ISSN 1045-6333

# PUNITIVE DAMAGES: AN ECONOMIC ANALYSIS

A. Mitchell Polinsky*
Steven Shavell**

Discussion Paper No. 212

7/97

Harvard Law School
Cambridge, MA 02138

The Center for Law, Economics, and Business is supported by a grant from the John M. Olin Foundation.

*Professor of Law and Economics, Stanford Law School.
**Professor of Law and Economics, Harvard Law School.

Forthcoming:
111 *Harvard Law Review* --- (1998)

# PUNITIVE DAMAGES: AN ECONOMIC ANALYSIS

**A. Mitchell Polinsky**

**and**

**Steven Shavell***

---

* Polinsky is the Josephine Scott Crocker Professor of Law and Economics at Stanford Law School; Shavell is Professor of Law and Economics at Harvard Law School. Research on this Article was supported by the Exxon Corporation. (The Exxon Corporation is appealing a punitive damage judgment against it in the *Exxon Valdez* oil spill litigation; we have served as consultants to the Exxon Corporation in connection with this litigation.) The views expressed below are our own and do not necessarily represent those of the Exxon Corporation. We received helpful comments on an earlier draft of this Article from Janet Alexander, Daniel Capra, Richard Craswell, Aaron Edlin, Jill Fisch, Jesse Fried, Victor Fuchs, Robert Hall, Louis Kaplow, Daniel Klerman, Thomas Krattenmaker, Richard McAdams, Kenneth Simons, and participants at seminars given at a number of law schools. At various stages in the preparation of this Article, we also benefitted from research assistance from Jennifer Brosnahan, Patrick O'Reilly, Janine Scancarelli, Debra Volland, and Sherri Lynn Wolson.

JEL Class K13

# PUNITIVE DAMAGES: AN ECONOMIC ANALYSIS

A. Mitchell Polinsky*
Steven Shavell**

*Abstract*

Punitive damages law is one of the more controversial features of the American legal system. Trial and appellate courts have been struggling for many years to develop a coherent set of principles for assessing when punitive damages should be awarded, and at what level. In this Article Professors Polinsky and Shavell use economic reasoning to provide a relatively simple set of principles for answering these questions, given the goals of deterrence and punishment. With respect to the deterrence objective, upon which their Article focuses, their main point is that punitive damages ordinarily should be awarded if, but only if, an injurer has a significant chance of escaping liability for the harm he caused. When this condition holds, punitive damages are needed to offset the deterrence-diluting effect of the chance of escaping liability. (They mention as well a deterrence rationale for punitive damages that does not rest on the possibility of escape from liability -- that punitive damages may be needed to remove the socially illicit gains that individuals obtain from malicious acts.) Professors Polinsky and Shavell also discuss the tension between the implications of the deterrence objective and present punitive damage law, including the law's emphasis on the reprehensibility of a defendant's conduct and on a defendant's wealth. With respect to the punishment objective, Professors Polinsky and Shavell stress that the imposition of punitive damages on corporations may fail to serve its intended purpose (although the imposition of punitive damages on individual defendants accomplishes punishment in a straightforward manner). This is primarily because the payment of punitive damage awards by corporations often does not lead to greater punishment of culpable employees, but instead punishes the corporation's shareholders and customers.

*Professor of Law and Economics, Stanford Law School.
**Professor of Law and Economics, Harvard Law School.

TABLE OF CONTENTS

## PUNITIVE DAMAGES: AN ECONOMIC ANALYSIS

## I. INTRODUCTION

One of the more controversial features of the American legal system is the law governing punitive damages. The courts have been struggling for many years to develop a rational set of principles for the imposition of punitive damages,[1] legislative bodies have passed or have considered a variety of statutes to remedy perceived problems with punitive damages,[2] academic commentators have debated the theory and significance of punitive

---

[1] The U.S. Supreme Court's struggle to develop coherent principles is exemplified by the skeptical views about the rationality of punitive damages that various Justices have expressed. Most recently, in *BMW of North America, Inc. v. Gore*, 116 S. Ct. 1589 (1996), Justice Scalia, in a pointed dissent, described the majority's guideposts for assessing punitive damages as "provid[ing] no real guidance at all." *Id.* at 1613; *see also TXO Prod. Corp. v. Alliance Resources Corp.*, 509 U.S. 443, 475 (1993) (O'Connor, J., dissenting) ("[T]he lack of clear guidance heightens the risk that arbitrariness, passion, or bias will replace dispassionate deliberation as the basis for the jury's verdict."); *id.* at 466-67 (Kennedy, J., concurring in part and dissenting in part) (arguing that the Court's vague formulation of a "reasonableness" standard for punitive damage awards is unsatisfactory and that "[t]his type of review, far from imposing meaningful, law-like restraints on jury excess, could become as fickle as the process it is designed to superintend. Furthermore, it might give the illusion of judicial certainty where none in fact exists, and, in so doing, discourage legislative intervention that might prevent unjust punitive awards"); *Browning-Ferris Indus., Inc. v. Kelco Disposal, Inc.*, 492 U.S. 257, 281 (1989) (Brennan, J., concurring) ("Without statutory (or at least common-law) standards for the determination of how large an award of punitive damages is appropriate in a given case, juries are left largely to themselves in making this important, and potentially devastating, decision."). Justice Brennan also noted that the instructions typically given to jurors, which advise them to consider the character and wealth of the defendant and the nature of the defendant's conduct, provide guidance that is "scarcely better than no guidance at all." *Id.*

[2] The statutes that we are referring to generally are aimed at reducing the number of punitive damages claims and the level of punitive damages awards. Notably, many states have enacted legislation imposing limits on the magnitude of punitive damages awards. Another type of statutory reform implemented by several states is the payment of a fraction of punitive damages to a state agency rather than to the plaintiff. Such legislation obviously reduces the incentive to bring suit for punitive damages (but does not impose a ceiling on defendants' payments). For a list of state statutes and proposed legislation on punitive damages caps and payment of punitive damages to state agencies, see the appendix to Justice Ginsburg's dissent in *BMW of North America v. Gore*, 116 S. Ct. at 1618-20; *see also* RICHARD L. BLATT ET AL., PUNITIVE DAMAGES: A STATE-BY-STATE GUIDE TO LAW AND PRACTICE (1991). Also, one state has passed legislation requiring post-trial review of punitive damages awards. *See* MONT. CODE ANN. § 27-1-221 7(c) (1996).

Additionally, as part of its tort reform efforts, Congress has considered legislation curbing punitive damages. Several bills have been proposed limiting punitive damages recovery to cases in which the plaintiff can demonstrate by "clear and convincing" evidence that the defendant displayed a conscious indifference to safety, or "actual malice." *See* H.R. 956, 104th Cong. § 2(a) (1995); H.R. 955, 104th Cong. § 8(A), (C) (1995); H.R. 917, 104th Cong. § 6(c)(I) (1995). Most of the bills have proposed some sort of cap on punitive damages in civil cases, usually $250,000 or three times the plaintiff's economic injury, whichever is greater. *See* H.R. 956, 104th Cong. § 201(b) (1995); H.R. 955 § 8(B); H.R. 917, 104th Cong. § 6(c)(2) (1995); H.R. 10, 104th Cong. § (C), (D) (1995). The proposed tort reform legislation is reviewed in Note, *"Common Sense" Legislation: The Birth of Neoclassical Tort Reform*, 109 HARV. L. REV. 1765, 1769 (1996); *see also* Neil A. Lewis, *Senate, 61-37, Approves Narrow Punitive-Damages Curb*, N.Y. TIMES, May 10, 1995, at B10 (reporting that the House of Representatives voted to limit punitive damages to the greater of $250,000 or three times the economic damages; and that the Senate voted for a narrower proposal which would limit punitive damages only in products liability cases to $250,000 or twice the

damages,[3] and the press has expressed strongly divergent opinions about the merits of punitive

---

amount of economic damages).

[3] With respect to the theory of punitive damages, commentators have disagreed, for instance, about the relevance of the wealth of the defendant. Compare, for example, Michael Rustad & Thomas Koenig, *The Historical Continuity of Punitive Damages Awards: Reforming the Tort Reformers*, 42 AM. U. L. REV. 1269, 1317 (1993) (endorsing the scaling of punitive damages to defendants' wealth because larger sanctions are required to influence the rich than the poor), with Kenneth S. Abraham & John C. Jeffries, Jr., *Punitive Damages and the Rule of Law: The Role of Defendant's Wealth*, 18 J. LEGAL STUD. 415 (1989) (concluding that a "defendant's wealth is irrelevant to the goal of deterring socially undesirable conduct and is an improper consideration in assessing the basis for retribution"), and Clarence Morris, *Punitive Damages in Tort Cases*, 44 HARV. L. REV. 1173, 1191 (1931) (noting that evidence of a defendant's wealth "instead of aiding the jury to assess a proper verdict, may prejudice them against the defendant and prevent an impartial judgment"). Additionally, commentators emphasize different goals in their consideration of punitive damages. Compare, for example, Thomas C. Galligan, Jr., *Augmented Awards: The Efficient Evolution of Punitive Damages*, 51 LA. L. REV. 3 (1990) (proposing a system of extra-compensatory damages based solely on deterrence), and Dan B. Dobbs, *Ending Punishment in "Punitive" Damages: Deterrence-Measured Remedies*, 40 ALA. L. REV. 831 (1989) (same), with Marc Galanter & David Luban, *Poetic Justice: Punitive Damages and Legal Pluralism*, 42 AM. U. L. REV. 1393 (1993) (discussing both punishment and deterrence goals), David G. Owen, *The Moral Foundations of Punitive Damages*, 40 ALA. L. REV. 705, 713 (1989) (same), and Jacqueline Perczek, Note, *On Efficiency, Punishment, Deterrence, and Fairness: A Survey of Punitive Damages Law and a Proposed Jury Instruction*, 27 SUFFOLK U. L. REV. 825, 856 (1993) (same).

