(64)

working paper
department
of economics

EFFICIENCY EFFECTS OF PUNITIVE DAMAGES

Peter Diamond

No. 97-17   September , 1997

massachusetts
institute of
technology

50 memorial drive
cambridge, mass. 02139

EFFICIENCY EFFECTS OF PUNITIVE DAMAGES

Peter Diamond

August 30, 1997

Efficiency Effects of Punitive Damages

Peter Diamond

Abstract

   This paper develops a typology of different behaviors that might be viewed as outrageous by a jury and subjected to punitive damages   The paper derives the level of punitive damages for achieving economic efficiency in four different situations - malice and three settings where a jury might find reckless disregard - a rational response to insufficient compensatory damages, a nonrational disregard of risk, and a rational response when compensatory damages are adequate.

   Since efficient deterrence is reached in very different ways in situations of malicious intent and reckless disregard, jury instructions should differ in these different situations. With malicious intent, the jury should be instructed to focus on the preferences of the defendant. With reckless disregard, the jury should be instructed to focus on the circumstances leading to the accident, particularly the inadequacy of attention to a cost falling on the plaintiff. Inadequate attention could come from a rational disregard of costs that the tort system will not assess or from a nonrational disregard of the costs that the tort system will assess. Jury instructions with reckless disregard should focus on the costs that are not adequately represented in the defendant's decision process. This focus is naturally stated in terms of a ratio of punitive to compensatory damages, or equivalently the ratio of the sum of compensatory and punitive damages to the level of compensatory damages. The goal is to pay appropriate attention to accident costs, not to have a zero accident risk. Therefore, in the absence of malice, the court should determine if an underassessment of tort liability was anticipated or if decision-making was nonrational. In the absence of malice, underassessment of liability, and nonrationality, punitive damages should not be allowed.

   It is argued that in some situations of malicious intent, courts should allow arguments that encourage the setting of damages proportional to the wealth of the defendant, but not in situations of reckless disregard. Civil and criminal fines are also part of the expected costs of the defendant, and should be deducted from the punitive damages that would give efficient incentives with zero fines. The analysis of insufficient compensatory damages is extended to a situation with defendants who differ, but in ways the court can not distinguish.

Efficiency Effects of Punitive Damages

Peter Diamond[1]

A great deal of controversy surrounds the use of punitive damages. Courts and academic writers alike have worked at developing a well-based set of rules to govern punitive damages. Liability for punitive damages can occur when a jury finds the behavior of a defendant outrageous. The courts and legislatures have identified two different classes of behavior they are willing to allow juries to consider outrageous and so a basis for punitive damages - malice and reckless disregard of the risk to others.[2][3]

---

2 As an example of instructions to the jury, the following was given in Re the Exxon Valdez.
"Plaintiff has the burden of proving that punitive damages should be awarded by a preponderance of the evidence. You may award punitive damages only if you find that defendant's conduct
    (1) was malicious; or
    (2) manifested reckless or callous disregard for the rights of others.
Conduct is malicious if it is accompanied by ill will, or spite, or if it is for the purpose of injuring another.
In order for conduct to be in reckless or callous disregard of the rights of others, four factors must be present. First, a defendant must be subjectively conscious of a particular grave danger or risk of harm, and the danger or risk must be a foreseeable and probable effect of the conduct. Second, the particular danger or risk of which the defendant was subjectively conscious must in fact have eventuated. Third, a defendant must have disregarded the risk in deciding how to act. Fourth, a defendant's conduct in ignoring the danger or risk must have involved a gross deviation from the level of care which an ordinary person would use, having due regard to all the circumstances.
Reckless conduct is not the same as negligence. Negligence is the failure to use such care as a reasonable, prudent, and careful person would use under similar circumstances. Reckless conduct differs from negligence in that it requires a conscious choice of action, either with knowledge of serious danger to others or with knowledge of facts which would disclose the danger to any reasonable person."
Phase I Jury Instruction No. 28, Clerk's Docket No. 5309.
3 Outrageous behavior without a malicious intent is referred to with 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.

The amount of punitive damages is supposed to punish and deter such behavior.[4] This paper focuses on deterrence, leaving for a later paper (1997a) the integration of punishment and deterrence concerns. This paper develops a typology of different behaviors that might be viewed as outrageous by a jury and derives the level of deterrence appropriate for achieving economic efficiency in the presence of each behavior.[5] The paper does not consider changing the measurement of compensatory damages or the levels of civil and criminal fines as alternatives to the use of punitive damages.

Four different situations are explored in the paper, malice and three settings where a jury might find reckless disregard - a rational response to insufficient compensatory damages, a nonrational disregard of risk, and a rational response when compensatory damages are adequate. Actual malice occurs in a situation where the defendant is seeking to gain (financially or psychologically) at the expense of the plaintiff. When a party acts with malice, society does not value the gain to the defendant the same way that the defendant does. Thus punitive damages can aid efficiency by offsetting the gains to the defendant that are not valued by society.

In situations of reckless disregard, the defendant is engaging in a socially legitimate activity that involves some risk to others, but the defendant has generated an outcome that results in a jury's finding of outrageous behavior. We identify three situations where this can happen. In the first situation, the defendant is acting in a rational way, but the defects of the tort system result in compensatory damages that are too small to achieve efficient deterrence.[6] In this situation, punitive damages can improve efficiency by offsetting the defects in the tort system. We will refer to reckless disregard when the defendant is acting rationally as "rational disregard." In the second situation, the defendant is not making the appropriate calculations, is not acting rationally.[7] We will refer to reckless disregard when the defendant is not acting rationally as "nonrational disregard." In this situation punitive damages can induce some individuals behaving nonrationally to behave more efficiently. These efficiency gains are offset by the overdeterrence of those who were making rational decisions, but might be subjected to punitive

_____

4 I do not consider specific as opposed to general deterrence, since issues around a particular defendant are not central for the analysis below.
5 For a full discussion of the law of punitive damages, see Boston, 1993.
6 In terms of the instructions in footnote 2, this corresponds to "a conscious choice ... with knowledge of serious danger to others."
7 This approach includes nonrational behavior that makes an outrageous decision "either with knowledge of serious danger to others or with knowledge of facts which would disclose the danger to any reasonable person." The assumption is that the knowledge is being ignored.

damages nevertheless. In this situation, an efficient level of punitive damages needs to balance these concerns, limiting the size of punitive judgments. The distinction between these two situations calls for trying to distinguish between nonrational and rational responses to risk. The third situation is where there is no malice, no underassessment of compensatory damages and a defendant who acted rationally. In this situation, any punitive damages are overdeterrence leading to inefficiency and should not be allowed. Hindsight bias might lead a jury to find rational decisionmaking reflecting full social costs to be outrageous.

This paper concentrates on situations involving individual defendants and accidents among strangers, although punitive damages are levied on corporations and after accidents involving people who have a contractual relationship (e. g., products liability). The issues analyzed here are relevant for the other situations, but additional issues arise. For example, products liability analysis needs to examine the impact on market pricing and market behavior. A major issue in considering liability for corporations is the role of multiple decisionmakers, for example, an employee such as a ship captain, and the supervisors who monitor ship captains. Insofar as a corporation can be considered as a unitary decisionmaker, the analysis in this paper is applicable. I will return to the corporate situation in a later paper about vicarious liability (1997b).

At present, the jury has wide latitude in setting punitive damages. Instructions to the jury to guide this choice are the same in malicious intent and reckless disregard situations. With the same instructions to juries in different situations, it is difficult for judges to review the appropriateness of the damage levels set by juries. Since the efficiency goals of deterrence are reached in very different ways in situations of malicious intent and reckless disregard, it is argued that jury instructions should be different in these two different situations. In particular, it is argued that with reckless disregard, but not with malicious intent, it is appropriate to direct the jury's thinking to be in terms of a ratio of punitive to compensatory damages.

The wealth of the defendant is admissible evidence to help the jury assess a suitable level of punitive damages. Plaintiffs' arguments commonly assume the appropriateness of having punitive damages be proportional to the wealth of the defendant (for example, twice the damages for a defendant twice as wealthy). It is argued that in some situations of malicious intent, courts should allow arguments that implicitly encourage the setting of damages proportional to the wealth of the defendant, but not in the situation of reckless disregard.

The paper begins (Section I) by briefly reviewing the standard law-and-economics analysis of when compensatory damages are sufficient to induce efficiency. Then the paper considers three shortcomings of compensatory damages that are widely recognized

in the literature on deterrence-based justifications for punitive damages in the presence of rational decision-making by the defendant (see, e. g., Ellis, 1982). Section II considers malicious intent - a desire by the defendant to deliberately inflict harm or to benefit at the expense of the plaintiff. This benefit is not recognized as a socially legitimate benefit. This section analyzes the efficiency role of punitive damages to deter malicious acts. Section III analyzes the amount of punitive damages that would be appropriate in order to have the efficient level of deterrence when the probability of detection is less than one or there is the possibility of uncompensated damages. The analysis in Section III assumes a homogeneous class of defendants, i. e., all the potential defendants are assumed to be the same. Section IV extends the analysis to a situation with heterogeneous defendants - defendants who differ, but in ways the court can not distinguish. The mathematical derivation is in Section V, a section that can be skipped without loss of continuity. Malicious intent can well be viewed as outrageous by a jury, it is necessary to understand why juries might sometimes find behavior outrageous in a situation when the defendant has no malicious intent. A jury might find outrageous the undertaking of risky actions that seem worthwhile to the defendant because of the underdetection, underassessment, or undercompensation of compensatory damages. If the activity is undertaken for such a reason, then there is a potential for punitive damages, if set appropriately, to increase efficiency.

However, the assessment of punitive damages is not restricted to situations with malicious intent or underassessment of liability. Thus, additional bases that might lead a jury to a conclusion of reckless disregard are examined. The paper draws a distinction between "rational disregard" of risk and "nonrational disregard" of risk.[8] Section VI analyzes the former case and concludes that there does not appear to be an efficiency justification for punitive damages in this situation. In Section VII, we turn to nonrational disregard, formally modelling an example of nonrational decision-making. In light of the difficulty for juries in assessing decision-making by defendants, we allow for some erroneous assessment of punitive damages even when decision-making was rational. Thus, the risk of being subject to punitive damages designed to deter behavior not based on full rationality also affects the precaution decisions of rational agents. The efficient level of punitive damages will balance the cost of overdeterrence of rational decision-makers with the benefit of increased care by nonrational decision-makers.

_____

8 This distinction is similar to but not identical to the distinction in the jury instructions quoted in footnote 2 between "knowledge of serious danger to others" and "knowledge of facts which would disclose the danger to any reasonable person." A rational decision-maker has knowledge of serious danger. A nonrational decision-maker may be ignoring either knowledge of danger or knowledge of facts.

In order to explore the common argument by plaintiffs that punitive damages should be set proportionally to the wealth of the defendant, Section VIII assumes that in deterring nonrational behavior, punitive damages work proportionally to the wealth of the defendant. (If this proportionality is not present, then it is not clear what deterrence basis there is for allowing such arguments by the plaintiff.) It is then asked how punitive damages should vary with wealth when nonrational deterrence has this form. Since rational behavior does not have systematic variation of precaution with wealth, variation of punitive damages with wealth affects the precaution of rational agents. Thus optimality again calls for balancing the alternative concerns. The derivations of the results in Sections VII and VIII are in Section IX, which can be skipped without loss of continuity. Implications of this analysis in setting punitive damages are discussed in the concluding section.

I. Compensatory damages and efficiency

The widely recognized law-and-economics argument that compensatory damages will induce efficiency (in the absence of a need for incentives for victims), points out that if compensatory damages equal the monetary value of the harm to others, then a rational decisionmaker will weigh the value of the harm to others along with the net gain of any activity in deciding whether to engage in the activity (see, e. g., Cooter and Ulen, 1997, 272-6). For analytical convenience, this argument is examined here in a situation of strict liability, not negligence, since any defendant at serious risk for being found liable for punitive damages because of malicious intent or reckless disregard is very likely to be held negligent and so liable for compensatory damages if there is a negligence standard.

