^R$.

29. For example, two of the jurors in the *Exxon Valdez* case have been quoted: "Our job was to make sure they'd never do it again" [Nancy Provost, cited by Munk (1994:89)]. "We wanted to make sure Exxon is never going to do this again" (Spann, quoted by Natalie Phillips, *Anchorage Daily News*, January 22, 1995, back page).

Similarly, when ideal retribution exceeds accident costs plus fines, setting punitive damages equal to the difference, $P = P^R - C$, we have

$$W'[P^R - C] = (A - P^R)\pi' < 0, \tag{A5}$$

where we have used the conditions $V[0] = V'[0] = 0$. Since $u[\pi[C+P]] - (1-\pi[C+P])A$ is decreasing in P for $P + C > P^R > A$, we have lower values of W above $P^R - C$ than at $P^R - C$. Thus the optimum, if not at the corner at zero, lies between 0 and $P^R - C$. Using the first-order condition of Equation (5), the condition $\pi' > 0$, and the result that at the optimum $V' < 0$, at an interior optimum we have $P + C - A > V[C + P - P^R]$, completing the proof of Equation (6).

Appendix B: Derivation of Optimal Punitive Damages for Section 5

Rewriting the first-order condition from Section 4 to make the role of wealth specific, with deviations from ideal punishment evaluated in relative terms, we have

$$\{A - C - P + V[(C+P)Y^{-1} - P^r]\}\pi'[C+P] - (1-\pi[C+P])V'[(C+P)Y^{-1} - P^r]Y^{-1} = 0. \tag{B1}$$

To calculate the variation of optimal punitive damages with respect to wealth, we differentiate Equation (B1) implicitly:

$$\frac{dP}{dY} = \frac{\{(C+P)V'\pi' - (1-\pi)(C+P)V''Y^{-1} - (1-\pi)V'\}Y^{-2}}{\{2V'\pi'Y^{-1} - (1-\pi)V''Y^{-2} - \pi' + (A-C-P+V)\pi''\}}. \tag{B2}$$

From the assumed concavity of W in P, the denominator is negative. With $V' < 0$, $\pi' > 0$, and $V'' > 0$, the first two terms in the numerator are also negative, although the third term is positive. One way of examining these terms is in terms of the elasticity of accident probability with respect to liability costs, $-(C+P)\pi'/(1-\pi)$. If the elasticity is approximately one, so that accident probability declines proportionately to liability costs, then the first and third terms roughly balance and the second term gives the expected negative sign. Large elasticities are needed to reverse the sign. Thus if P is increasing with Y, the numerator is smaller in absolute value than the sum of the first two terms. In the denominator, with $\pi'' < 0$ and $C + P > A + V$, the first three terms are negative, but the fourth is positive.

Converting Equation (B2) into an elasticity of total costs with respect to wealth, we have

$$\frac{Y}{C+P}\frac{dP}{dY} = \frac{\{(C+P)V'\pi' - (1-\pi)(C+P)V''Y^{-1} - (1-\pi)V'\}Y^{-1}}{(C+P)\{2V'\pi'Y^{-1} - (1-\pi)V''Y^{-2} - \pi' + (A-C-P+V)\pi''\}}. \tag{B3}$$

The positive sign of the third term in the numerator, the factor 2 in the denominator, and the negative sign of the third term in the denominator push toward an elasticity below one, but the fourth term in the denominator pushes the other way. We can conclude that when π'' is small, the elasticity is less than one. This can be expressed alternatively in terms of u''' being small.

Turning to the situation where the deviations from ideal punishment are judged in absolute terms, we rewrite the disutility of deviations from ideal punishment as $V[C+P-P'Y]$. The first-order condition of Equation (5) becomes

$$\{A-C-P+V[C+P-P^rY]\}\pi'[C+P] - (1-\pi[C+P])V'[(C+P)-P^rY]=0. \tag{B4}$$

To calculate the variation of optimal punitive damages with respect to wealth, we differentiate Equation (B4) implicitly:

$$\frac{dP}{dY}=\frac{P^rV'\pi'-(1-\pi)P^rV''}{2V'\pi'-(1-\pi)V''-\pi'+(A-C-P+V)\pi''}. \tag{B5}$$

The numerator is negative, so the derivative is positive. Converting Equation (B5) into an elasticity, we have

$$\frac{Y}{C+P}\frac{dP}{dY} = \frac{P^rY\{V'\pi'-(1-\pi)V''\}}{(C+P)\{2V'\pi'-(1-\pi)V''-\pi'+(A-C-P+V)\pi''\}}. \tag{B6}$$

The term $-\pi'$ in the denominator and the factor 2 in the denominator both contribute toward an elasticity of less than one. However, the term with π'' pushes in the other direction, as does the presence of $P^rY/(C+P)$, which is greater than one. The elasticity will be less than one when π'' is small and the punishment is close to ideal.

Appendix C: Derivation of Optimal Punitive Damages for Section 7

Precaution, π, is an increasing function of expected total damages assessed, $a(bA+F+P)=a(C+P)$, which is given implicitly by the first-order condition for the maximization of Equation (10):

$$u'[\pi]+a(C+P)=0. \tag{C1}$$

We continue to assume that $\pi'>0$ and $\pi''<0$.

As above, we substitute the level of precaution, $\pi[a(C+P)]$, as defined by maximization of Equation (9), in the social welfare function of Equation (10), giving the objective function to be maximized:

$$W[P]=u[\pi[a(C+P)]]-(1-\pi[a(C+P)]) \times\{A+aV[C+P-P^R]+(1-a)V[-P^R]\}. \tag{C2}$$

Calculating the derivative of W, we have

$$W'[P]=\{u'+A+aV[C+P-P^R]+(1-a)V[-P^R]\}a\pi' - (1-\pi)aV'[C+P-P^R]. \quad \text{(C3)}$$

Which, using Equation (C1) to simplify Equation (C3), gives the first-order condition presented above.