The empirical importance of punitive damages also is the subject of dispute among academic commentators. A number of commentators suggest that punitive damages are widespread and problematic. See, for example, Dorsey D. Ellis, Jr., *Punitive Damages, Due Process, and the Jury*, 40 ALA. L. REV. 975, 975-77, 987-88 (1989) (arguing that "courts . . . have continued to uphold ever larger awards in cases in which defendants' conduct falls far short of the intentionally injurious behavior that traditionally characterized punitive damages cases"), Peter Huber, *No-Fault Punishment*, 40 ALA. L. REV. 1037, 1037-47 (1989) ("As the new tort revolution has taken hold, courts and juries have developed an even sharper, and for plaintiffs more lucrative, sense of outrage."), John Calvin Jeffries, Jr., *A Comment on the Constitutionality of Punitive Damages*, 72 VA. L. REV. 139, 139 (1986) (arguing "punitive damages are out of control"), and Malcolm E. Wheeler, *A Proposal for Further Common Law Development of the Use of Punitive Damages in Modern Product Liability Litigation*, 40 ALA. L. REV. 919, 919 (1989) (noting the "dramatic growth in both the frequency and size of punitive damages awards"). See also MARK PETERSON, SYAM SARMA & MICHAEL SHANLEY, PUNITIVE DAMAGES: EMPIRICAL FINDINGS (Institute for Civil Justice, Rand Corp., No. R-3311-ICJ, 1987) (empirical study concluding that punitive damage awards for business/contract cases are increasing while punitive damage awards for personal injury cases are stable). Others suggest that the fraction of cases in which punitive damages are awarded is not significant. See Stephen Daniels & Joanne Martin, *Myth and Reality in Punitive Damages*, 75 MINN. L. REV. 1 (1990) (empirical study concluding that the magnitude of the punitive damages problem is overstated by reformers); William M. Landes & Richard A. Posner, *New Light on Punitive Damages*, REGULATION, Sept.-Oct. 1986, at 33 (empirical study suggesting that "the incidence of punitive-damage awards may be exaggerated"); Michael Rustad, *In Defense of Punitive Damages in Products Liability: Testing Tort Anecdotes with Empirical Data*, 78 IOWA L. REV. 1 (1992) (empirical study concluding that there are very few punitive damages awards in product liability cases). These and other studies are reviewed by Marc Galanter, *Real World Torts: An Antidote to Anecdote*, 55 MD. L. REV. 1093, 1126-40 (1996). The most thorough, recent empirical study of punitive damages also emphasizes this point. *See* Theodore Eisenberg, John Goerdt, Brian Ostrom, David Rottman & Martin T. Wells, *The Predictability of Punitive Damages*, 26 J. LEGAL STUD. 623 (1997). *But see* A. Mitchell Polinsky, *Are Punitive Damages Really Insignificant, Predictable, and Rational? A Comment on Eisenberg et al.*, 26 J. LEGAL STUD. 663 (1997) (in part explaining why the study by Eisenberg et al. may understate the significance of punitive damages by ignoring the effects of such damages on settlements).

damages.[4]

Our goal in this Article is to develop a coherent and relatively simple set of principles for determining when punitive damages should be awarded and, in circumstances in which they are appropriate, what their level should be. We separately consider two social objectives in our analysis, deterrence and punishment.[5] Our methodology is economic in the sense that we organize our inquiry around an examination of how rational parties will respond to the threat of punitive damages, and whether their response will promote, or fail to promote, social welfare.[6]

The analysis of the deterrence objective comprises the first and major part of our Article. Our conclusions in this part flow from the basic principle that, to achieve appropriate

---

[4] *Compare Curtailing Civil Justice*, N.Y. TIMES, Apr. 8, 1995, § 1, at 22 (noting that "[p]unitive damages are often the best deterrent to destructive corporate behavior" and arguing against artificially low limits on awards), Editorial, ST. PETERSBURG TIMES, May 23, 1996, at 14A ("Punitive damages clearly have a place in society. They are designed to punish and awaken."), and Editorial, *Long shadow of the Exxon Valdez*, N.Y. TIMES, Sept. 21, 1994, at A22 ("[T]he jury ... clearly understood that only a sizable civil penalty would accomplish the purpose for which punitive damages are designed: to penalize flagrant wrongdoing and deter others from similar gross negligence."), *with* Editorial, *Trial Lawyers' Triumph*, WASH. POST, Mar. 19, 1996, at A16 (stating that "[l]egislation is needed because punitive damages are wildly unpredictable, so arbitrary as to be unfair"), Editorial, *Sue? Just Say No*, WALL ST. J., Sept. 28, 1994, at A18 (noting that punitive damages suits "have widespread, silent costs in frivolous filings being `settled out' and legitimate business activities curtailed for fear of exposure to jury risk"), and *No paint, no gain: product liability*, ECONOMIST, May 25, 1996, at 67 (arguing that the current punitive damages system yields "bizarre" and excessive awards that keep useful products off the American market and chill research and development of new products).

[5] These are traditionally said to be the goals of punitive damages. *See City of Newport v. Fact Concerts, Inc.*, 453 U.S. 247, 266-67 (1981) ("Punitive damages . . . are . . . intended to . . . punish the tortfeasor whose wrongful action was intentional or malicious, and to deter him and others from similar extreme conduct."); *Gertz v. Robert Welch, Inc.*, 418 U.S. 323, 350 (1974) ("[Punitive damages] are . . . private fines levied by civil juries to punish reprehensible conduct and to deter its future occurrence.").

[6] The concept of rationality in individual decisionmaking is discussed, for example, in Amartya Sen, *Rational Behaviour, in* 4 THE NEW PALGRAVE: A DICTIONARY OF ECONOMICS 68-76 (John Eatwell, Murray Milgate & Peter Newman eds., 1987). The notion of social welfare is reviewed, for instance, in DAVID KREPS, A COURSE IN MICROECONOMIC THEORY 149-82 (1990), and Amartya Sen, *Social Choice in* 4 THE NEW PALGRAVE, *supra*, at 382-93. Social welfare is determined by the well-being of individuals. Thus, social welfare generally rises if individual well-being rises and falls if individual well-being falls. In particular, social welfare reflects the deterrence objective of punitive damages, for the avoidance of harm raises the well-being of persons; and social welfare reflects the punishment objective of punitive damages, for the punishment of wrongdoers may be desired by individuals. We should add that, from the viewpoint of economics, there is no objective basis for saying that a specific formulation of social welfare (such as utilitarianism, which defines social welfare as the sum of individuals' utilities) is correct. Any measure of social welfare can be studied to determine what social policies or legal rules are best with respect to that measure. However, certain relatively simple measures of social welfare are often investigated for analytical convenience.

deterrence, injurers should be made to pay for the harm their conduct generates, not less, not more. If injurers pay less than for harm caused, underdeterrence will result -- that is, precautions will be inadequate, product prices will be too low, and risk-producing activities will be excessive. Conversely, if injurers are made to pay more than for harm caused, wasteful precautions may be taken, product prices may be inappropriately high, and risky, but socially beneficial, activities may be undesirably curtailed.

It follows from these observations that a crucial question for consideration is whether injurers sometimes escape liability for harms for which they are responsible. If injurers are found liable with certainty, then making them pay compensatory damages alone will result in their paying for the harm that their conduct generates. The imposition of extra-compensatory damages in this case would result in overdeterrence. But if there is a possibility that injurers will escape liability for harms for which they are responsible,[7] then the level of liability imposed on them when they *are* found liable needs to exceed compensatory damages so that, on average, they will pay for the harm that they cause. This excess liability can be labeled "punitive damages," and failure to impose it would result in inadequate deterrence. In summary, *punitive damages should ordinarily be awarded if, but only if, an injurer has a chance of escaping liability for the harm he caused.*[8]

This principle often will have transparent implications for the circumstances under which punitive damages should be awarded in practice. Consider a company that is responsible for trucking toxic waste to a dump site where it will be charged disposal fees. To reduce its fees, suppose the company allows some of the waste to leak onto the highway, knowing that the leak is unlikely to be noticed and traced to its source. Punitive damages

---

[7] We discuss in Section II.B several reasons why injurers might be able to escape liability: difficulty of detecting harm; inability to identify the injurer; problems in proving that the injurer is liable even if he can be identified; and the plaintiff's failure to sue due to litigation costs. *See* text accompanying notes 44-45 *infra*.