For convenience of presentation, let us introduce some formal notation. While we could do the analysis in terms of actions that affect the probability of an accident, it is more convenient to consider the probability of an accident directly as the control variable of the defendant (within limits). The defendant can lower the probability of an accident by expending resources and changing behavior in a wide variety of ways. There are decreasing gains from pursuing such precautions, and the defendant will expend resources as long as the cost is less than the expected benefits from a lower accident probability.

We let p equal the probability of <u>avoiding</u> an accident, where p defines the level of precaution being taken by the defendant.[9]

---

[9] It is common in the law and economics literature to consider an explicit care or precaution variable and then to relate both the cost of care and the probability of an accident to this explicit care variable. Since both the cost of care and probability of an accident functions are monotonic, we can simplify the notation by using probabilities as control variables, since there is a direct functional link between the cost of care and the probability of an accident.

Let u[p] 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. Recognizing both the costs to the defendant of any accident and the cost of avoiding accidents, we assume that u is first increasing, then decreasing in p. We also assume that it is impossible to have a zero probability of an accident while engaging in this activity, and that the cost of avoiding accidents rises without limit as the probability of avoiding all accidents increases toward its upper limit. The specific mathematical assumptions about u are detailed in the footnote, along with an example.[10] In Figure 1, we show a typical pattern of utility relative to precaution.

We denote by A the cost to the plaintiff in the event of an accident, including non-economic costs. Thus A is the amount of compensation required to restore the plaintiff to the position held before the accident.[11] If the social evaluation of the choice of the level of precaution is utilitarian, denoted in monetary terms, the social value, W, of equilibrium with a chosen level of precaution of p satisfies:

(1) $W[p] = u[p] - (1-p)A$.

This social evaluation recognizes the utility of the defendant, u[p], and the expected accident costs of the plaintiff, $(1-p)A$.[12][13] The payment of damages by the defendant to the plaintiff is viewed as a transfer without direct social significance. For convenience, the

---

10 We assume that u[p] is strictly concave in p, $u''[p]<0$. We also assume that $u'''[p]<0$. 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[p] might have the form: $u[p] = k_0 - k_1(1-p) - k_2/(1-p)$ for some positive constants k_i. Thus $k_0-k_1-k_2$ is utility of the activity (ignoring liabilty) if an accident is certain (p=0). With this utility function, we have the derivatives: $u'[p] = k_1 - k_2(1-p)^{-2}$; $u''[p] = -2k_2(1-p)^{-3}$; $u'''[p] = -6k_2(1-p)^{-4}$.
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 to preserve the plausible description of the accident environment.
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 Note that this expression involves no concern about the incentives for the plaintiff to avoid accidents or to lower the costs of accidents that do occur.
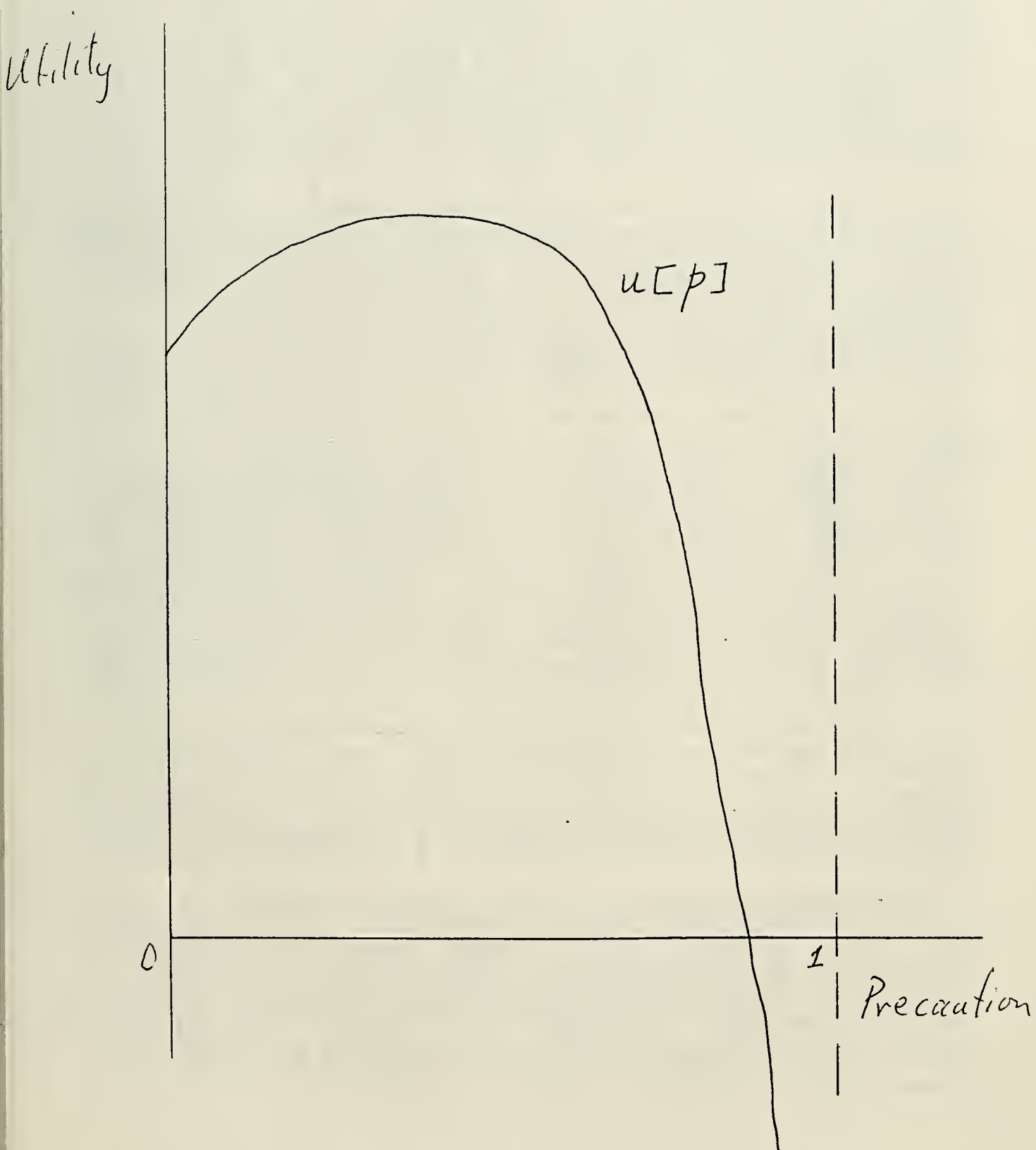
Figure 1

legal costs are taken to be zero.[14]

If the defendant is liable for the plaintiff's accident costs, then the individual utility of the defendant, denoted U[p], equals u[p] - (1-p)A, which coincides with W[p]. In this setting, a rational decision-maker would indeed select the level of precaution that is socially efficient. This analysis is structured for the situation of an accident involving strangers (in the contractual sense), as with environmental damage. Additional complications arise when there is a contractual relationship and so one must examine the impact on prices. Thus, this analysis does not directly apply to issues of products liability.

Implicit in this formulation of social evaluation is an absence of externalities other than the possibility of an accident. If the activity of the defendant has social values which 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. In addition, the undertaking of large projects can involve consumer surplus, which again is not captured by market transactions. Below, we will be concerned with overdeterrence. To the extent that the activities deterred involve positive externalities, then the social concern for overdeterrence becomes larger. "Reputation costs" raise a similar issue. When reactions to an accident do not accurately reflect true risks, then the defendant's costs from an accident can be larger than the social costs. Such reputation costs are thought to be particularly relevant with consumer product risks. Conversely, when deterring accident-generating behavior also deters other negative externalities then there is less concern about overdeterrence. In different particular situations one or the other of these concerns may be larger.

As has been noted in the literature, the argument for efficiency coming from compensatory damages breaks down if the probability of paying damages is less than the probability of an accident, 1-p, or if the level of compensatory damages is less than the costs inflicted on the plaintiff, A. We return to these issues below. But first we consider a model of a maliciously inflicted harm.

---

14 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.

II. Malicious intent

In this section, we assume that the reason an act is undertaken is precisely to cause harm, not to pursue utility in the presence of a risk of accident. Included in this category is a desire to make a profit that can only be accomplished at the expense of the plaintiff, as with fraud or theft. Such motivation, by itself, does not alter the analysis. But, such individual gains are not a positive part of social evaluation. That is, we distinguish between the pleasure of a defendant in hurting the plaintiff and the pleasure of a defendant in carrying out activities that may have an accident as an undesired side effect. We do not ask if a rapist really enjoyed the rape. We do not ask about the enjoyment of the deliberate destruction of someone's property in order to inflict the pain of having it destroyed or in order to flaunt the willingness to do (and pay for) the destruction. Similarly, we do not recognize envy as a good social reason to block people's private gains. Thus, we introduce a further difference between individual utility and social welfare in order to capture this perspective and derive its implications for deterrence.[15]

Since variation in the probability of success in the maliciously intended harm is not central to this analysis, we can simplify the analysis by suppressing the role of precaution - we assume that any attempt to inflict malicious harm is successful. Denote by u the utility of the defendant if the malicious act is not committed. Let us denote the additive pleasure of the malicious act (gross of any legal liability) as v. For simplicity, we take v to be a constant, not varying with the magnitude of damages paid. This is the defendant's utility from the malicious act that does not enter social welfare.[16] As before, we denote the cost inflicted on the plaintiff by A.[17] We introduce the possibility of civil or criminal fines, denoted C, which are socially costless transfers. Now, we can write the defendant's utility assuming the act is carried out as:

$$(2) \quad U = u + v - (A + C + P).$$

That is, we assume that the defendant is held liable and then pays compensatory damages of A, civil and criminal fines of C, and

---

15 This approach to defining social welfare as excluding some individual utility is discussed in Baker, 1978, Schwartz, 1979, Ellis, 1982, and Shavell, 1985.
16 In general, the utility of the defendant from the act and the ignored utility for social evaluation need not coincide exactly. There may be multiple parts to an act that includes a malicious component.
17 As before, we do not consider actions by the plaintiff to avoid this cost.

punitive damages of P.[18][19] Alternatively, if the defendant is deterred, his utility is equal to u - the defendant doesn't receive the utility v from committing the act and the plaintiff does not receive the cost A from being subject to the act. Thus the act will be carried out if v is larger than A+C+P, and the act will be deterred otherwise.

Without social recognition of the utility of inflicting harm, v, social welfare, if the malicious act is performed, is written as

(3) W = u - A.

That is, society recognizes the utility of the defendant other than the part that comes from the malicious act, and society recognizes the accident costs of the plaintiff. Thus, it is socially preferred to deter the act whenever A is positive. The distinction between this situation and the analysis in the first section is that society does not recognize as legitimate the gain from a malicious act, while the gain from an act that has a risk of harming others is recognized. Indeed this can be viewed as the definition of malicious that is being employed.

The minimum level of punitive damages needed to deter the act is determined by equating the defendant's utility levels with and without the act:

(4) u = u + v - (A + C + P).

or, solving for P:

(5) P = (v - A - C).

That is, the defendant is deterred if P > (v - A - C) and the defendant is not deterred if P < (v - A - C). The need to pay compensatory damages and civil and criminal fines deters some malicious acts (where v<A+C). From the definition of malicious, it is socially desirable to deter all malicious acts, and compensatory damages and fines may not be large enough to accomplish this. Thus, there is a potential efficiency role for punitive damages. This situation is different from that where there is a risk to others but the gain from the activity is recognized as socially legitimate.

---

18 It would be straightforward to modify this formula to reflect the possibility that the defendant will not pay damages or will pay less than the full costs of the plaintiff. These two issues are central to the next section. It would also be straightforward to have punitive damages assessed over some but not all malicious acts.

19 We do not consider the possibility that the defendant gets some utility from being assessed punitive damages, for example as proof of willingness to bear financial costs in order to inflict harm.