Let us examine $W'[P]$ evaluated at $P=P^D-C$ and at $P=P^R-C$, assuming these values are positive:

$$W'[P^D-C]=\{aV[P^D-P^R]+(1-a)V[-P^R]\}a\pi' - (1-\pi)aV'[P^D-P^R], \quad \text{(C4)}$$

where we have used Equation (12), $aP^D=A$. If $P^D<P^R$, $V'<0$ and so $W'[P^D-C]>0$. If $P^D=P^R$, W' is still positive because of the gain from deterring accidents that do not result in liability and so escape the desire to have retribution. If $P^D>P^R$, W' may be either positive or negative, depending on the balance between the cost of overpunishing and the gain from deterring improperly punished accidents.

Considering $P=P^R-C$, we have

$$W'[P^R-C]=\{A-aP^R+(1-a)V[-P^R]\}a\pi'. \quad \text{(C5)}$$

If $P^D>P^R$, implying that $A>aP^R$, we have $W'[P^R-C]>0$. If $P^D=P^R$, W' is still positive because of the gain from deterring accidents that do not result in liability and so escape retribution. If $P^D<P^R$, W' may be either positive or negative, depending on the balance between the cost from overdeterring relative to economic efficiency and the gain from further deterring accidents that do not result in liability and so escape the desire to have retribution.

We might have the situation that the fines, F, are sufficiently large that the desire to punish the defendants who are identified is satisfied by the fines, together with the compensatory damages paid, bA. To examine this we consider $W'[0]$:

$$W'[0]=\{A-aC+aV[C-P^R]+(1-a)V[-P^R]\}a\pi' - (1-\pi)aV'[C-P^R]. \quad \text{(C6)}$$

This might be negative for C large enough.

We can conclude that optimal punitive damages might be zero if C is large enough. Alternatively, when optimal punitive damages are positive, they are larger than the smaller of the ideal levels for just economic efficiency or retribution, and may be larger than both of them if they are not too far apart and it is particularly important to avoid the social cost from a lack of retribution for those who escape liability.

References

Breyer, Stephen. 1993. *Breaking the Vicious Circle: Toward Effective Risk Regulation.* Cambridge, MA: Harvard University Press.

Cooter, Robert, and Thomas Ulen. 1997. *Law and Economics,* 2nd ed. Reading, MA: Addison-Wesley.

Diamond, Peter. 1997. "Efficiency Effects of Punitive Damages," working paper, MIT.

Diamond, Peter, and Daniel McFadden. 1974. "Some Uses of the Expenditure Function in Public Finance," 3 *Journal of Public Economics* 3–21.

Ellis, Dorsey D., Jr. 1982. "Fairness and Efficiency in the Law of Punitive Damages," 56 *Southern California Law Review* 1–78.

Fischhoff, Baruch. 1982. "Debiasing," in D. Kahneman, P. Slovic, and A. Tversky, eds., *Judgment under Uncertainty: Heuristics and Biases.* Cambridge: Cambridge University Press, pp. 422–44.

Goldman, Alan H. 1979. "The Paradox of Punishment," 9 *Philosophy and Public Affairs* 42–58.

Huber, Peter W., and Robert E. Litan. 1991. *The Liability Maze.* Washington DC: Brookings Institution.

Kahneman, Daniel, David Schkade, and Cass Sunstein. 1998. "Shared Outrage and Erratic Awards," 16 *Journal of Risk and Uncertainty* 49.

Mirrlees, J. A. 1971. "An Exploration in the Theory of Optimum Income Taxation," 38 *Review of Economic Studies* 175–208.

Morris, Clarence. 1931. "Punitive Damages in Tort Cases, 44 *Harvard Law Review* 1173–209.

Munk, Nina. 1994. "We're Partying Hearty!" October 24 *Forbes* 84–90.

Owen, David G. 1985. "Deterrence and Desert in Tort: A Comment," 73 *California Law Review* 665–76.

———. 1994. "A Punitive Damages Overview: Functions, Problems and Reform," 39 *Villanova Law Review* 363–413.

Packer, Herbert L. 1968. *The Limits of the Criminal Sanction.* Stanford, CA: Stanford University Press.

Polinsky, A. Mitchell, and Steven Shavell. 1998. "Punitive Damages: An Economic Analysis," 111 *Harvard Law Review* 869–962.

———. 2000. "The Fairness of Sanctions: Some Implications for Optimal Enforcement Policy," 2 *American Law and Economics Review* 223–37.

Posner, Richard A. 1975. "The Economic Approach to Law," 53 *Texas Law Review* 757–83.

Rawls, John. 1955. "Two Concepts of Rules," 64 *Philosophical Review* 3–32.

Shavell, Steven. 1981. "A Note on Efficiency vs. Distributional Equity in Legal Rulemaking: Should Distributional Equity Matter Given Optimal Income Taxation?" 71 *American Economic Review* 414–18.

———. 1992. "Liability and the Incentive to Obtain Information about Risk," 21 *Journal of Legal Studies* 259–70.

Sunstein, Cass, Daniel Kahneman, and David Schkade. 1998. "Assessing Punitive Damages," 107 *Yale Law Journal* 2071.

Wertheimer, Alan. 1975. "Should the Punishment Fit the Crime?" 3 *Social Theory and Practice* 403–23.