[8] We say "ordinarily" because we discuss circumstances in which it might not be desirable or necessary to impose punitive damages even if there is a chance of escaping liability (punitive damages might not be desirable when the probability of escaping liability is low, *see* Section II.B *infra*, and they might not be necessary when harm occurs to purchasers of products, *see* Section III.J *infra*). We also discuss reasons why punitive damages might be desirable even if there is *no* chance of escaping liability (specifically, when the injurer's act is malicious, *see* text accompanying note 11 *infra* and Section III.A).

obviously would be called for under our principles because there is a significant chance that the company will escape liability for the harm it caused. Alternatively, suppose the gross negligence of the firm that is responsible for treating the waste at the dump site leads to a substantial and highly visible spill from the firm's waste storage tanks. Punitive damages would not be called for because the firm is unlikely to be able to escape detection and liability for this harm.

When an injurer has a chance of escaping liability, the proper level of *total damages* to impose on him, if he is found liable, is the harm caused multiplied by the reciprocal of the probability of being found liable. Thus, for example, if the harm is $100,000 and there is a 25 percent chance that the injurer will be found liable for the harm for which he is legally responsible, the harm should be multiplied by 1/.25, or 4, so total damages should be $400,000. Since this amount will be paid by the injurer every fourth time he causes harm, his average payment will be $100,000 (= $400,000/4).[9] Thus, on average, the injurer will be paying for the harm he causes, and appropriate deterrence will result. Once the proper level of total damages is calculated in this way, punitive damages can be determined by subtracting compensatory damages from the total. In the example, since compensatory damages would equal the harm of $100,000, punitive damages would equal $400,000 less $100,000, or $300,000.

If punitive damages are needed according to this theory, we believe that courts and juries often will be able to obtain enough information about the likelihood of escaping liability to apply the theory reasonably well. We will discuss how our analysis relates to several leading punitive damages cases,[10] and we will provide model jury instructions that could be used to aid jurors in applying the principles that we develop.

We also will relate our analysis of the deterrence rationale for punitive damages to the criteria commonly applied by the courts in imposing such damages. Importantly, we will

---

[9] We are presuming here that the injurer engages in repetitive conduct. Our point, however, applies even if the injurer commits the harmful act only once. It is still true then that the probability-discounted or "expected" value of what he pays is $100,000. *See* note 46 *infra*.

[10] *See* Section II.D *infra*.

explain why the reprehensibility of a corporate defendant's conduct generally should not be a factor in deciding on punitive damages for purposes of promoting deterrence (although the reprehensibility of the conduct of a person who is a defendant may be relevant to punitive damages and deterrence). In addition, we will argue that the wealth of a corporate defendant presumptively should not be taken into account in determining the level of punitive damages (although again the conclusion may be different in the case of a person who is a defendant). We also will consider other aspects of punitive damages policy from a deterrence perspective, including the appropriateness of caps on punitive damages, the relevance of potential harm for punitive damages, the insurability of punitive damages, and the importance for punitive damages of the distinction between victims who are customers of an injurer and victims who are strangers to the injurer.

One further observation about our analysis of deterrence is worth noting. We ordinarily assume that the benefits that injurers obtain from engaging in the conduct that gives rise to harm counts in social welfare. We also will discuss, however, the possibility that such benefits should not be included -- notably, when a wrongdoer maliciously derives pleasure from his victim's suffering. We will explain that if the injurer's benefits are excluded from social welfare, then punitive damages may be needed for proper deterrence *even when there is no chance of escaping liability*.[11]

In our discussion of the second objective of punitive damages -- punishment -- we focus on the assumption that the underlying goal of society is to penalize especially blameworthy *individuals*. Achieving this goal is reasonably straightforward if the defendant is a person who has been found to have acted culpably -- imposing punitive damages on that person punishes *him*. But if the defendant is a corporation, imposing punitive damages on it may or may not lead to the punishment of blameworthy individuals within the corporation, for a variety of reasons that we will discuss. To the extent that such individuals are not punished, imposing punitive damages on the corporation does not advance the punishment goal. Moreover, we will explain why much of the sting from imposing punitive damages on corporations may be

---

[11] *See* Section III.A *infra*.

borne by individuals who are usually *not* thought to be culpable, namely shareholders and customers. In light of these points, we conclude that the extent to which imposing punitive damages promotes the punishment goal may be significantly different when defendants are corporations from when they are individuals, and that the importance of punitive damages as a form of punishment may be considerably attenuated for corporate defendants.

The plan of our Article is as follows. In Section II, we review the economic theory of deterrence and develop the basic principles determining when punitive damages should be awarded, and at what level. In this Section we also apply these principles to certain aspects of punitive damages law and legislation, as well as to several prominent punitive damages cases. Section III relates the basic principles to a number of criteria that are employed by the courts to determine the appropriateness and/or magnitude of punitive damages awards, and also examines a variety of other factors and policies that bear on punitive damages. Section IV discusses the punishment goal of punitive damages, and Section V briefly summarizes our main points.[12] An Appendix contains the model jury instructions.

---

[12] Our article builds on many other contributions. The general point that, to achieve proper deterrence, sanctions must be inflated if injurers can escape liability dates back at least to Bentham, JEREMY BENTHAM, PRINCIPLES OF PENAL LAW, *reprinted in* 1 THE WORKS OF JEREMY BENTHAM, 365, 401-02 (John Bowring ed., 1962) (1838-1843), and has been applied to the subject of punitive damages by many commentators. The first explicit references to the factor of escaping liability as a justification for punitive damages apparently are RICHARD A. POSNER, ECONOMIC ANALYSIS OF LAW 77-78 (1st ed. 1972) and Dorsey D. Ellis, Jr., *Fairness and Efficiency in the Law of Punitive Damages*, 56 S. CAL. L. REV. 1, 25-26 (1982). This justification for punitive damages has been developed most thoroughly by Robert D. Cooter, *Punitive Damages for Deterrence: When and How Much?*, 40 ALA. L. REV. 1143 (1989) [hereinafter Cooter, *Deterrence*].

In addition to the items cited in the previous paragraph, there are many others that consider punitive damages using economic analysis. Those in print include Darryl Biggar, *A Model of Punitive Damages in Tort*, 15 INT'L REV. LAW & ECON. 1 (1995); James Boyd & Daniel E. Ingberman, *Noncompensatory Damages and Potential Insolvency*, 23 J. LEGAL STUD. 895 (1994); Bruce Chapman & Michael Trebilcock, *Punitive Damages: Divergence in Search of a Rationale*, 40 ALA. L. REV. 741 (1989); Robert D. Cooter, *Economic Analysis of Punitive Damages*, 56 S. CAL. L. REV. 79 (1982) [hereinafter Cooter, *Economic Analysis*], Richard Craswell, *Damage Multipliers in Market Relationships*, 25 J. LEGAL STUD. 463 (1996); David Friedman, *An Economic Explanation of Punitive Damages*, 40 ALA. L. REV. 1125 (1989); Galligan, *supra* note 3; David D. Haddock, Fred S. McChesney & Menahem Spiegel, *An Ordinary Economic Rationale for Extraordinary Legal Sanctions*, 78 CAL. L. REV. 1 (1990); Jason Johnston, *Punitive Liability: A New Paradigm of Efficiency in Tort Law*, 87 COLUM. L. REV. 1385 (1987); Marcel Kahan & Bruce Tuckman, *Special Levies on Punitive Damages: Decoupling, Agency Problems, and Litigation Expenditures*, 15 INT'L REV. LAW & ECON. 175 (1995); William Landes & Richard Posner, *An Economic Theory of Intentional Torts*, 1 INT'L REV. LAW & ECON. 127 (1981); George L. Priest, *Insurability and Punitive Damages*, 40 ALA. L. REV. 1009 (1989) [hereinafter Priest, *Insurability*]; George L. Priest, *Punitive Damages and Enterprise Liability*, 56 S. CAL. L. REV. 123 (1982); Paul H. Rubin, John E. Calfee & Mark F. Grady, BMW v. Gore*: Mitigating the Punitive Economics of Punitive Damages*, 5 SUP. CT. ECON. REV. (forthcoming June 1997). Recent unpublished articles

include James Boyd & Daniel E. Ingberman, Do Punitive Damages Promote Deterrence? (Bus. Law and Econ. Center, Wash. Univ., St. Louis, BLE-95-08, August 1995); Andrew F. Daughety & Jennifer F. Reinganum, Everybody Out of the Pool: Products Liability, Punitive Damages and Competition (Nov. 1995) (unpublished manuscript, on file with authors) [hereinafter Daughety & Reinganum, Products Liability]; Andrew F. Daughety & Jennifer F. Reinganum, Settlement, Deterrence and the Economics of Punitive Damages Reform (Department of Econ. & Bus. Admin., Vanderbilt Univ. Working Paper No. 97-W04, 1997) [hereinafter Daughety & Reinganum, Settlement]; Peter Diamond, Efficiency Effects of Punitive Damages (May 17, 1997) (unpublished manuscript on file with authors) [hereinafter Diamond, Efficiency Effects]; Peter Diamond, Integrating Punishment and Efficiency Concerns in Punitive Damages for Reckless Disregard of Risks to Others (May 20, 1997) (unpublished manuscript on file with authors) [hereinafter Diamond, Punishment and Efficiency].

Our treatment of deterrence and punitive damages is more comprehensive than that of these earlier economically-oriented articles, and our analysis of certain issues that have been considered previously differs from what has been written. Moreover, our examination of punishment and punitive damages is substantially different from what is found in the literature on punitive damages. We discuss the relationship between our Article and some of the preceding literature in notes 14, 47, 85, 105, 116, 143, 160, 171, 181, 183, 196, 218, and 241 *infra*. *See also* note 191 *infra*.