We have assumed that the utility of inflicting pain, v, is measured in dollars. This is the basis for an important distinction. If the utility of inflicting pain is directly in financial terms, then the punitive damages needed to make them unprofitable are also in financial terms and have no reason to vary with the wealth of the defendant. For example, this would be the case with fraud committed solely for profit. In contrast, if the utility of inflicting pain is in utility terms, as with a desire to inflict harm per se, then the utility must be converted into financial terms. To make this conversion, we need to consider how the utility from a malicious act might vary with wealth and how the needed financial disincentive varies with wealth, which depends on the shape of the function relating utility to wealth. I know of no evidence for how v varies with wealth and, lacking an alternative, will assume that it does not vary with wealth. With individual defendants, it is plausible that people with higher incomes or wealth have a lower marginal utility of wealth. There is then an argument that the level of punitive damages needed for deterrence is likely to vary with the wealth of the defendant.[20] For example, if the utility of the defendant as a function of the defendant's wealth, Y, is logarithmic, $u[Y] = \log[Y]$, then the level of punitive plus compensatory damages plus fines needed to deter a given utility gain from a malicious act is indeed proportional to wealth.[21] On the other hand, when considering widely-held corporate (as opposed to individual) defendants, the risk neutral case seems more appropriate (Cooter and Ulen (1997), page 47), leaving no reason to vary punitive damages with wealth in this situation.

This analysis includes the situation of ill-gotten gains - a desire to profit from an exchange that is taking advantage of another party, although it is the gain and not the inflicting of harm that is the goal of the defendant. Fraud cases fall into this category. The formulation above seems to apply well to these cases, since society does not want to recognize the gain obtained in this way as part of its welfare evaluation. We do not explore the problem of deriving a definition of ill-gotten gains from more basic considerations; presumably they would come from sources of profits associated with lower efficiency in the economy, for example violations of trust, especially when such actions violate laws or regulations.

Of course, any larger level of punitive damages than that given in equation (5) would also deter the defendant. In the model as formulated, if taken literally, deterrence works fully -

---

20 For a contrary claim about plausible patterns, see Abraham and Jeffries, 1989.
21 To solve for punitive damages as a function of wealth, we have:
$\log(Y) = \log(Y - A - C - P) + v.$
Solving for P, we have:
$P = kY - A - C$, for some constant k.

defendants never commit an act that would lead to punitive damages. If punitive damages are never assessed, there is no reason to worry about punitive damages being too large, only a reason to worry about their being too small to deter. But this model should not be taken literally (indeed no model should be taken literally). Sometimes, punitive damages will be assessed. This can happen for a number of reasons. Some defendants may not be deterred - they may not be aware of the level of punitive damages, or may not incorporate the possibility of such damages correctly into their decision calculations. To incorporate this concern, one would need to determine the social evaluation of different levels of punitive damages that do get assessed as a result of the failure of punitive damages to deter fully.

Also, juries can make errors. Of particular concern is assessing punitive damages for what appears to be a malicious act when the act was done by someone else or the act was not deliberately intended and only appears malicious. Recognition of the social cost of assessing punitive damages in such cases has two implications. One is a need to have a suitable standard of evidence before assessing punitive damages. The other is not assessing damages at excessive levels relative to the needs for deterrence so as to hold down the costs from erroneous assessment - costs that include both efficiency elements (from overdeterrence) and equity elements (from incorrect punishment). Below, in the context of reckless disregard, we consider the efficiency costs associated with punitive damage assessments when there is poor decision-making by some defendants and when there are errors by some juries. We do not extend the analysis of malicious intent to include these factors. Nor do we consider underassessment of damages when there is malicious intent (as is relevant in some fraud situations).

In the rest of the paper, we focus on punitive damages based on reckless disregard of others, focusing on situations where there is no malicious intent. Reckless disregard is distinguished from malicious intent by the gain aspect of the activity - with reckless disregard, there is no desire to inflict pain on the plaintiff, nor does the defendant seek benefits that require inflicting costs on others. That is, distinguishing reckless disregard from ill-gotten gains is based on the link between the gain of the defendant and the loss of the plaintiff. Ill-gotten gains require a loss from the plaintiff (although not necessarily equal); with reckless disregard, the cost to the plaintiff is an undesired (but inadequately considered) side effect of the activity. With reckless disregard the social welfare function includes the complete utility of the defendant, as well as that of the plaintiff; with malicious intent, the social welfare function omits part of the utility of the defendant. With this distinction in mind, we turn to consideration of reckless disregard.

III. Underdetection and undercompensation

We began in Section I by considering how compensatory damages could result in efficiency. As has been noted in the literature, this argument breaks down if the probability of paying damages is less than the probability of an accident, 1-p, or if the level of compensatory damages is less than the costs inflicted on the plaintiff, A. We begin our consideration of reckless disregard by examining efficiency where compensatory damages may be too small as an incentive for precaution (despite the presence of fines). After examining the effect of punitive damages on efficiency whenever compensatory damage are too small, we turn to the issue of when behavior in such a situation might be considered outrageous, and so subject to punitive damages.

Let a be the probability of being held liable conditional on the occurrence of an accident and let b be the fraction of actual accident costs for which the defendant is held liable (assumed to be less than one), so that compensatory damages paid are equal to bA and expected compensatory damages are abA.[22] The possible inadequacy of compensatory damages can come from defendants who might escape detection or might not be sued, from the possible incompleteness of accident losses allowed in the determination of compensatory damages, and from legal costs that are not compensated.[23] Let us denote by C the fines paid by the defendant. For simplicity, we assume that the defendant will be subject to fines with the same probability that there is liability for compensatory damages. Direct fines paid to the government are included in C. C does not include expenses borne by the defendant to remediate or restore the damage from the accident. Such costs are part of the social cost of the accident, and can be considered as part of u[p], assuming that these costs are paid for sure or part of A if they are paid with probability a.[24]

Let U[p] denote the utility of the defendant, including the payment of damages, but ignoring legal costs of the defendant that are not payments to the plaintiff (which we assume are zero). Then, U[p] satisfies:

---

22 We consider a class of cases with similar levels of a and b and do not explore how, on average, a and b might vary with the size of harm, A.
23 This formulation ignores the likely variation in the probability of being held liable with the level of precaution taken by the defendant. With such variation, the simple conclusions here on the efficiency gains from expanded liability are not necessarily valid. For a careful discussion, see Calfee and Craswell, 1984.
24 If there is inefficiency coming from remediation expenditures that exceed the socially efficient amount, then those excess expenditures are a social cost of the accident for purposes of analysis of punitive damages. Such additional costs are part of punishment and part of deterrence. But they are not part of the externalities caused by the defendant and so in need of correction by punitive damages for efficiency.

(6) U[p] = u[p] - a(1-p)(bA + C).

As above, we assume that social welfare is the utility of the defendant less the expected accident costs of the plaintiff:

(1) W[p] = u[p] - (1-p)A.

Comparing W[p] and U[p], we see that they coincide when a and b are both equal to one (provided C is equal to zero) and will not coincide when either or both of a and b are less than one unless the fines C are just the right amount. For an efficient decision, we want the defendant to choose the level of precaution, p, that maximizes W[p]. A rational decision-maker will choose the level of precaution that maximizes U[p]. Thus the defendant's choice will be efficient when the two functions W[p] and U[p] coincide.

Adding punitive damages, denoted P, to the liability of the defendant can restore the match between U[p] and W[p], provided that the level is set appropriately (see, e. g., Cooter and Ulen, 1997, 314-7). We need to recognize that the defendant might be subjected to compensatory damages but not punitive damages, even though the probability of liability for compensatory damages is less than one. We assume that the probability of being assessed punitive damages equals f times the probability of being assessed compensatory damages. In this case, the utility of the defendant is

(7) U[p] = u[p] - a(1-p)(bA + C + fP),

where fP is the expected level of punitive damages, conditional on liability for compensatory damages. To have U[p] match with W[p], we need the expected liability costs of the defendant to match with the expected social costs of accidents. Thus, equating (1) and (7), we need the punitive damages, P, to satisfy the equation:

(8) A = a(bA + C + fP),

or, solving for P,

(9) P = (A{(1/a) - b} - C)/f.

Of course, this only results in a positive level of punitive damages when fines are small enough. Recognizing that compensatory damages are equal to bA, we can also express the punitive damages needed for efficient deterrence in terms of the compensatory damages as:

(10) P = bA({1/(ab)} - 1)/f - C/f.

Thus with either a probability of liability below one or an assessment of liability below accident costs, compensatory damages

do not give incentives that induce efficient precaution from a rational decision-maker unless fines are sufficient; punitive damages can be used to restore the level of deterrence that matches bearing full social costs if fines are too low. Thus, punitive damages can be seen as a corrective to legislated punishments as much as a corrective to compensatory dmaages.

To see the quantitative significance of equation (10), we can calculate a few examples. If there is a probability of 1/2 of escaping liability, with the bearing of all accident costs and punitive damages when there is liability and no fines (a=0.5, b=1, f=1, C=0), then the ratio of punitive to compensatory damages given in equation (10) is 1 - punitive and compensatory damages should be equal. Alternatively, if compensatory damages are only 1/2 the level of accident costs and there is no probability of escaping liability for this level of compensatory damages or for punitive damages, and there are no fines (a=1, b=0.5, f=1, C=0), then the ratio of punitive to compensatory damages given in equation (10) is again 1 - punitive and compensatory damages should be equal.

There are three familiar bases for differences between the social costs of accidents and costs included in compensatory damages. Some costs are viewed as highly speculative and not allowed as part of compensatory damages because of the difficulty of accurate measurement. The presence of outrageous behavior might tip the balance in the willingness to try to measure such costs. When pain and suffering were less frequently compensated than currently, punitive damages might have been a way to introduce compensation in a class of cases, without opening up similar liability after all negligently caused accidents.

Legal fees of the plaintiff (which are part of the social cost of an accident and the ensuing compensation) are not normally compensated, although there are situations where legislation allows compensation for legal fees. Just as the balance is struck differently in some of these situations, so the presence of outrageous behavior might be seen as a reason to compensate the legal fees of the plaintiff - a different balance in the effects of allowing such recovery.

The third familiar basis for a possible deviation between social costs and compensated costs comes from various legal rules that restrict recovery of remote economic losses. These include the <u>Robins Dry Dock</u> rule of maritime law that conditions recovery of economic loss on physical impact, as well as various other common law doctrines that may bring about similar results under the rubrics of "particular foreseeability" or "proximate" causation. From the perspective of matching social costs and compensation, there is an important distinction among situations. The distinction is based on whether there are offsetting gains which are part of the social effects of the accident, but which are also not counted in determining compensatory damages. In some

settings, individuals without direct physical injury from the accident nevertheless suffer economic losses from delay, lost business, etc. If fines are not adequate deterrence for accident generating behavior, then there may be scope for improved efficiency from punitive damages. However, in other settings, the losses come naturally with offsetting gains, for example, when business is diverted from one supplier to another. If one were to count the uncompensated losses in setting punitive damages, but not count the implied gains, then one would be overdeterring such accidents. Thus, within the framework underlying equation (10), one can include legal fees, hard-to-measure costs, and costs to others that are not accompanied by gains to still others.

This paper focuses on the efficiency implications of alternative levels of punitive damages, taking as given both the set of situations resulting in punitive liability and the workings of compensatory damages and government fine setting. Nevertheless, it is interesting to briefly ask whether any category of costs that might be thought of as a basis for an efficiency gain from punitive damages, could be better dealt with by changing compensatory damage rules (which involve more guidance for juries than do punitive damage instructions) rather than using punitive damages. Insofar as the set of situations resulting in punitive damages differs significantly from the set of situations resulting in compensatory but not punitive damages, there may be a basis for preferring the current division between compensatory and punitive damages. Analysis of this question would need to consider the reasons underlying the rules for compensatory damages (e. g., considerations of measurability, fairness and incentives for plaintiffs) before one could judge whether there was a good basis for a distinction in this class of situations. If there was a reason for different damage rules in this set of situations, it must lie in the difference between the set of situations resulting in punitive damages and the set of all situations resulting in compensatory damages.