## II. DETERRENCE: THE BASIC THEORY

In this Section, we summarize the basic principles of the economic theory of deterrence and explain what these principles imply for the use of punitive damages. By deterrence, we mean what is often called *general deterrence*, namely, the effect that the prospect of having to pay damages will have on the behavior of similarly situated parties in the future (not just on the behavior of the defendant at hand).[13]

We should add that the basic theory that we are about to review is the standard theory of deterrence, widely agreed upon by economically-oriented scholars.[14] As noted, we will usually make the conventional assumption that the benefits that injurers obtain from engaging in the conduct that gives rise to harm count in social welfare.[15] Thus, for example, we will

---

[13] General deterrence may be contrasted with *specific deterrence*, which is the effect that the imposition of a sanction on a party will have on *that* party's future behavior. *See generally* JEREMY BENTHAM, PRINCIPLES OF PENAL LAW, *reprinted in* 1 THE WORKS OF JEREMY BENTHAM, *supra* note 12, at 365, 396; HERBERT L. PACKER, THE LIMITS OF THE CRIMINAL SANCTION 39-48 (1968).

[14] The theory of deterrence -- the elaboration of the effect on rational actors of the possible imposition of sanctions for violations of law -- was first articulated in detail by Jeremy Bentham, *see generally* BENTHAM, PRINCIPLES OF PENAL LAW, *reprinted in* 1 THE WORKS OF JEREMY BENTHAM, *supra* note 12, at 365, and has been developed intensively in the last several decades, stimulated largely by an important article by Gary S. Becker, *Crime and Punishment: An Economic Approach*, 76 J. POL. ECON. 169 (1968). This literature is synthesized and surveyed in, for example, R. A. CARR-HILL & N. H. STERN, CRIME, THE POLICE, AND CRIMINAL STATISTICS (1979); WILLIAM A. LUKSETICH & MICHAEL D. WHITE, CRIME AND PUBLIC POLICY (1982); DAVID J. PYLE, THE ECONOMICS OF CRIME AND LAW ENFORCEMENT (1983). More recent contributions, up to 1992, are collected in BIBLIOGRAPHY OF LAW AND ECONOMICS 504-26 (Boudewijn Bouckaert & Gerrit De Geest eds., 1992).

Beginning with GUIDO CALABRESI, THE COSTS OF ACCIDENTS (1970), many writers have applied the general theory of deterrence to the subject of tort liability. The literature on deterrence and tort liability is presented in an integrated way in WILLIAM M. LANDES & RICHARD A. POSNER, THE ECONOMIC STRUCTURE OF TORT LAW (1987) and STEVEN SHAVELL, ECONOMIC ANALYSIS OF ACCIDENT LAW (1987).

[15] This assumption is consistent with a standard definition of social welfare employed in deterrence theory, in which social welfare equals the benefits that parties obtain from their activities less various costs (including the expense of precautions taken to avoid harm, the harm that does occur, and any costs associated with use of the legal system).

Note that this definition of social welfare does not incorporate the compensation of victims as a social benefit, even though most individuals consider compensation to be a social goal. Taking the compensation goal into account in social welfare would be relatively straightforward, but doing so is unnecessary for our purposes, for several reasons. First, and most importantly, punitive damages are generally extra-compensatory; thus, whether or not they are paid typically does not affect fulfillment of the compensatory objective. Second, victims often have insurance -- so called "first-party" insurance -- that compensates them for their losses, at least partially, so that the extent to which it is necessary to rely on the liability system to achieve the compensation goal may be limited. Third, the insurance system is generally a much less expensive way to achieve compensation than the liability system. *See, e.g.*, DON DEWEES, DAVID DUFF & MICHAEL TREBILCOCK, EXPLORING THE DOMAIN OF ACCIDENT LAW: TAKING THE FACTS SERIOUSLY 421-24 (1996) (summarizing empirical findings concerning the relative cost of the insurance

assume that the time saved by a speeding driver, or the cost saved by a company that chooses not to purchase certain pollution control equipment, constitutes a social benefit that is to be weighed against the harm from speeding or polluting. We will consider the implications for punitive damages of the alternative assumption -- that the benefits from harmful conduct do *not* count in social welfare -- when we examine the reprehensibility criterion in Section III.A.

We first discuss deterrence in a very simple setting in which, whenever a party causes harm, he will be sanctioned for sure for having caused that harm. We then discuss the situation in which parties sometimes escape sanctions for harms for which they are responsible. It is in this latter case, as we indicated above, that there is a reason to impose damages that exceed the harm, and thus for punitive damages to be used.

### A. Optimal Damages When the Defendant is Found Liable for Sure

The central point that we want to explain here is that *if a defendant will be found liable for sure for the harm for which he is responsible, the proper magnitude of damages is the harm the defendant has caused.*[16] If damages are either lower or higher than the harm, various socially undesirable consequences will result (described below). We first illustrate these points when liability is strict -- an injurer is supposed to be liable for any harm caused[17] -- and then when liability is fault-based (in which case our conclusions are somewhat qualified). We assume for purposes of this discussion that harm is properly measured.[18]

There are a number of reasons why it is best for damages to equal harm under strict liability. One concerns the *level of precautions* taken by parties when engaging in their activities. We interpret the term "precautions" very generally. For example, it can refer to

---

system and the liability system in providing compensation to accident victims).

[16] Readers familiar with the economic logic supporting this claim may want to proceed directly to Section II.B.

[17] For ease of analysis, we will presume that issues of contributory negligence do not arise.

[18] In Section III.L, *infra*, we discuss the possibility that compensatory damages do not correctly reflect the harm that actually occurred. Also, in note 145 *infra* in Section III.E, we observe that the total social harm caused by an adverse event includes litigation costs, but for simplicity we will not take this refinement into account in Section II.

safety devices -- such as valves to release excess pressure on tanks used to store dangerous chemicals -- or to the actions of individuals that reduce harm -- such as inspecting the brakes of trucks. Additionally, and importantly, precautions include the variety of ways in which firms monitor and screen their employees -- such as an airline testing its pilots for their use of controlled substances. Any action that reduces the risk or the level of harm is what we mean by a precaution.

If damages equal harm, potential injurers will in theory have socially correct incentives to take precautions. Specifically, they will be induced to spend money on precautions if the expenditure is socially worthwhile in the sense that the expenditure reduces the harm by a greater amount. Suppose, for example, that by spending $50,000 on a precaution, a firm can prevent a harm of $100,000. It is socially desirable that such a precaution be taken. And if the level of liability is equal to the harm of $100,000, a firm will be led to spend $50,000 to prevent harm. But if the level of liability is less than $100,000, then a firm might not take precautions when it should. For instance, if the level of liability is only $30,000, then a firm would not take the precaution costing $50,000, even though this precaution is socially desirable.[19]

Conversely, if damages exceed harm, firms might be led to take socially excessive precautions. A socially excessive precaution is a precaution that costs more than the reduction of harm occasioned by it. In the previous example, suppose that the precaution costs $250,000 (instead of $50,000). Then the precaution would be socially excessive because it would be wasteful to spend $250,000 to avoid a harm of $100,000. Yet if damages exceed harm, a firm might be led to take the precaution. Assume, for instance, that punitive damages of $200,000 are added to the compensatory damages of $100,000, so that the firm's total damages will be $300,000 if it does not take the precaution. Since the cost of the precaution is $250,000, the firm will be led to take it, even though the precaution is socially wasteful.

Although the notion of excessive spending on precautions might seem counterintuitive to the reader, it is quite real and often recognized as such. For example, reference frequently

---

[19] In this example, and in those in the remainder of the Article, we consider relatively simple fact situations. The principles that these examples illustrate carry over to more complicated and realistic circumstances.

is made to "defensive medicine," by which is meant physicians' wasteful use of tests and diagnostic procedures in response to the threat of liability.[20] On reflection, it is not difficult to imagine that excessive expenditures could be made on safety precautions in almost any context. Consider, for instance, how much could be spent on cement traffic dividers for city streets (suppose that all streets had dividers); on additional personnel to monitor employees' safety practices at oil refineries (each employee at a refinery could be accompanied by another watching over his activities); or on sensors to detect railroad track switching problems (a costly sensor might be installed on every switch on every track).[21]

In our numerical examples above, we have been discussing precautions as if they would completely eliminate the risk of harm if they are taken. However, everything we have said carries over to situations in which precautions reduce, but do not eliminate, the risk of harm. In particular, the proper magnitude of damages continues to be the harm the defendant has caused.[22] This level of damages again induces potential injurers to take appropriate precautions, but the determination of appropriate precautions now involves a comparison of the cost of the precaution to the reduction in the *expected harm* that results if the precaution is taken, where expected harm refers to the harm multiplied by the probability of its

---

[20] Some studies show this to be an important phenomenon. *See, e.g.*, Daniel Kessler & Mark McClellan, *Do Doctors Practice Defensive Medicine?*, 110 Q.J. ECON. 353 (1996) (finding that "malpractice reforms that directly reduce provider liability pressure lead to reductions of 5 to 9 percent in medical expenditures without substantial effects on mortality or medical complications"). *See also* Stanley Joel Reiser, *Malpractice, Patient Safety, and the Ethical and Scientific Foundations of Medicine*, *in* THE LIABILITY MAZE: THE IMPACT OF LIABILITY LAW ON SAFETY AND INNOVATION 227, 231-33, 238-39 (Peter W. Huber & Robert E. Litan eds., 1991) (discussing interviews with doctors and medical administrators who claim that certain tests, procedures, and paperwork are responses to the threat of malpractice suits and do not improve the care that patients receive).