If sometimes defendants escape all liability (a<1), then the alternatives to the use of punitive damages to directly offset this limitation are greater fines and encouragement of additional detection and initiation of suits against defendants. There is an appropriate level of encouragement of detection and initiation of suits, given the social costs and benefits associated with private suits (see, e. g., Shavell, 1997). Whatever residual underdetection remains when the appropriate level of encouragement is present and fines are at whatever level has been set by legislation is a basis for a potential efficiency gain from appropriately set punitive damages.

There are several elements to notice in equations (9) and (10). First, the level of punitive damages (adjusted for the expected fines to be paid) is proportional to the level of costs that should be compensated and to the level of compensatory damages. This is in contrast with the analysis of malicious intent where it was the

utility of the defendant, not the uncompensated costs of the plaintiff that were the focus of the analysis. Second, again in contrast with the situation of malicious intent, the appropriate level of punitive damages has no variation with the wealth of the defendant, unless one of a, b, C, or (1-f), the levels of underdetection, undercompensation, fines and avoidance of punitive damages, vary systematically with the wealth of the defendant. If there were a reason to believe that any of these levels did vary systematically with wealth, the effect on optimal punitive damages could be explored explicitly.

This suggests that the current undifferentiated approach to punitive damages should be refined - guidance to juries should be different in the different situations of malicious intent and underassessment. In situations of malicious intent, the jury should be instructed to consider the preferences of the defendant in selecting a suitable amount for deterrence purposes. In a situation of reckless disregard, the jury should be instructed to consider the circumstances leading to the accident, not the preferences of the defendant, in selecting an amount to offset the inadequacy of compensatory damages.

The modelling of deterrence treating all dollars as equally important ignores the equity issues raised by having a defendant pay more than the damage actually caused because someone else may have caused a similar harm and not been caught. A similar issue arises when additional harm might have been caused by this defendant, although it is not proved in court. While these are substantial equity issues, they will not be explored here.

Punitive damages are supposed to be triggered by outrageous behavior. While malicious intent can be readily seen as outrageous, it is natural to examine how a link between underassessment of compensatory damages and a jury's perception of outrageous behavior might be constructed. Possibly it comes from identification of behavior based solely on self-interested economic calculations (actual liability) rather than a socially inclusive economic calculation. That is, running a high risk of an accident because one expects to pay only part of the cost might be viewed as outrageous behavior by a jury. In other words, such a jury is looking to the spirit of compensatory damage law rather than the letter of the law - individuals who take "too much" advantage of the weaknesses of the law are subjected to punitive liability. Punitive damages then creates concern for costs that remian on others because of the shortcomings of the compensatory damages assessment process. While responding to inadequate incentives might be viewed as outrageous, merely allowing an accident to occur (especially after efficient precaution) should not be viewed as outrageous.

Currently, neither the determination of liability for punitive damages nor their level conform with the patterns that would be appropriate if underassessment of compensatory damages were the

sole basis for liability when malicious intent was not present.
There is no required discussion of the likelihood of
underdetection or of the magnitude of inadequate compensation in
the trial. Where the motivation for punitive damages is derived
from the inadequacies of compensatory damages, one would expect
to require extensive discussion of those inadequacies. Moreover,
punitive damages would not then be allowed where lack of
detection of the harm and its the cause are not significant issues
and where economic damages, as regularly measured, are a good
approximation of total costs inflicted. In addition, juries would
be instructed to adjust the level of punitive damages for the
amounts of fines paid. Either these assessments of punitive
damages are erroneous, or there is some additional basis for
identifying situations where punitive damages are appropriate.
We consider another basis for a finding of reckless disregard in
Section VII, based on nonrational decision-making. First, we
generalize the model in this section to one with heterogeneous
defendants. In such a situation, with asymmetric information
about the defendants, punitive damages can not work perfectly, as
they do in the current model. Thus it is appropriate to examine
the efficiency implications of punitive damages when defendants
are heterogeneous and to derive the rules that should be followed
in this situation for (second-best) efficiency. This section and the
mathematical derivation in Section V can be skipped without loss
of continuity.

IV.   Defendants who differ in ways the court can not distinguish

     In the situations above, punitive damages achieve efficient
deterrence. With homogeneous defendants, the level of punitive
damages needed for efficiency is the same for all defendants. The
assumption is not that all defendants are the same, but that the
court can distinguish among defendants who are different and
therefore is able to set down rules that are suitable for each class
of defendants, with homogeneity within each class. In contrast, if
we have heterogeneous defendants who can not be distinguished
by the legal system when setting punitive damages, then we can
not achieve full efficiency. That is, we assume that there is a
class of defendants who must be treated the same since it is not
possible to tell them apart. Yet, there are in fact differences
among these defendants. In this situation the efficient level of
punitive damages is the solution to a second-best problem. In
simple situations of this kind, the second-best efficient level of
punitive damages will be a weighted average of the levels that
would be set if different types of defendants were the only ones
subject to punitive damages. The weights in this weighted average
depend not only on the numbers of defendants of different types
but also on their responsiveness to financial incentives. This
familiar property of second-best optimization is brought out below
in a model with two types of defendants.[25]

_____

25 For a discussion of second-best pricing to correct for
externalities, see Diamond, 1973.

We now assume there are two classes of defendants who differ in the probabilities of underdetection, although we assume complete compensation for mathematical convenience. We also assume that there are no fines paid. We assume that each person knows his own type - each person knows the likelihood of being liable for the consequences of an accident. We refer to these two classes as types 1 and 2 and assume that the conditional probabilities of assessment of damages for the two classes satisfy $a_1 < a_2 < 1$. We assume that after an accident, the legal system can not distinguish between types. It is inherently very difficult to measure ex ante accident probabilities from an ex post perspective. While some differences in precaution will indeed be plausibly observable after an accident, other differences will not be. We assume that the legal system knows that there is a fraction, g, of defendants who are type 1 and a fraction (1-g) who are type 2. And the legal system knows the probabilities of assessment for each type. But, the legal system must set the same level of punitive damages for both types.

Since they differ in the probability of detection, the two types differ in behavior. Their precaution decisions depend on the sum of compensatory and punitive damages to which they are exposed when found liable. We denote these precaution decisions by the functions $p^1[A+fP]$ and $p^2[A+fP]$, and note that type 1 takes less precaution than does type 2, since type 1 is less likely to have to pay damages after an accident. For efficiency calculations, we use a social welfare function that is a sum of individual utilities less expected accident costs. To contrast the optimal level of punitive damages in this situation with that when the defendants are homogeneous, we define a weighted average of the probabilities of detection, denoted by a_0:

(11) $\quad a_0 = \{ga_1p^{1\prime} + (1-g)a_2p^{2\prime}\} / \{gp^{1\prime} + (1-g)p^{2\prime}\}.$

Thus the weights in this calculation use the fractions g and 1-g to average the changes in precaution for the types, the derivatives $p^{1\prime}$ and $p^{2\prime}$. The weights are not simply the numbers of each type or the numbers of accidents involving each type. This is a standard finding in second best analysis. Because incentives are not exactly right for each person, the efficiency concern is focused on the responsiveness of the behavior that we want changed. Note that this is a proper weighted average, with weights between zero and one and adding to one.[26]

As derived in the next section, which can be skipped without loss of continuity, the optimal level of punitive damages satisfies the implicit equation:

(12) $\quad P = A\{(1/a_0) - 1\}/f.$

---

[26] As a proper weighted average, we have $a_1 < a_0 < a_2$.

This is the same expression as with a homogeneous population, (9), with a for the homogeneous population replaced by the weighted average a_0. Thus the inability of the courts to distinguish types results in the type with lower probability of detection underdeterred, while the type with higher probability of detection is overdeterred. All defendants are subject to the same punitive damages, P. If the court could distinguish between the types, it would want to raise the punitive damages for one type and lower them for the other. This would give better incentives for the levels of precaution. A central point to notice is that the responsiveness of precaution, $p^{1\prime}$, plays a major role in the averaging. This reflects the fact that punitive damages only play an efficiency role when they alter the behavior of potential defendants.

This analysis has focused narrowly on the deterrence (efficiency) dimension of punitive damages with heterogeneous defendants. We do not discuss the fairness of assessing such damages based solely on these efficiency characteristics, which reflect incentives for a wide class of potential defendants, not all of whom are the same.

V. Derivation of optimal punitive damages

Since they differ in the probability of detection, the two types differ in behavior. The two utility functions are:

(13) $\quad U^1[p] = u[p] - a_1(1-p)(A + fP).$
$\qquad U^2[p] = u[p] - a_2(1-p)(A + fP).$

Thus their precaution decisions, which we denote by $p^1[A+fP]$ and $p^2[A+fP]$, satisfy the two first order conditions:

(14) $\quad u'[p^1] + a_1(A+fP) = 0;$
$\qquad u'[p^2] + a_2(A+fP) = 0.$

Comparing these equations, we can conclude that type 1 takes less precaution than does type 2 since type 1 is less likely to have to pay damages after an accident. That is, since u' is decreasing, $a_1 < a_2$ implies $p^1 < p^2$.

For efficiency calculations, we use a social welfare function that is a sum over individuals of utilities less expected accident costs:

(15) $\quad W[P] = g\{u[p^1[A+fP]] - (1-p^1[A+fP])A\}$
$\qquad\qquad + (1-g)\{u[p^2[A+fP]] - (1-p^2[A+fP])A\}.$

To calculate the optimal level of punitive damages, we differentiate W[p] with respect to P and set the derivative equal to zero:

(16) $g\{u'[p^1] + A\}fp^{1\prime} + (1-g)\{u'[p^2] + A\}fp^{2\prime} = 0$.

Substituting from the first order conditions for the precaution decisions, (14), we can rewrite (16) as:

(17) $g\{A - (A+fP)a_1\}fp^{1\prime} + (1-g)\{A - (A+fP)a_2\}fp^{2\prime} = 0$.

To contrast the optimal level of punitive damages in this situation with that when the defendants are homogeneous, we define a weighted average of the probabilities of detection, denoted by a_0 and given in equation (11). Solving equation (17) for P we have equation (12) above.

(12) $P = A\{(1/a_0) - 1\}/f$.

VI.   Reckless disregard

Punitive damages are only supposed to be assessed when the behavior of the defendant is viewed as outrageous. Thus, to have punitive damages assessed when there is no malice and a rational choice of precaution by a defendant in the presence of inadequate assessment of compensatory damages, the jury must view it as outrageous (reckless disregard) to ignore the costs for which one is not liable, at least when those costs are sufficiently large. In Section III, we showed that in such a situation, punitive damages can increase efficiency provided they are not set too high.

Yet punitive damages are sometimes assessed where there is little or no reason to think that the compensatory damage system has inadequate incentives (and there is no sign of actual malice). In order to consider an efficiency basis for punitive damages in a situation where compensatory damages adequately reflect the costs falling on others, we need to consider the behavior of defendants which a jury might find to be a reckless disregard of others despite a full measure of compensatory damages. That is, for a jury to find outrageous behavior, the decision-making that led to the accident must be viewed as seriously flawed - mere occurrence of an accident should not be a basis for an inference of outrageous behavior. When the defendant faces all the costs of an accident, we need to identify another reason for a jury to conclude that there is a reckless disregard of others. We shall explore two potential bases for this conclusion. A jury might conclude that the defendant did not give adequate weight to the risk of an accident even though the defendant was acting rationally. This might occur as a consequence of the process used by the defendant when gathering information about accident risks. We will call this "rational disregard" - the situation when the defendant makes rational decisions about information gathering. The second basis will be called "nonrational disregard."[27]

---

27 The distinction among malicious intent, rational disregard, and nonrational disregard is similar to the distinction among three mental states in Owen, 1982, page 106: deliberate (or malicious), evaluative (or reflective) and inadvertent (or impulsive).