[21] For example, the Federal Railroad Administration has studied proposals for sensor systems to prevent derailments on railroad bridges, but it concluded that "they would cost as much as $40,000 per bridge to install. For the country's 100,000 railroad bridges, the agency reported, the cost could reach billions for installation and $60 million more a year for maintenance." Richard Perez-Pena, *Rail Accident Stirs Debate About Sensors*, N.Y. TIMES, Nov. 29, 1996, at B1, *available in* LEXIS, News Library, Nyt File. The Federal Railroad Administration determined that the cost was too high because of the low frequency of railroad accidents that occur on bridges. *Id.*

[22] This is true assuming that individuals are *risk neutral*. The meaning of this assumption, and the justification for making it, is discussed in the text at the end of Section II.A.

occurrence.[23]

Note that even when proper precautions are exercised, some accidents will occur, because of residual, hard-to-eliminate risks. For example, even the safest automobile tire may blow out and cause an accident. That a blowout occurs and results in an accident does not necessarily mean that inadequate precautions were taken in the design and manufacture of the tire. Likewise, even if employees are screened and monitored with appropriate vigilance, there will be occasional employee misbehavior, possibly egregious in character.[24] It is important to stress that such misbehavior does not necessarily signal a lack of proper oversight by the employer.[25]

Let us now turn to a second reason why it is best for damages to equal harm. This reason concerns the extent to which individuals and firms participate in risky activities -- what we will refer to, in brief, as their *level of activity*.[26] A party's level of activity affects the magnitude of expected harm, whatever precautions are taken by the party when engaging in the activity. For example, the more miles a person drives -- his level of activity -- the greater the number of accidents that he is likely to cause, whatever is his level of care when he drives.

---

[23] For example, if a harm of $10,000 occurs with a probability of 20%, the expected harm is $2,000 (= 20% x $10,000). This expected harm also can be interpreted as the average harm per instance of some conduct that, each time it occurs, has a 20% chance of causing a $10,000 harm. More generally, the expected harm is the sum of products of each possible magnitude of harm and its probability. Thus, if there is a 20% chance of a $10,000 harm, as well as a 5% chance of a $30,000 harm, the expected harm is $3,500 (= (20% x $10,000) + (5% x $30,000)).

[24] An example may clarify this point. Suppose that 3% of the applicants for employment at a firm are "rotten apples" who will misbehave on the job and cause a $100,000 accident; that by spending $500 per applicant on investigating the applicant's background it is possible to detect one-third of the rotten apples; that by instead spending $4,000 per applicant on more intensive screening, it is possible to detect two-thirds of the rotten apples; but that it is impossible to detect the remaining one-third of the rotten apples. In this example, it is socially desirable for the firm to spend $500 per applicant on screening: this results in a 1% chance of detecting a rotten apple (one-third of the 3% of rotten apples in the applicant pool are identified) and thereby avoiding a $100,000 harm, which is to say, it reduces the expected harm by $1,000 (= 1% x $100,000). It is not socially worthwhile, however, to spend $4,000 per applicant, because that would reduce the expected harm by only $2,000 (= 2% x $100,000). Significantly, whichever amount is spent, there will still be at least a 1% chance of employee misconduct.

[25] In the example in the previous footnote, the occurrence of employee misconduct does not imply that the employer's screening activity was improper -- misconduct can occur even when proper screening is undertaken.

[26] The distinction between level of activity and level of care was introduced in Steven Shavell, *Strict Liability versus Negligence*, 9 J. LEGAL STUD. 1 (1980).

(Of course, the more care he takes when driving, the lower the expected number of accidents per mile driven.) Similarly, the more units of the product a firm produces and sells -- its level of activity -- the greater the number of accidents that will be caused by its product, whatever are the safety features of the product (which affect the expected harm per unit sold).

If damages equal harm, potential injurers will have the socially correct incentives to engage in risky activities. In particular, they will be induced to engage in an activity if and only if the benefit they derive exceeds the additional harm caused by their decision to engage in it. If damages equal harm, an individual will tend to participate in an activity like hunting if and only if the benefit he obtains from this activity exceeds the expected accident costs that hunting imposes on others. Likewise, a firm will produce a product if and only if its value, as reflected in the willingness of customers to pay for it, exceeds the full costs of its production, including accident losses. Specifically, if damages equal harm, the cost of production will include the harm. To cover its costs, a firm will have to sell its product at a higher price -- a price that reflects the average harm caused per unit of output. Consumers will only buy the product, therefore, if they value it more highly than its full cost of production, including the harm. This will lead to the socially correct level of consumption of the product. In other words, the fact that the product price will rise in response to the firm's liability costs is desirable because, if damages equal harm, this price increase appropriately discourages consumers of the product from, in effect, causing an excessive number of accidents by consuming too much of the good.

It is obvious from what has been said that if damages are less than harm, then parties will engage in activities to an excessive extent -- that is, engage in activities even when their benefits are outweighed by the harms caused. Conversely, if damages exceed harm, then parties may be led to curtail their activities to an inappropriate extent -- not engage in them even when their benefits exceed the harms caused. In particular, a firm might be induced to withdraw its product from the marketplace even though consumers value the product more than its full cost of production, which includes the average harm caused by the product.

The preceding possibilities -- excessive or inadequate engagement in activities that cause harm -- are realistic. For example, it appears that much of the harm from automobile

pollution is not reflected in the price of gasoline or of automobiles.[27] Hence, individuals will tend to drive too much.[28] Conversely, there is evidence to suggest that manufacturers of certain socially desirable products (for example, childhood vaccines) may have stopped selling their products due to the prospect of damages exceeding harm.[29]

We have now explained the fundamental reasons why damages should equal harm under strict liability. We next discuss why we assume that damages should equal harm under the negligence rule.

By definition of the negligence rule, if a potential injurer fails to take proper precautions -- does not meet the negligence standard -- he is said to be negligent and must pay damages.[30] The economic interpretation of the proper negligence standard involves a comparison of the cost of taking the precaution to the reduction in harm that results from taking it; if the former amount is less than the latter amount, the precaution should be taken

---

[27] *See* Kenneth A. Small & Camilla Kazimi, *On the Costs of Air Pollution from Motor Vehicles*, 29 J. TRANSP. ECON. & POL'Y J. 7, 27 (1995) (demonstrating that an additional tax of about 50% on the price of gasoline would be required to account for air pollution costs).

[28] *See* JAMES J. MACKENZIE, THE GOING RATE: WHAT IT REALLY COSTS TO DRIVE 5 (1992) ("The net effect of [federal and state] policies is to make driving seem cheaper than it really is and to encourage the excessive use of automobiles and trucks.").

[29] A number of articles have discussed the withdrawal of products from the marketplace in response to actual and prospective liability costs (including liability insurance premiums). *See* DEWEES, DUFF & TREBILCOCK, *supra* note 15, at 241-42 (1996) (discussing the reduction of vaccine manufacturing due to the expansion of liability); Louis Lasagna, *The Chilling Effect of Product Liability on New Drug Development*, *in* THE LIABILITY MAZE, *supra* note 20, at 334, 337-41 (describing the litigation which led to the voluntary withdrawal of the anti-morning sickness drug, Bendectin); Robert Martin, *General Aviation Manufacturing: An Industry under Siege*, *in* THE LIABILITY MAZE, *supra* note 20, at 478 (concluding that product liability litigation "threatens the very existence" of the corporate and private airplane industry); W. KIP VISCUSI, REFORMING PRODUCTS LIABILITY 8 (1991) (noting product liability litigation has forced some companies to stop producing private airplanes). It should be observed that these discussions do not address the question whether the reduction in productive activity was due to the imposition of damages exceeding harm. However, if as some commentators believe, product liability is socially excessive, *see, e.g.*, PETER W. HUBER, LIABILITY: THE LEGAL REVOLUTION AND ITS CONSEQUENCES 9-11 (1988) (describing the excessive costs imposed by product liability suits), then such reductions in the availability of products may be partially attributed to excessive damages.

[30] *See* RESTATEMENT (SECOND) OF TORTS § 282 ("[N]egligence is conduct which falls below the standard established by law for the protection of others against unreasonable risk of harm."); PROSSER AND KEETON ON THE LAW OF TORTS 169 (5th ed. 1984) ("[N]egligence is not necessarily the absence of solicitude for those who may be adversely affected by one's actions but is instead behavior which should be recognized as involving unreasonable danger to others.").

and the failure to do so is negligent.[31] For example, it would be negligent not to take a precaution costing $50,000 that would prevent a harm of $100,000.