The distinction between these potential bases for punitive damages parallels a distinction between two methods of decision-making in the presence of risks that has been drawn by Tamerin and Resnik (1972). On one hand they identify some risk taking that is based on conscious deliberation, adequate preparation, positive aims, and responsibility for outcomes. In contrast, there is risk taking "with little deliberation, incomplete planning (or no planning at all), for negative reasons, and without responsibility for outcome. This kind of risk taking is generally associated with individual psychopathology." (Page 74.) This section will consider rational disregard, while the next section considers nonrational disregard.

There are two bases for a reasoned decision that undervalues the risk of an accident - one is low awareness of the risk (associated with limited information), while the second is cognitive error in processing the information about the risk. The world is full of risks; when people think about risks, perhaps by constructing scenarios where accidents might occur, they recognize that the list of risks they have constructed is incomplete. Thus there is a need for a basis for deciding what level of resources to allocate to thinking about accident scenarios, with the everpresent possibility that "too few" scenarios will be thought about. To define "too few" we need to consider how such an evaluation "ought" to be made. Formally, we can consider the usual Bayesian formulation of rational decision-making. In this situation, one would use one's prior beliefs about accident risks and about the possible revisions of beliefs about accident risks that would come from further information gathering. With these beliefs one would estimate the gain from devoting further resources to information gathering and assessment and compare the expected benefit of that information gathering with its cost. Given the complexity in this process and the unavoidably subjective nature of prior beliefs - beliefs that may reflect a great deal of experience, but not formal evidence - it is very difficult to assess whether an individual or firm devoted "enough" resources to thinking about possible accidents. Thus, one is tempted to infer inadequate thought from the accident that happened - an inference fraught with potential error. The type of risk evaluation that would be viewed as reckless as opposed to either negligent or even nonnegligent is unclear. The usual model of negligence assumes that the agents should make a reasonable calculation and will be liable for compensatory damages if they get the calculation sufficiently wrong and act on that wrong calculation. An extension to punitive damages in this situation seems to be based on getting the calculation so wrong that additional liability is inflicted. But, if decisions to seek information are being rationally made (and prior beliefs are right), then compensatory damages are the correct incentive and the use of punitive damages will induce inefficiently excessive investment in gathering information about risks (Proposition 1 in Shavell, 1992).

It will often be difficult to collect evidence about the actual decision process on which to base an evaluation of how badly the decision was made. It is also difficult to formulate criteria for judging when a serious consideration of risks is adequate. Thus there will be a tendency to look at the accident that happened as the primary proof that the level of precaution was too low. Even in the presence of evidence about the decision process, the accident outcome may play a large role in a jury's conclusion that the decision was poorly taken. This approach to liability suffers severely from the widespread presence of hindsight bias.

Cognitive psychologists have studied hindsight bias. It is described by Fischhoff (1982) as: "In hindsight, people consistently exaggerate what could have been anticipated in foresight. They not only tend to view what has happened as having been inevitable, but also to view it as having appeared "relatively inevitable" before it happened. People believe that others should have been able to anticipate events much better than was actually the case." (Page 428.) Not only is hindsight bias common, it is also difficult to avoid. Quoting Fischhoff (1982) again: "Research on this bias has included investigations of most of the possible debiasing strategies included in the previous section. Few of these techniques have successfully reduced the hindsight bias; none has eliminated it." (Page 428.)

Belief that the outcome of an accident was both more likely and more predictable than was the case will tend to lead juries to find behavior outrageous, rather than finding the outcome to be due to bad luck. Thus, we should expect frequent errors from juries that are trying to assess whether the expectations that resulted in a level of precaution that, in turn, resulted in an accident, were badly formed, indeed outrageously formed. By making the accident outcome seem more likely and more predictable, hindsight bias works to increase the likelihood of a finding of punitive damages whatever the true nature of the underlying behavior. A further complication, in both the thinking of juries and that of potential defendants is the difficulty in thinking about risk even when the information is available. Indeed cognitive psychology has found that people do tend to make systematic errors in thinking about uncertain situations.[28] But such widespread errors are not obviously connected with outrageous behavior, nor is there reason to think that juries can identify when such errors are particularly large; hindsight bias may make juries conclude that risks were far more apparent (transparent) than they really were. Concern to avoid such cognitive errors should be part of jury instructions. Identification of flaws in the decision process should be a necessary part of assessing liability.

---

28 The problems people have with assessing risks has been a major subject in cognitive psychology; see Kahneman and Tversky (1979), Tversky and Kahneman (1992).

If rational decisions are being made about information gathering, then the risk of punitive damages, based either on underappreciation of risks at the conclusion of rational information gathering or on jury hindsight bias, will induce inefficiency where compensatory damages are adequately assessed. The analysis of the efficiency from compensatory damages, discussed in Section I, continues to apply when we include the decision as to how much information to gather along with the decision of the level of precaution. In other words, shifting back one step from the precaution decision to the information gathering decision does not alter the basic underlying logic of the analysis. For example, pharmaceutical firms need to allocate research and development resources between the pursuit of uses for new drugs (or new uses for existing drugs) and the examination of the risks associated with drug use (side effects). With punitive damages added to compensatory damages, a rational allocation of resources between these two aspects of this activity leads to too few (but safer) drugs; the presence of punitive damages in the presence of the decision of how many resources to devote to the activity of drug research and development leads to too few resources devoted to this activity. We do not formally model this circumstance.[29]

It has long been an element in discussion of legal rules that part of their purpose is to instruct agents about risks - risks that are made more salient because of their legal consequences. In the absence of undercompensation, it is unclear how relevant punitive damages are for this purpose since collecting compensatory damages in a new situation also serves to inform the public. Additional salience coming from punitive damages is not of obvious importance.

---

29 One example of such a meta analysis is the paper by Radner and Rothschild (1975). They modelled choice in a setting of limited resources of attention by management. While their focus was on management of a firm, the approach also makes some sense for individuals. The basic idea is that there are multiple issues that evolve over time and need attention if the agent is to respond to them differently than in the past. But limited attention ability implies that the agent can only attend to some of these at any time, and paying more attention to some issue involves paying less attention to others. In this framework, a conclusion of reckless disregard by a jury may be an assessment that the defendant did not allocate enough attention to this issue, compared to the attention it deserved. Assuming a rational allocation of attention, this meta analysis leaves us in the same place as the analysis above - there is a need for additional deterrence if compensatory damages are inadequate or if society does not recognize some of the utility of the defendant. In the absence of these two issues, there does not appear to be an efficiency basis for punitive damages - indeed the overdeterrence coming from punitive damages in such a setting involves both the misallocation of attention to avoid punitive damages and the underproduction of activities that are at risk of punitive damages.

We conclude that when there are not significant defects in compensatory damages, punitive damages will decrease efficiency unless there is a nonrational basis for the precaution decision. Because of potential hindsight bias by juries, it is appropriate to identify the source of the irrationality directly, not to try to infer it from the occurrence of an accident.

VII. Nonrational disregard

The assumption of rational decision-making gives a structure to the analysis of the effects of legal rules, a structure that has considerable relevance since many agents are trying to pursue their self-interest in a reasonably coherent way. However, there are situations where such a model does not adequately describe the behavior of a particular defendant. One example is some drunk driving accidents, where the drivers (at least when sober) are aware of the risks of accident, and may not be underestimating accident probabilities, but rather are proceeding to drink anyway, recognizing that later driving behavior will not be decided rationally.[30] After drunk driving accidents, juries sometimes award punitive damages, and sometimes do not. The analysis that follows considers the efficiency implications of the level of punitive damages, assuming assessment after findings of outrageous behavior; I do not explore the degree of irrationality that would make a useful distinction between outrageous and nonoutrageous irrational behavior.

It is a common human trait to hope for the best and attempt some activity without fully crediting the possibilities of accidents, whether they involve costs to the actor or costs to others. However, beyond some level or when consequences are too serious, this behavior is sometimes in the category labeled reckless disregard and made a basis for punitive damages when the costs fall on others. In this section, we consider irrational behavior that ignores accident costs that initially fall on others, but, with liability for compensatory damages, actually fall on the defendant.[31]

The analytical stance here is of wishful thinking, not a deliberate decision to ignore accident costs, which would seem a

---

30 There is an awkwardness in the fit between punitive damages for drunk driving and judicial descriptions that the risk be a probable outcome of the conduct of the defendant. While statistics are not gathered about the extent of driving by drunks that does not result in accidents, no doubt there is a great deal. Thus, in any single episode, there may well be a low probability of an accident, not a high one, although with frequent drunk driving there may be a high aggregate probability of an accident.
31 We ignore the situation where the defendant is judgment proof, in which case punitive damages don't add to incentives.

more suitable description of the behavior analyzed in Section III, where costs remaining on others were ignored.[32] We assume that some risk is irrationally ignored, it is treated as too small to warrant attention even though it is important enough to warrant attention, and attention to the ignored risk would change precautionary behavior. Thus there is an efficiency cost to this inadequate attention to risk. If the imposition of punitive damages reduces the likelihood of ignoring this risk, then the inefficiency from this inadequate attention can be decreased. If punitive damages did not alter the likelihood of attending to such a risk, then there would be no deterrence basis for having them in this situation. We continue the analysis assuming that there is an impact on the likelihood of rationality.

It seems useful to have a crisp model of this sort of behavior in order to explore the link between punitive damages and efficiency. Assume that when a defendant acts with nonrational disregard, the consequences of an accident are ignored. This results in some probability of avoiding an accident which we denote p^n and the resulting utility (gross of liability costs) of $u[p^n]$. We assume that p^n optimizes $u[p]$ - it would be optimal if there were no accident liability. That is, we do not assume that people act randomly; their behavior is systematically trying to generate utility, but utility is misspecified.

We continue to assume that when the defendant acts with attention to accident risks, then the probability evaluations are correct, the decision-making is rational. Let q denote the probability that the defendant acts rationally and thus 1-q is the probability that the defendant acts with nonrational disregard. That is, with probability 1-q, the defendant has a probability of avoiding an accident p^n. And we assume that with probability q the defendant acts in a deliberate way, maximizing utility net of legal costs, as in the analysis above. In order to describe this behavior we need to describe the workings of the legal system. For simplicity, we assume that all accidents from this activity lead to liability for compensatory damages and that the defendant bears all of the costs of the accident as compensatory damages. The central new issue to be modeled is the extent to which the legal system can distinguish between accidents happening after nonrational disregard and accidents happening despite a rational and accurate evaluation by the defendant of the risks involved. For example, driving after a little bit of alcohol consumption is legal and, if the driving is careful, is not negligent, although even a low level of alcohol may affect the probability of an accident. That is, the drinking may not have resulted in negligent driving, but may have reduced the probability of avoiding an accident.

---

32 In terms of the instructions in footnote 2, this seems closer to "knowledge of facts which would disclose the danger to any reasonable person" than to "knowledge of serious danger to others." However, the fit is not tight since one could irrationally ignore a knowledge of danger as well as a knowledge of facts.