Under the negligence rule, it is readily seen that if damages equal harm, potential injurers will be led to comply with the negligence standard (assuming that it is chosen properly) and thus to take appropriate precautions. If a precaution costing $50,000 would prevent a harm of $100,000, then the threat of having to pay damages of $100,000 for not taking the precaution would clearly induce a party to spend $50,000 on the precaution. However, if damages are less than harm, then the negligence standard might not be met and underdeterrence would result. In the example, if damages are only $40,000 (even though harm is $100,000), then the party would not be led to take the precaution costing $50,000.[32] Conversely, if damages exceed harm, a potential injurer will have a stronger motive to meet the negligence standard than if damages are equal to harm. If damages are $200,000 (even though harm is $100,000), a party will have a greater incentive to spend $50,000 on the precaution than if damages are $100,000. But he will not take more precautions than are required to meet the negligence standard, assuming that the negligence determination is free of errors. For in the absence of errors, there is no reason to do more than just satisfy the negligence standard, even if the damages that would be imposed if negligence is found far exceed the harm.

Realistically, however, there will be errors in the negligence determination, implying that damages exceeding harm could lead to excessive precautions. There are reasons why parties may be found liable under the negligence rule, despite their effort to act non-

---

[31] The economic interpretation of the negligence rule is encapsulated in Judge Learned Hand's algebraic formula for determining the due care standard. In his opinion in *United States v. Carroll Towing Co.*, 159 F.2d 169 (2d Cir. 1947), Hand said that a party is negligent if he fails to take a precaution when the burden of the precaution is less than the reduction of the expected loss occasioned by taking the precaution -- in other words if $B < PL$, where $B$ is the burden of taking the precaution, $P$ is the probability of the loss if the precaution is not taken, and $L$ is the magnitude of the loss. Id. at 173. For further development of this idea, see, for example, John Prather Brown, *Toward an Economic Theory of Liability*, 2 J. LEGAL STUD. 323 (1973); *see also* William M. Landes and Richard A. Posner, *The Positive Economic Theory of Tort Law*, 15 GA. L. REV. 851, 892-903 (1981) (arguing that judicial practice is consistent with the economic interpretation of the negligence rule).

[32] Note that damages must be sufficiently below harm before the party would find it worthwhile to act negligently. For if damages exceeded the $50,000 cost of the precaution, the party would be induced to take the precaution even if damages were less than the harm of $100,000.

negligently. Notably, they may inaccurately assess what the negligence standard is, or the courts may erroneously observe their behavior and find them negligent when they were not negligent. Because of the risk of mistakes, parties may well have an incentive to take greater precautions than they would otherwise, in order to reduce the chance that they will incorrectly be found negligent.[33] If, as a result, they are induced to take socially excessive precautions, raising the level of damages imposed on them will exacerbate this problem.

Next consider the relationship between damages and the level of activity under the negligence rule. In this regard, observe that, in the absence of mistakes, the negligence rule may result in parties participating in risky activities to a socially excessive extent.[34] This is because, once a party takes the precautions required by the negligence standard, he will not be found liable for any harms that he causes. For example, a person who drives with reasonable care will not be found negligent, and therefore will not have to pay for any harm caused by his driving; consequently, he will drive more than is socially desirable. Or a manufacturer that takes appropriate care in the design of its product will not be liable under a negligence rule for harms that result if its product nonetheless turns out to be flawed; as a result, too much of the product will tend to be produced.[35]

---

[33] Although parties will generally reduce the chance of being found negligent by mistake by taking greater care, it does not necessarily follow that this will lead them to take more care than they would if there were no mistakes in the determination of negligence. The reason is that, to the extent that there is a random component in the assessment of care, the exercise of greater care will only be partially rewarded. For instance, if half the time a party's care is not observed and courts make a guess about its level, then increasing the level of care would only benefit the party half the time. The condition under which parties will take greater care is, roughly, that the assessment of care is not too imprecise. For details, see John E. Calfee & Richard Craswell, *Some Effects of Uncertainty on Compliance with Legal Standards*, 70 VA. L. REV. 965 (1984), and Richard Craswell and John Calfee, *Deterrence and Uncertain Legal Standards*, 2 J.L. ECON. & ORG. 279 (1986); see also SHAVELL, *supra* note 14, at 93-99. *But see* Mark F. Grady, *A New Positive Economic Theory of Negligence*, 92 YALE L.J. 799, 817-21 (1983) (showing that under a different but plausible interpretation of the negligence rule, mistakes will not cause parties to take excessive care); Marcel Kahan, *Causation and Incentives to Take Care under the Negligence Rule*, 18 J. LEGAL STUD. 427, 437-39 (1989) (same).

[34] The result that the negligence rule leads to socially excessive participation in risky activities was originally developed in Shavell, *supra* note 26; *see also* A. Mitchell Polinsky, *Strict Liability vs. Negligence in a Market Setting*, 70 AM. ECON. REV. 363 (1980) (demonstrating the result in the context of firms).

[35] The point of this paragraph may be clarified by a numerical illustration. Suppose that taking proper precautions would cost an individual $100 each time he engages in an activity and would reduce the risk of an accident that would cause harm of $100,000 to 1/2%. Then, assuming that the individual would take the precautions in order to avoid liability for negligence, he will engage in the activity whenever the benefit to him exceeds the $100

However, because non-negligent parties sometimes will be found liable by mistake, they *will* sometimes bear damages. In principle, this could improve matters with respect to the problem that the negligence rule may induce parties to participate in risky activities to an excessive degree. It also is possible, however, that finding parties negligent by mistake will result in their bearing damages in excess of the harm they have caused,[36] and thereby discourage their participation in the activity to an inappropriate extent. This effect, if it occurs, will be exacerbated by raising the level of damages.

The discussion of the negligence rule shows that the optimal level of damages is not as easily determined as under the strict liability rule. Under strict liability, we concluded that damages should equal harm. Under the negligence rule, we have observed that in the absence of mistakes, damages equal to harm will appropriately encourage parties to take precautions, but so will higher levels of damages.[37] In the presence of mistakes, the optimal level of damages under the negligence rule is difficult to ascertain, although it is clear from what we have said that if damages are set too high, parties will tend to be induced to take excessive precautions; moreover, they will not participate in their activities to an appropriate degree. In the light of the preceding points, and recognizing that there is not a simple, theoretically correct, answer to the question of what level of damages is optimal under the negligence rule, we will assume for purposes of our analysis that optimal damages under the negligence rule are

---

cost of precautions. However, each time he engages in the activity, he causes total social costs of $600: the $100 cost of precautions *plus* the expected harm of $500 (= 1/2% x $100,000). Consequently, from society's perspective, he should engage in the activity only if his benefit exceeds $600. If his benefit lies between $100 and $600, however, he will engage in the activity even though doing so is socially undesirable; in other words, his participation in the activity will be socially excessive.

[36] More precisely, expected damages could exceed harm for two reasons. First, a party who actually caused harm but who was not negligent might be found negligent by mistake and made to pay such a high level of damages that, on average, he will pay for more than the harm he caused. Second, a party who in fact did not cause harm might mistakenly be found both to be the cause of harm and to be negligent. Obviously, any damages imposed on such a party are excessive and will chill participation in activities in which such mistakes can occur.

[37] Some lower levels of damages will as well, provided that these levels still exceed the cost of precautions. *See* note 32 *supra*.

equal to the harm.[38]

This concludes our review of the implications of deterrence theory for the optimal level of damages under the rules of strict liability and negligence when injurers will be found liable for sure. Because damages should equal harm under the strict liability rule, and because we assume damages should equal harm under the negligence rule for the reasons given, we generally will not distinguish between the rules in our subsequent discussion.

In passing, we want to note that the principal conclusion reached in this Section -- that damages should equal harm -- depends on how potential injurers respond to risk. We have implicitly assumed that they are *risk neutral*. This means that in considering situations of risk, parties care only about the expected value of a risky situation -- that is, the magnitude of a potential loss or gain multiplied by the probability of the loss or gain occurring.[39] If injurers are *risk averse* (they dislike uncertainty itself)[40] and cannot purchase liability insurance, then the optimal level of damages tends to be lower than harm, both to reduce the imposition of risk on injurers and because damages do not need to be as high to induce injurers to behave appropriately.[41]

Notwithstanding this last point, we will usually assume below that, when injurers are found liable for sure, the level of damages that is optimal with respect to deterrence is equal to

---

[38] This assumption is made mainly for convenience and does not affect our main point that the benchmark level of damages -- the level that would be appropriate for deterrence if injurers were found liable for sure -- should be inflated using a specific multiplier formula if injurers can sometimes escape liability. If the benchmark level of damages is different from harm, this benchmark quantum, whatever its magnitude, should be inflated by the multiplier we give below in order for deterrence to be appropriate when injurers can escape liability. *See* Section II.B *infra*.

[39] A risk-neutral injurer would be indifferent between paying damages of, say, $10,000 for sure, and facing a risky situation in which he will have to pay either nothing or $20,000 with equal probability. This is because the risky situation involves an expected payment of $10,000 (= 50% x $20,000). For discussion of the concept of risk neutrality, see, for example, ROBERT S. PINDYCK & DANIEL L. RUBINFELD, MICROECONOMICS 146 (3d ed. 1995).

[40] A risk-averse injurer would prefer to pay damages of $10,000 for sure than to face a risky situation in which he will have to pay either nothing or $20,000 with equal probability. This is true even though the risky situation involves an expected payment of $10,000. *See, e.g., id.* at 146.