Driving after more alcohol consumption raises the sort of efficiency question examined here, but there is not a bright line separating the two sorts of decisions. We assume that all accidents occurring after the care chosen with nonrational disregard result in punitive damages, while a fraction f of accidents that happen after an accurate evaluation of risks also result (erroneously) in punitive damages. This failure to perfectly discriminate circumstances is realistic and central to the concerns analyzed.[33] When a rational defendant is at risk of punitive damages, the incentive for precaution will be too large, generating inefficiency. This inefficiency offsets the efficiency gain from lessening the likelihood of nonrational disregard. If there were no impact of punitive damages on the probability of nonrational disregard, there would be no efficiency case for punitive damages.[34]

Returning to the formal model, we assume that compensatory damages equal accident costs, A. For convenience, we also assume that there are no fines, although behavior such as drunk driving resulting in accidents does generate fines - it is just that incorporating such fines in the analysis does not change the thrust of the argument, it just lowers the amount of punitive damages that would be appropriate if there were zero fines. When doing an accurate evaluation, the defendant maximizes utility as given by

$$(18) \quad U[p] = u[p] - (1-p)(A + fP).$$

where f is the probability of being held liable for punitive damages after an accident, despite having chosen a level of precaution with an accurate evaluation of risks. Thus f is a measure of the mistakes made in the assessment of punitive damages. We denote by p^r the level of precaution chosen in this circumstance. Note that the level of precaution with accurate evaluation, p^r, depends on the level of punitive damages, P, since there is a risk of punitive damages even when due precaution has been taken. In contrast, the level of precaution when behaving

_____

33 For a discussion of erroneous assessment of punitive damages, see Ellis, 1982.
34 An alternative way to proceed that would lead to a similar analysis would be to consider an earlier decision point about engaging in the activity at all (or how frequently to engage). If this early decision is rational and recognizes that sometimes there will be nonrational disregard later in carrying out the activity, then we would again have a balancing between discouraging nonrational disregard and discouraging the activity beyond the point that comes from its bearing its full social costs. Both of these models appear to capture realistic concerns about punitive damages that result in inefficiency. Given the similarity in their structure relative to efficiency, only the earlier formulation is developed formally here. The other form of model will be considered in the situation of vicarious liability of corporations in Diamond, 1997b.

with nonrational disregard, p^n, does not depend on the level of punitive damages since the possibility of an accident (and so liability for both compensatory and punitive damages) is ignored. Thus we write the function $p^r[A+fP]$, but treat p^n as a scalar, not a function. This is illustrated in Figure 2, which illustrates the assumption that the role of punitive damages is to switch some agents from ignoring a risk to evaluating it properly, not to affect the precaution of those who are not evaluating the risk appropriately.

In other words, we are describing nonrational behavior as a two-step process. The first step is whether a defendant behaves appropriately - evaluating the risks at hand and making a suitable precaution decision. The probability of appropriate behavior is affected by punitive damages.[35] In the second step, if the defendant is not behaving appropriately, there is no impact at all from the possibility of punitive damages. Let us consider the two steps in terms of the drunk driving example. Someone who assesses risk accurately will call a taxi rather than driving after having drunk enough, with the level of drinking warranting a taxi influenced by the risk of accident and the risk of punitive damages. Someone who nonrationally disregards the risk will drive no matter how much has been drunk. Punitive damages are assumed to change the probability that the driver will assess the risks at all and then make a conscious choice between driving and calling a taxi. Thus the probability of risk assessment and the behavior if there is a risk assessment are both influenced by the possibility of punitive damages, but the behavior of someone who does not assess the risks is not influenced.

Behavior conditional on an accurate assessment of risks is given by the function $p^r[A+fP]$, which is defined implicitly by the first order condition for individual utility maximization. Differentiating (18) with respect to p and setting the derivative equal to zero, we have the condition that implicitly defines p^r:

(19) $u'[p^r] + A + fP = 0.$

Since $u[p]$ is concave in p, p^r is increasing in P as long as f>0. That is, an increase in punitive damages increases precaution, lowering efficiency. We also assume that the rate of increase in precaution with the level of punitive damages is decreasing (over the relevant range). Thus, we are assuming $p^{r'}>0$ and $p^{r''}<0$.[36]

---

35 The possibility of mistakenly assessing punitive damages despite rational behavior presumably lowers the impact of punitive damages on the likelihood of nonrational disregard, with this effect incorporated in the function q.

36 We can express these derivatives in terms of the utility function: $u''p^{r'} = -1$; $u''p^{r''} + u'''(p^{r'})^2 = 0$. The concavity of u ($u''<0$) implies that $p^{r'}>0$. A necessary and sufficient condition for $p^{r''}<0$ is $u'''<0$.
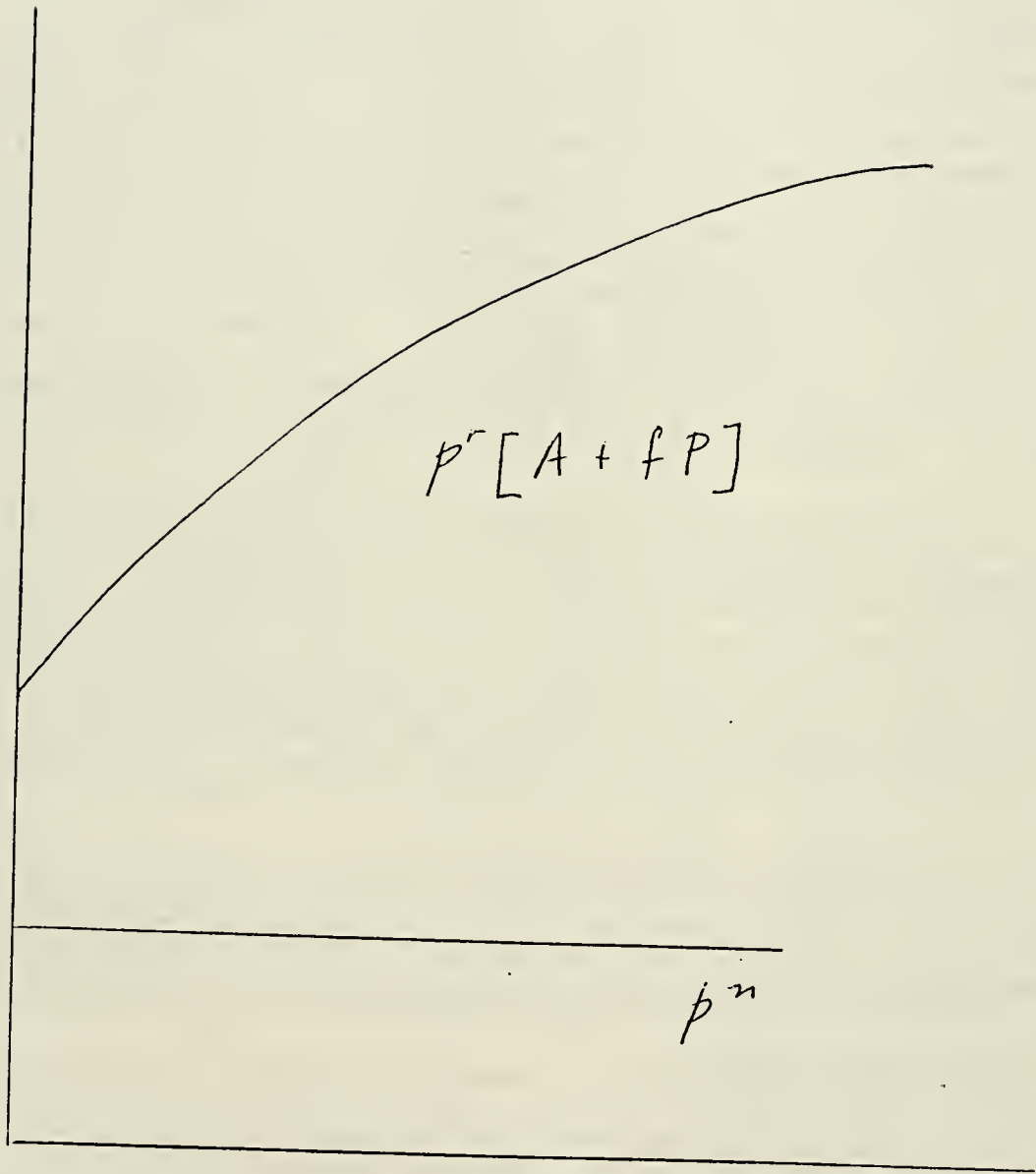
Figure 2

The efficiency purpose of punitive damages is to direct attention to the possibility of an accident and its consequences, in order to reduce the probability of nonrational disregard. Indeed, sending a message to deter ignoring this risk appears to be in the mind of juries. There are two simple ways that we can we write the probability of nonrational disregard as a decreasing function of damages - one is to relate $1-q$ to punitive damages, P, the other is to relate $1-q$ to the total level of damages, compensatory and punitive, $A+P$, and fines, if any. For analysis with a representative agent facing a given level of compensatory damages, the two formulations are equivalent. However, for considering punitive damages levied on different defendants, the two formulations are slightly different. If we ask how punitive damages should vary with the level of compensatory damages, the presence or absence of A in the determination of $1-q$ will make a difference. Similarly, when we consider how punitive damages should vary with the wealth of the defendant, the formulation again makes a difference. Rather than picking one particular form, we use the more general structure, making q an increasing function of both A and P, $q[A, P]$.[37] In order to have an efficiency role for punitive damages, we assume that nonrational disregard is deterred by punitive damages, $q_P>0$, but that the effect of a marginal dollar gets smaller as the level of punitive damages increases, $q_{PP}<0$.[38] This is illustrated in Figure 3.

The function $q[A, P]$ reflects behavioral responses to the attention drawn by punitive damages, not a rational calculation, since such rationality appears to be inconsistent with the concept of nonrational disregard. We assume that there is no direct resource cost to the change in the level of regard to this decision. That is, we assume that the individual either does or does not make a rational decision about the level of precaution. We ignore investments undertaken to affect the defendant's ability to make rational decisions, such as hiring a driver or participating in a program to learn better self-control.

We can now write social welfare as a function of the level of punitive damages by recognizing two possibilities of individual behavior - nonrational disregard and accurate evaluation - and two possibilities of jury behavior - accurate assessment of nonrational disregard and mistaken assessment of accurate evaluation as nonrational behavior[39]:

---

37 This formulation does not necessarily assume that punitive damages are assessed whenever there is an accident caused by someone with precaution level p^n; whatever the likelihood of assessment is incorporated into the function $q[A, P]$.
38 For example, the fraction showing nonrational reckless disregard might shrink toward an irreducible minimum as P rises without limit. An example of a function with this property is $q[A, P] = k_0 - k_1/(1+A+P)$, for some constants k_i.
39 We assume that, over the relevant range, $W[P]$ is concave in P. Thus there will be a unique optimal level of punitive damages, given as the unique solution to the first order condition for maximization of W.