[41] This result is demonstrated in Steven Shavell, *On Liability and Insurance*, 13 BELL J. ECON. 120 (1982) and in SHAVELL, *supra* note 14, at 218-21 (1987). The particular level of damages that is optimal depends on the degree of risk aversion of injurers and on whether victims are insured or, if not, how risk averse they are.

harm,[42] for two reasons. First, even if parties are risk averse, if they can purchase liability insurance it can be shown that the optimal level of damages still equals the harm.[43] Second, it can be demonstrated that publicly held firms should be treated as approximately risk neutral -- implying that damages should equal harm -- if their shareholders have well-diversified portfolios, which often, if not usually, will be the case.[44]

### B. Optimal Damages When the Defendant Can Sometimes Escape Liability

The main point that we will develop here is that *if a defendant can sometimes escape liability for the harm for which he is responsible, the proper magnitude of damages is the harm the defendant has caused multiplied by a factor reflecting the probability of his escaping liability*. As we will explain shortly, use of such a multiplier will make defendants pay on average for harm actually done and thus will lead to socially desirable behavior in terms of precautions and participation in risky activities.

There are several reasons why injurers sometimes escape liability for harms for which they should be liable under a liability rule. First, it may be difficult for the victim to determine that the harm was the result of some party's act -- as opposed to simply being the

---

[42] In Section III we will discuss two situations in which this assumption is not appropriate. One is, as we have noted, when behavior is malicious; then optimal damages may exceed harm. *See* Section III.A *infra*. The second situation is when victims of harm are customers of the defendant and are relatively well informed about risk, in which case it may not be necessary to impose damages equal to harm to achieve optimal deterrence. *See* Section III.J *infra*.

[43] This conclusion is formally demonstrated in Shavell, *supra* note 41, and in SHAVELL, *supra* note 14, at 222-27. It should seem intuitively plausible if injurers purchase full coverage against liability, for then their risk aversion is irrelevant. If injurers do not purchase full coverage (suppose there is coinsurance or a deductible), the conclusion is not obvious because injurers do bear residual risk, but it is true nonetheless for reasons explained in the above-cited sources.

[44] The point that diversified shareholders will want a firm to be operated in an approximately risk-neutral manner is well accepted in the economic literature concerning corporate finance. *See, e.g.*, RICHARD A. BREALEY & STEWART C. MYERS, PRINCIPLES OF CORPORATE FINANCE 148-149 (4th ed. 1991). The reason shareholders prefer this is, roughly speaking, that each one, being diversified, will not worry about the riskiness of any particular firm in which he has ownership rights; thus he will vote to have the firm maximize its expected return. The result that the optimal level of damages equals harm when shareholders want the firm to act in a risk-neutral way follows directly from one of the key results in Harry A. Newman & David W. Wright, *Strict Liability in a Principal-Agent Model,* 10 INT'L REV. L. & ECON. 219 (1990). This result holds even though employees of firms will generally be risk averse. (This point does not apply, of course, if the wealth of the owners of a particular firm depends in a significant way on the profitability of that firm, as would often be the case for privately held firms owned by relatively few individuals.)

result of nature, of bad luck. This might be the case, for instance, if an individual develops a form of cancer that could have been caused by exposure to a naturally occurring carcinogen, such as radon gas, but which was in fact caused by exposure to a man-made carcinogen that was inadequately controlled by the injurer.

Second, even if the victim knows that he was injured by some party's conduct, it might be difficult for him to prove who caused the harm. The owner of a parked car that was damaged might know that it had been struck by another vehicle but not be able to identify the injurer. Or the residents near a polluted lake might know both that pollution is responsible for an unusually high rate of disease in their neighborhood and who the polluters are, but not be able to establish this in court.

Third, even if the victim knows both that he was wrongfully injured and who injured him, he might not sue the injurer. A person will not bring a suit if the legal cost and the value of the time and effort he would have to devote to the suit exceed the expected gain from suit. This will often be the case when the harm he has suffered is relatively small or when the likelihood of establishing causation is low. (Additionally, a victim might not sue if he has a distaste for the legal process.)

For one or more of the above reasons, there will be many circumstances in which injurers can escape liability -- and with a substantial probability -- for harms for which they should be held responsible.[45]

The consequences of the possibility that injurers can escape liability are clear: *If damages merely equal harm, injurers' incentives to take precautions will be inadequate and their incentive to participate in risky activities will be excessive.* Suppose that there is only a

---

[45] Consider, for example, the following evidence: (1) The likelihood of obtaining compensation for medical negligence has been found to be about 6 percent. HARVARD MEDICAL PRACTICE STUDY, PATIENTS, DOCTORS, AND LAWYERS: MEDICAL INJURY, MALPRACTICE LITIGATION, AND PATIENT COMPENSATION IN NEW YORK 7-1 (1990). (2) The average probability that an oil spill in excess of 10,000 gallons will be detected and traced to its source is approximately 60 percent. Mark A. Cohen, *Optimal Enforcement Strategy to Prevent Oil Spills: An Application of a Principal-Agent Model with Moral Hazard*, 30 J.L. & ECON. 23, 44-45 (1987). (3) The average probability of the detection of fraud is estimated to be 30 percent. Jonathan M. Karpoff & John R. Lott, Jr., *The Reputational Penalty Firms Bear from Committing Criminal Fraud*, 36 J.L. & ECON. 757, 790 (1993).

We should also note that there are reasons why parties may sometimes be found liable for harms for which they are *not* responsible. Although we do not consider this possibility in our analysis, were we to do so, it would often lower the level of damages that otherwise would be optimal.

one-in-four chance that an injurer will be found liable for a $100,000 harm and that, if found liable, will have to pay damages of $100,000. Then *on average* the injurer will pay $25,000 when he causes the harm, only a fraction of the harm caused. If the harm could have been prevented each time by taking a $50,000 precaution, the injurer will not have an adequate incentive to take the precaution -- the precaution cost will exceed his average liability cost by a substantial margin. Moreover, because the injurer will pay only $25,000 on average for a $100,000 harm, he will engage in the risky activity to an excessive degree. If the injurer is a firm, the price of its product will rise by an amount reflecting only one-quarter of the harm caused, leading consumers of the product to buy more of it -- and thereby cause more harm -- than is socially desirable.

*To remedy these problems of underdeterrence, damages that are imposed in those instances when injurers are found liable should be raised sufficiently so that injurers' average damages will equal the harm they cause.* In the example in the preceding paragraph, where the chance of being found liable for having caused a $100,000 harm is only one-in-four, suppose damages are raised to $400,000. Then, on average, the injurer will pay $100,000 when he causes the harm -- on average, every four times he causes harm he will be found liable once for $400,000. Equivalently, his total damages will tend to equal the total amount of harm that he has caused.[46] As we emphasized in Section II.A above, making injurers liable for the harm they cause will induce them to take proper precautions and to appropriately participate in risky activities.

This discussion suggests a simple formula for assuring that injurers will pay for the harms they cause: *The total damages imposed on an injurer should equal the harm multiplied by the reciprocal of the probability that the injurer will be found liable when he ought to be.*[47]

---

[46] If the injurer does not engage in an activity repeatedly, but, say, only *once*, then the injurer obviously will not pay for the harm done, even approximately: he either will pay $400,000 in this one instance (more than the $100,000 harm he caused) or will escape liability altogether. However, the injurer's *expected* damages -- the damages he will have to pay if he is found liable multiplied by the probability of being found liable -- are equal to the harm of $100,000 (he has a one-in-four chance of being found liable and made to pay $400,000).

[47] It may be helpful to state this formula algebraically. If $H$ is the harm and $P$ is the probability of being found liable, then the injurer should pay $H \times (1/P)$ -- that is, $H/P$ -- when he is found liable. Thus, the injurer's expected damages will be $P \times (H/P) = H$. The earliest mention of this formula (although in words) apparently is in

We will refer to this multiplier as the *total damages multiplier*. In the example in the preceding paragraph, the probability that the injurer would be found liable was one-in-four, or .25; thus, the multiplier is 1/.25, or 4. Since the harm was $100,000, this results in total damages imposed on the injurer of $400,000. Similarly, if the injurer would be found liable with a one-in-two chance, damages should be $200,000 -- the $100,000 harm multiplied by 2. And if the chance of liability is only one in ten, damages should be $1,000,000 -- the harm multiplied by 10. The application of this formula will guarantee that injurers pay, on average, for the harms they cause, and therefore will take proper precautions and appropriately participate in risky activities.[48]

It is important to stress that the level of damages given by the formula is optimal not only because this level remedies problems of underdeterrence, but also because it avoids problems of overdeterrence. The latter problems, described in Section II.A, would arise if damages were to exceed the optimal amount.

We will refer to the excess of total damages over compensatory damages as *punitive damages*. Thus, the *optimal* level of punitive damages from the perspective of deterrence is the level of total damages determined by the formula less compensatory damages. If an injurer has a one-in-four chance of being found liable for causing a $100,000 harm, the formula implies that total damages should be $400,000. Since $100,000 of this total represents compensatory damages, the $300,000 remainder is the optimal punitives damage amount.

The optimal level of punitive damages also can be described as a multiple of harm or, equivalently, of compensatory damages. Specifically, punitive damages should equal the harm multiplied by a factor that we will refer to as the *punitive damages multiplier*: the ratio of the

---

POSNER, *supra* note 12, at 77; see also SHAVELL, *supra* note 14, at 148, 161-62, and ROBERT COOTER & THOMAS ULEN, LAW AND ECONOMICS 391-92 (1st ed. 1988).