Probability
of rational
decision

q [A, P]
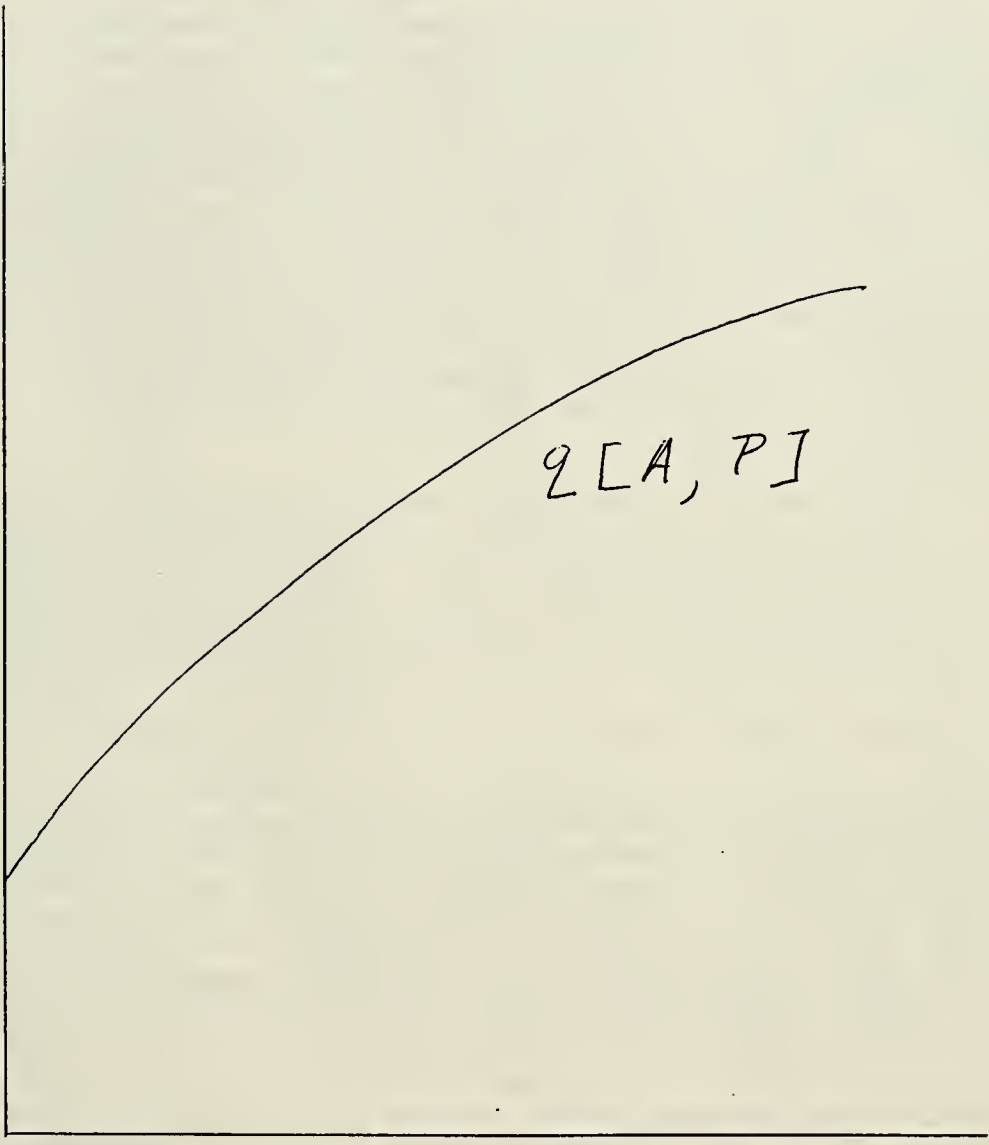
Punitive damages

Figure 3

(20) $W[P] = (1-q[A, P])\{u[p^n] - (1-p^n)A\}$
                  $+ q[A, P]\{u[p^r[A+fP]] - (1-p^r[A+fP])A\}.$

That is, social welfare is the value of the utility of the activity
less the expected accident costs, summed over both types of
defendants. In this formulation, we continue to assume that there
are no direct social costs from the cash transfer as a result of
punitive damages - the only social concern with the payment of
punitive damages is associated with the incentives for the
defendant. The solution to this problem will balance the
improved efficiency from reducing nonrational disregard with the
decreased efficiency from overdeterrence in the selection of
precaution by those who are not recklessly disregarding the risk of
accident. The balancing will depend on several factors including
the responsiveness of nonrational disregard to punitive damages,
the importance of the difference in precaution between proper
evaluation of risk and nonrational disregard, and the
responsiveness of the level of precaution of those not recklessly
disregarding the risk.

     In order to have the first order condition in an intuitive
form, we first define the social gain from having the defendant
make an accurate assessment of the risk, rather than showing
nonrational disregard. This social gain depends on the level of
punitive damages since the level of punitive damages affects the
chosen level of precaution. We denote this gain by G[P]:

(21) $G[P] = \{u[p^r[A+fP]] - (1-p^r[A+fP])A\} - \{u[p^n] - (1-p^n)A\}.$

Thus the social gain compares the gross utility from the activity
of the defendant less the expected accident cost when there is a
rational level of precaution with that when there is a nonrational
level of precaution. When punitive damages are increased, this
gain will decrease since punitive damages do not affect the
behavior of those who remain nonrationally reckless, while
lowering the efficiency of those who take precaution. The
lowered efficiency comes from the expected punitive damages
falling on a rational defendant, fP, which acts like a distortionary
tax on this decision. The importance of this distortionary tax
depends on the responsiveness of the precaution decision to
financial decisions. Thus, as shown in Section IX, the derivative
of the social gain satisfies:

(22) $G'[P] = -(fP)fp^{r\prime}[A+fP] < 0.$

Increasing punitive damages increases expected punitive damages
by f times the increase. The increase in expected punitive
damages increases precaution by $p^{r\prime}$ times the amount of increase.
The social cost of increasing precaution is fP times the amount of
increase, since fP is equivalent to the size of a distorting tax on
the precaution decision.

Bearing in mind this structure of the social gain from avoiding nonrational disregard, the first order condition for the optimal level of punitive damages can be written as:[40]

(23) $-q_P[A, P]G[P] = q[A, P]G'[P]$.

When punitive damages are increased, there will be less nonrational disregard, with the magnitude given by the derivative of the fraction showing nonrational disregard, 1-q, with respect to the level of punitive damages. The social gain per person deterred is given by the gain function G[P]. The second effect of raising punitive damages is to increase the precaution of potential defendants who are rational. Since these defendants would be taking the socially efficient level of precaution in the absence of punitive damages, any increase in punitive damages lowers the efficiency of the chosen level of precaution. The first order condition balances these two effects - the efficiency gain from encouraging more people to make accurate assessments and the efficiency loss from overdeterring the people who are making accurate assessments. Among other elements, the balance in equation (23) depends on the ability to distinguish between the rational and nonrational decision processes, modeled here as an error rate of f after rational decision-making. That is, all of the terms in equation (23) depend on the error rate of juries, as measured by f. When the error rate is large enough, the level of punitive damages satisfying this first order condition may be small enough so that it is not worth the trouble and cost to have punitive damages. This possibility is more important if there are also civil or criminal penalties that may be mistakenly applied. Sometimes, as with drunk drivers, there will be direct evidence about decision processes and not just the outcome of an accident. In other situations, particularly corporate situations, it will sometimes be difficult to distinguish between these cases. The presence of hindsight bias, discussed above, adds to the risk of inappropriate assessment of punitive damages. The first order condition makes it clear that it is insufficient to just look to deterring the nonrational when setting punitive damages; the marginal gain from this deterrence should equal the marginal loss from overdeterrence if the jury sometimes makes mistakes.

VIII. Variation of punitive damages with wealth

It is common for plaintiffs to argue that punitive damages should be proportional to the wealth of the defendant. To consider the efficiency implications of this practice, let us consider the possibility that the likelihood of nonrational disregard does vary with income or wealth. Let us denote by Y the measure of income or wealth that is used in assessing punitive damages, and refer to it as wealth. If the wealth of the defendant did not enter the behavioral elements in (20), then there would be

---

40 This equation is a special case of the derivation in Section IX.

no deterrence reason to vary the level of punitive damages with the wealth of the defendant. Moreover, rational choice of a level of precaution depends on the cost of an accident, not the cost relative to wealth. However, it is possible that the probability of nonrational disregard does vary with the wealth of the defendant. While there are many channels of difference (for example the variation of both wealth and probability of nonrational disregard with personal traits and previous decisions, such as education) we want to focus on the role of wealth in affecting the impact of punitive damages on nonrational disregard, ignoring the other channels by which wealth affects or is correlated with the probability of nonrational disregard. In particular, to review the argument that a constant fraction of the wealth of the defendant is an appropriate level of punitive damages, we make the strong assumption that the role of wealth is to scale proportionally the impact of punitive damages on the probability of nonrational disregard. This assumption ignores many ways in which wealth influences behavior. Yet this assumption is implicit in arguments that it is appropriate to scale punitive damages proportionally to wealth on the grounds that proportionality is needed to achieve the same level of deterrence for people with different levels of wealth. The purpose of analyzing this situation is to ask: _if_ one assumes that proportionality is the correct scaling for influencing the probability of nonrational disregard, does it then follow that punitive damages ought to be scaled proportionally? That is, we assume that scaling punitive damages proportionally to the wealth of the defendant results in the same level of nonrational disregard for defendants with different levels of wealth; our purpose is to examine the implications for efficiency of this strong assumption.

We continue to write care, assuming correct evaluation, p^r, as a function of $A+fP$, not $(A+fP)/Y$. Having p^r depend on $A+fP$ follows from the utility maximization problem we have used to describe rational decision makers. For example, this would be the case for a profit maximizing corporation.

Scaling punitive damages in this way, we write the social welfare associated with the decisions of individuals of wealth level Y as

$$(24) \quad W[P; Y] = (1-q[A, P/Y])\{u[p^n] - (1-p^n)A\} \\ + q[A, P/Y]\{u[p^r[A+fP]] - (1-p^r[A+fP])A\}.$$

This has exactly the same form as equation (20) above. For overall social welfare, we sum this expression over all individuals of different wealth levels. The central analytical question is how optimal punitive damages vary with the wealth of the defendant. In designing incentives, we need to balance the incentives that depend on the ratio of punitive damages to wealth with the incentives that depend on the absolute level of punitive damages.[41]

_____

41 There would also be a, presumably small, efficiency cost from implicitly taxing wealth since the same activity results in higher damages for people who saved more or earn more. Similarly, there would be an efficiency cost for corporations seeking an optimal size with this implicit tax on being larger.

That is, even with the attention-drawing role of punitive damages proportional to wealth, it does not follow that the optimal setting of punitive damages has total damages proportional to wealth. The absence of this implication comes from the fact that the inefficiency associated with the correctly perceived risk of punitive damages does not vary with wealth. Rather, like any distorting tax, the magnitude of the tax relative to the economic value of the activity is important, not its size relative to the wealth of the actor. Once we are in the realm of imperfect correction of decisions, the realm of second-best analysis, then additional elements in the reaction to incentives become relevant. Specific results about punitive damages do not have great generality - many patterns are possible. The mathematics that follows (in section IX, which can be skipped with no loss of continuity) shows that in this situation and with these assumptions, the optimal level of punitive damages for nonrational disregard does generally vary with wealth. While it is plausible that optimal punitive damages increase with wealth, an increase is not a necessary conclusion. When optimal punitive damages do increase with wealth, it is plausible, but again not necessary that they increase less than in proportion to wealth.[42] The many factors that enter into the first order conditions make it difficult to reach any firm conclusions without substantial empirical work. One would need to know about the behavioral responses to punitive damages of both rational and nonrational potential defendants and how these behavioral responses varied with wealth.

We know very little about the impact of punitive damages on the behavior of people who might nonrationally recklessly disregard a risk - for example, the extent to which drunk driving is deterred by adding punitive damages to compensatory ones; nor would we think that a jury has particular insight into this impact. Further, as long as it is hard to distinguish between cases where there was nonrational disregard followed by an accident and cases where there was proper calculation followed by an accident, we will be overdeterring some activities, with an efficiency cost.

The admissibility of the wealth of defendants has been controversial. This analysis suggests that if admissibility is continued, the court should limit the kind of arguments that are made about wealth. In particular, allowing arguments suggesting that proportionality to wealth is a reasonable basis for setting punitive damages lacks justification.

IX. Derivation of optimal punitive damages

To examine the variation in the optimal level of punitive

---

42 If the probability of reckless disregard had been written as $q[(A+P)/Y]$, then the result would apply to A+P increasing less than in proportion to Y rather than P increasing less than in proportion to Y.

damages with the wealth of the defendant, we maximize $W[P; Y]$ with respect to P and then examine the variation of the optimal P with Y. First, we restate the social welfare function:

(24) $W[P; Y] = (1-q[A, P/Y])\{u[p^n] - (1-p^n)A\}$
$+ q[A, P/Y]\{u[p^r[A+fP]] - (1-p^r[A+fP])A\}.$

Differentiating (24) with respect to P, the first order condition for the maximization of W is:

(25) $q_P[A, P/Y]\{u[p^r[A+fP]] - (1-p^r[A+fP])A - \{u[p^n] - (1-p^n)A\}\}/Y$
$+ q[A, P/Y]\{u'[p^r[A+fP]] + A\}fp^{r\prime}[A+fP] = 0.$

It is convenient to state this in terms of the social gain from having someone do an accurate evaluation of risks rather than behave with nonrational disregard. We denote this gain by $G[P]$:

(26) $G[P] = u[p^r[A+fP]] - (1-p^r[A+fP])A - \{u[p^n] - (1-p^n)A\};$

Differentiating (26) and using the individual first order condition, (19), and the properties of p_r in footnote 34, we have:

(27) $G'[P] = (u'[p^r[A+fP]] + A)fp^{r\prime}[A+fP]$
$= -(fP)fp^{r\prime}[A+fP] < 0.$
$G''[P] = u''(fp^{r\prime})^2 + (u' + A)f^2 p^{r\prime\prime}$
$= u''(fp^{r\prime})^2 - (fP)f^2 p^{r\prime\prime}$
$= u''(fp^{r\prime})^2 + (fP)f^2\{u'''(p^{r\prime})^2/u''\}.$
$= -f^2 p^{r\prime} - (fP)f^2 p^{r\prime\prime}.$

The sign of G'' is ambiguous with our assumption that $p^{r\prime\prime}<0$ (that is, $u'''<0$).

Using this social gain function, we can rewrite the social first order condition (25) in the more convenient form:

(28) $q_P[A, P/Y]G[P]/Y + q[A, P/Y]G'[P] = 0.$

Since $G'<0$ and $q_P>0$, at an optimum $G>0$.

Differentiating (28) implicitly, we have:

(29) $\dfrac{dP}{dY} = \dfrac{(Pq_{PP}G/Y + Pq_P G' + q_P G)/Y^2}{q_{PP}G/Y^2 + 2q_P G'/Y + qG''}.$

From the assumed concavity of W in P, the denominator is negative. We note that the first two terms in the numerator are negative, but the third term is positive. Thus it is not necessarily the case that punitive damages should increase with wealth. When the derivative is positive, the third term is smaller in absolute value than the sum of the first two.

Converting (29) into an elasticity, we have

$$(30) \quad \frac{Y}{P} \frac{dP}{dY} = \frac{q_{PP}G/Y^2 + q_PG'/Y + q_PG/(PY)}{q_{PP}G/Y^2 + 2q_PG'/Y + qG''}.$$

If G'' is negative or, more generally, sufficiently small, then the elasticity is less than one when it is positive. If G'' is positive, then it would have to be sufficiently large to offset both the positive sign of the third term in the numerator and the presence of the multiplicative factor of 2 in the second term in the denominator if it were to make the elasticity larger than one.

X. Conclusions

This paper examined equilibrium in a variety of situations where punitive damages might be assessed after an interaction between individuals who are not in a contractual relationship. The focus was on the efficiency implications of alternative rules for determining the level of punitive damages. The paper takes as given the set of cases where punitive damages are assessed. It examines the usefulness of dividing those cases between malicious intent cases and reckless disregard cases. It then examines the usefulness of subdividing the reckless disregard cases between rational and nonrational reckless disregard. Several conclusions follow from the analysis.

Guidance for juries and judicial oversight do not explicitly differentiate between situations of malicious intent and situations of reckless disregard. In both cases, juries are told to select a damage amount that is suitable for deterrence and punishment. In contrast, some of the caps legislated by states do differentiate between malicious intent and reckless disregard. For example, Oklahoma has different caps on damage judgments in the two types of cases - with reckless disregard, the cap is the greater of $100,000 or actual damages, while with an intentional and malicious act, the cap is the greatest of $500,000, twice actual damages, or the benefit accruing to the defendant. This pattern of legislative caps has received some judicial approval. For example, Koenig and Rustad (1993) have observed: "The West Virginia Supreme Court endorsed caps on punitive damage awards in cases in which the defendant behaved stupidly, but felt there should be no cap when the defendant had been "really mean" (419 S. W.2d 870, 887 [W.Va. 1992])." (Fn. 30, p. 30.) Based on the analysis in this paper, it seems useful to move away from the current undifferentiated treatment of different bases for punitive damages and toward greater structure in placing the issue before juries.

The analysis in this paper strongly supports the importance of distinguishing between malicious intent and reckless disregard. In particular, in the presence of malicious intent the court should examine the defendant while selecting a suitable punitive amount

to deter behavior that generates benefits to the defendant which are not suitable for inclusion in social accounting. The focus is precisely on the preferences (including any potential for ill-gotten gains) of the defendant. 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. The inadequate attention could come from a rational disregard of costs that the tort system will not assess or from a nonrational disregard of the costs that the tort system will assess. 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. This focus is naturally stated in terms of a ratio of punitive to compensatory damages, or equivalently the ratio of the sum of compensatory and punitive damages to the level of compensatory damages. The court should determine what multiple of the actual costs of the plaintiff would be sufficient for the defendant to pay appropriate attention to these costs. The goal should be to pay appropriate attention to accident costs, not to have a zero accident risk. There is no basis for such a comparison of punitive and compensatory damages in the situation of malicious intent. Thus, procedurally, there should be different instructions to juries and different standards for judicial oversight in these two different situations.

The analysis of reckless disregard distinguished two different types of behavior by defendants - referred to as rational and nonrational. Behavior that pays attention to one's own costs but not social costs is rational and might be viewed as outrageous when the costs falling on others is large. This situation can be identified, to some extent, by inquiring directly into the likelihood of and likely magnitudes of inadequacies in compensatory damages. This then leads, in a straightforward way, to a measure of punitive damages suitable for improving efficiency, again one that keys off the amount of compensatory damages. One also needs to recognize civil and criminal fines that are part of the expected costs of the defendant, and should be deducted from the punitive damages that would give efficient incentives with zero fines.

A separate circumstance where a jury might find a defendant has behaved outrageously is where the defendant acts as if the risk of an accident were much lower than the jury thinks was an appropriate measure. For punitive damages, this misreading of risk must be sufficiently large as to be not only negligent but grossly so, outrageously so. This approach to punitive damages was argued to be unsatisfactory on two grounds. One is that hindsight bias makes the risk of inappropriately assessing punitive damages very large. Second, the act of gathering information about risks is something that itself is subject to decisions, decisions that will involve some mistakes when they are made rationally. The costs of perfect decision making, indeed the impossibility of perfect decision-making, makes this an area where the actual occurrence of a serious accident may well lead to

erroneous interpretations about previous behavior. Moreover, insofar as decisions are being made rationally, any punitive damages will cause overdeterrence. Optimal deterrence for rational decision-makers involves imposition of the costs falling on others; any larger level of damages is overdeterrence. While overdeterrence will reduce accident probabilities, it will raise total costs and lower efficiency.

In the absence of malice, the court should look for evidence of the likelihood of underassessment or evidence of irrational decision-making. In the absence of both elements, punitive damages should not be allowed. That is, if all decision-making were rational, it would be hard to find a basis for inferring outrageous behavior when compensatory damages were roughly adequate. Thus the paper entered the somewhat slippery domain of nonrational decision-making, recognizing that it will sometimes be the case that decisions are made which omit rational benefit cost analysis, indeed which avoid it. Drunk driving was taken to be an example of such behavior. It is difficult to design a suitable measure of additional damages for deterrence in this situation. With rational decision-making, one can look to omitted costs as the correction that will then be used in a rational way in making decisions. With nonrational decision-making, one is in the realm of trying to decide how large an impact an expansion of actual costs would have on a nonrational decision-maker. Such analysis has less of an anchor in studies of behavior. Moreover, while any application of punitive damages where it is not appropriate lowers efficiency, this may be particularly a problem in this situation. That is, with malicious intent, one looks to evidence of malice. With nonrational decision-making one should look for direct evidence of a nonrational decision process, although it may sometimes be hard to distinguish rational from nonrational decision-making, harder than to distinguish malice from an accident. The presence of hindsight bias underlines the difficulty of accurate assessment of decision-making. Thus the multiple of compensatory damages suitable for deterrence needs to be tempered for the risk of error in this situation. While overdeterring potential malicious acts is not a concern, encouraging excessive precaution lowers economic efficiency.

Beyond distinguishing different bases for a possible determination of outrageous behavior, the paper examined the efficiency basis for varying punitive damages with the wealth of the defendant. No basis was found in the situation of inadequate compensatory damages. Bases were studied in the (different) situations of malice and of nonrational disregard. However, it was pointed out that varying damages with wealth represented a tax on behavior where the additional deterrence was not needed. In a situation of nonrational disregard, the second best nature of the impact of punitive damages on efficiency when both rational and nonrational decision-makers are present makes for a complex interaction of efficiency considerations with equilibrium. Nevertheless, it seemed plausible that even if there was a case for

varying punitive damages with wealth, they should vary less than in proportion to wealth. This suggests that arguments should not be allowed that implicitly assume that the correct punishment is proportional to wealth. In the case of reckless disregard, a focus on multiples of compensatory damages, as recommended above, would leave room for varying the multiple with wealth but would appropriately discourage the award of punitive damages simply as a proportion of wealth.

This paper has focused solely on deterrence. Later papers will discuss the integration of punishment and deterrence concerns (1997a) and will address these issues in the corporate situation, exploring the implications of vicarious liability (1997b).

References

Abraham, Kenneth S., and John C. Jeffries, Jr., Punitive Damages, 18 Journal of Legal Studies, 415-425, 1989.

Baker, C. Edwin, 1978, Counting Preferences in Collective Choice Situations, 25 UCLA Law Review, 381-416.

Boston, Gerald W., 1993, Punitive Damages in Tort Law, Deerfield IL: Clark, Boardman, Callaghan.

Calfee, John E. and Richard Craswell, 1984, Some Effects of Uncertainty on Compliance with Legal Standards, 70 Virginia Law Review, 965-1004.

Cooter, Robert and Thomas Ulen, 1997, Law and Economics, second edition, Reading, Ma.: Addison-Wesley.

Diamond, Peter, 1973, Consumption Externalities and Imperfect Corrective Pricing, Bell Journal of Economics and Management Science, 4 (2), Autumn, 526-538.

Diamond, Peter, 1997a, Integrating Punishment and Efficiency Concerns in Punitive Damages for Reckless Disregard of Risks to Others.

Diamond, Peter, 1997b, Vicarious Liability for Punitive Damages, forthcoming.

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

Fischhoff, Baruch, 1982, Debiasing, Ch. 31 in D. Kahneman, P. Slovic and A. Tversky (eds.), Judgment under Uncertainty: Heuristics and Biases, Cambridge: Cambridge University Press, 422-444.

Kahneman, Daniel, and Amos Tversky, 1979, Prospect Theory: An Analysis of Decisions under Risk, Econometrica, 47, 263-291.

Koenig, Thomas, and Michael Rustad, 1993, The Quiet Revolution Revisited: An Empirical Study of the Impact of State Tort Reform of Punitive Damages in Products Liability, 16 The Justice System Journal, 2, 21-44.

Owen, David G., 1982, Civil Punishment and the Public Good, 56 Southern California Law Review, 1, 103-122.

Radner, Roy, and Michael Rothschild, On the Allocation of Effort, Journal of Economic Theory v10 n3 June 1975, pp. 358-76.

Schwartz, Gary T., 1979, Economics, Wealth Distribution, and Justice, Wisconsin Law Review, 799-813.

Shavell, Steven, 1981, A Note on Efficiency vs. Distributional Equity in Legal Rulemaking: Should Distributional Equity Matter Given Optimal Income Taxation?, American Economic Review, 71, 2, May, 414-18.

Shavell, Steven, 1985, Criminal Law and the Optimal Use of Nonmonetary Sanctions as a Deterrent, Columbia Law Review, 85, 6, 1232-1262.

Shavell, Steven, 1992, Liability and the Incentive to Obtain Information about Risk, Journal of Legal Studies, 21 (2), 259-270.

Shavell, Steven, 1997, The Fundamental Divergence Between the Private and the Social Motive to Use the Legal System, Journal of Legal Studies.

Tamerin, J. S. and H. L. P. Resnik, 1972, Risk taking by individual option - Case study - Cigarette smoking, in Perspectives on Benefit Risk Decision Making, pp. 73-84, Washington D. C.: National Academy of Engineering; reprinted in P. Diamond and M. Rothschild, Uncertainty in Economics: Readings and Exercises, 1989, San Diego: Academic Press.

Tversky, Amos, and Daniel Kahneman, 1992, Advances in Prospect Theory: Cumulative Representation of Uncertainty, Journal of Risk and Uncertainty, 5, 297-323.

7000 570