[48] Two qualifications should be mentioned. First, total damages should be less than the amount given by the formula in this paragraph if injurers are risk averse and cannot obtain liability insurance. This follows from our discussion in the second-to-last paragraph of Section II.A *supra*. Second, if there is a difference between an injurer's subjective belief about the probability of being found liable and the true probability, the former should be used in the formula. For simplicity, we assume in this Article that parties are aware of the actual probability of being found liable.

injurer's chance of escaping liability to the injurer's chance of being found liable.[49] In the example in the previous paragraph, the injurer has a three-in-four chance of escaping liability and a one-in-four chance of being found liable. The punitive damage multiplier therefore is .75/.25, or 3. Since the harm was $100,000, punitive damages should be three times this amount, or $300,000.

Although we refer to the excess of total damages over compensatory damages as punitive damages, the adjective "punitive" may sometimes be inapt. This is because extra-compensatory damages may be needed for deterrence purposes in circumstances in which the behavior of the defendant would not call for *punishment*. As we have explained, the deterrence goal leads us to impose such damages when injurers may escape liability. But injurers might escape liability even when their conduct is not strongly blameworthy -- suppose an injurer accidentally (perhaps not even through negligence) causes harm but the victim does not sue because he is unable to trace the harm to its source or, even if he is, because of the cost of litigation. In other words, non-blameworthy conduct might call for punitive damages to achieve proper deterrence.[50] Despite a certain inappropriateness, therefore, in using the label "punitive damages" to refer to extra-compensatory damages needed for deterrence reasons, we will continue to employ it because it is the common term for extra-compensatory damages in private civil litigation.[51]

We have several comments to make about the punitive damages formula presented in this Section. First, courts and juries[52] often will be able to apply the formula without difficulty because the formula transparently (if trivially) implies that no punitive damages are needed. In

---

[49] To see this formally, recall from note 47 *supra* that the proper level of *total* damages is $H/P$, where $H$ is the magnitude of harm and $P$ is the probability that the injurer will be found liable. This amount would be comprised of a payment of $H$ in compensatory damages and $(H/P) - H$ in punitive damages. The punitive payment can be rewritten as $[(1 - P)/P]H$. The term in brackets is the punitive damages multiplier referred to in the text.

[50] We will elaborate on this point when we discuss the reprehensibility criterion in the context of deterrence in Section III.A *infra*.

[51] For essentially the reasons we have given, Galligan, *supra* note 3, at 12-13, prefers the term "augmented awards" to "punitive damages."

[52] For simplicity, we sometimes will refer hereafter to courts, where we mean both courts and juries.

other words, there will be many circumstances in which it is obvious that there is virtually *no* chance that the injurer can escape liability -- say because the harm occurred openly and the magnitude of the harm is such that the victim(s) almost surely will bring suit. This would be the case if a building collapsed as a result of a plainly defective design[53] or if a supertanker went aground and spilled a large quantity of oil on the shoreline, where it would be observed by many people.[54] In such cases, the proper total damages multiplier is one -- that is, total damages should equal harm. Punitive damages are not needed to deter, and imposing them would result in the problems of overdeterrence discussed in Section II.A.

Second, in other circumstances, when the chance of escaping liability clearly is positive, the probability of liability still might be relatively easily calculated. For example, suppose a firm dumps toxic waste at night along an infrequently used road, but is caught as a result of the report of a driver who happened to notice the firm's activities. In such a case, pressure-sensitive recording devices laid across the road could be used to determine the volume of traffic on the road at night, and the resulting data could be employed to calculate the odds that someone would drive by during a particular interval of time. The reciprocal of this probability could then be used as the total damages multiplier.[55] In general, a careful consideration of the facts in a case often will allow a jury to make a reasonable estimate of the probability of escaping liability.[56] Expert witness testimony also may be helpful in calculating this probability.[57]

---

[53] Consider, for example, the collapse of two pedestrian skywalks at the Hyatt Regency Hotel in Kansas City in 1981. *See In re* Federal Skywalk Cases, 97 F.R.D. 380 (W.D. Mo 1983). The defendants in that case agreed not to contest liability. *Id.* at 389.

[54] *See infra* text accompanying notes 91-98 (discussing the Exxon Valdez oil spill litigation).

[55] Some estimate also would have to be made about the probability that a driver who observed suspicious behavior would report it. Obviously, the lower this probability, the higher the proper damage multiplier. (If this consideration is ignored because of the difficulty of estimating the probability, then the multiplier discussed in the text would be a lower bound for the ideal multiplier.)

[56] We discuss a number of punitive damages cases in Section II.D *infra* and explain there how the facts in those cases bear on determining the appropriate damage multiplier.

[57] It should be noted that juries often are required to consider and determine probabilities in contexts other than that of this paragraph, such as in the assessment of negligence. In deciding whether a person is negligent, the jury must ascertain how much an additional precaution would lower the probability of harm. If the reduction in the

Third, there will inevitably be circumstances in which the chance of escaping liability is difficult to estimate. To reduce the decisionmaking burden on jurors, a court could present to the jury a table with a limited number of values for the probability, such as 0.1 through 0.9 by increments of one-tenth, from which the jury must choose.[58] While this does not directly resolve the problem of determining the probability of detection, it may aid jurors in settling on a single number.[59] Even if jurors make significant errors in estimating the probability, such errors will not necessarily create a serious problem for achieving optimal deterrence: provided that the errors are not systematically biased upwards or downwards, a potential injurer will know that *on average* juries will be approximately correct, which will induce the injurer to behave properly. A further option is for the legislature to set in advance damage multipliers for separate categories of wrongful conduct, based on a rough assessment of the different chances of escaping liability in the various settings. This approach might be desirable if it is believed that jury determination of the probabilities, and therefore of the damage multipliers, would be systematically biased (but it would prevent juries from making use of any information they have about particular cases).

Fourth, it might seem problematic for application of the multiplier formula that the probability of escaping liability, and thus the multiplier, may depend on the way in which an accident is categorized. For example, consider spills of dangerous chemicals from a tanker truck on a highway. Such spills can be categorized relatively narrowly -- one category might

---

probability multiplied by the harm exceeds the cost of the precaution, it is negligent not to have taken it. *See* RESTATEMENT (SECOND) OF TORTS § 291 ("Where an act is one which a reasonable man would recognize as involving a risk of harm to another, the risk is unreasonable and the act is negligent if the risk is of such magnitude as to outweigh what the law regards as the utility of the act or of the particular manner in which it is done."). Thus, when a jury considers whether the failure to install a safety device was negligent, it must determine how much the device would have lowered the accident probability. There is no reason to believe that juries would have more difficulty in appraising probabilities in the context of calculating punitive damages than in the context of determining negligence, which is not thought to be especially problematic.

[58] Such a table is provided in the Appendix to this Article.

[59] A byproduct of restricting attention to a limited number of probabilities is that this prevents the jury from picking a damage multiplier above a certain value -- for example, the damage multiplier cannot exceed 10 if the lowest probability offered is 0.1. We discuss caps on punitive damages in Section II.C *infra* and point out that they cannot be justified as a matter of principle.

be a spill resulting from the rupture of the tank and another category might be a spill resulting from a slow leak of the tank -- or broadly -- encompassing both types of accident. A spill obviously will be much more easily detected if it is the consequence of a rupture than of a slow leak. Thus, if a rupture-caused spill is treated as a separate category, a low multiplier would be used, whereas if it is treated as an instance of a broader category of spills that includes leaks, a higher multiplier would be implied. How then should the categorization of an accident be determined for purposes of deterrence? Should categorization be narrow with separate multipliers employed for different types of chemical spills, or should categorization be broad with a single multiplier employed? The answer is, essentially, that narrow definitions of accidents and separate multipliers should be used, other things being equal. Otherwise, incentives to prevent specific types of chemical spills would tend to be distorted. This is because the multiplier used for a specific type of accident would not be tailored to it but instead would reflect the likelihood of escaping liability for a broader category of accidents. If a single multiplier is employed for all chemical spills, then, since it would exceed one, firms would have excessive incentives to avoid chemical spills for which they would definitely be found liable; thus they might spend excessively on reinforcing the tanks or on testing the tanks' pressure (say every fifteen minutes rather than every trip). Further, since the multiplier would be lower than is appropriate for slow leaks that are very likely to escape notice, firms would have inadequate incentives to prevent these spills; for instance, they might not check frequently enough for cracks in difficult-to-inspect parts of the tanks. The general point, then, is that when actors can take precautions that are particular to a type of accident, the categorization of the accident should be narrow.[60]

---

[60] If, however, all precautions are general rather than specific to a kind of accident, it does not matter whether accidents are categorized narrowly. In other words, failing to categorize narrowly will not distort which type of precaution is taken if there are no specific types of precautions to take.

To illustrate, suppose that there is one type of precaution and two kinds of accident, A and B; that the two kinds of accident are equally likely to occur; that each would cause harm of $10,000; that if A occurs, there is no likelihood of escaping liability; and that if B occurs, the likelihood of escaping liability is 50%. Under these assumptions, we will show that an injurer's expected damages if an accident occurs will equal the harm of $10,000 regardless of whether accidents are categorized narrowly or broadly. If the categorization is narrow, the multiplier for an accident of type A will be 1 and the multiplier for an accident of type B will be 2. Thus, if an accident of type A occurs, the injurer's expected damages will be $10,000 since he will pay $10,000 for sure; and if an accident of type B occurs, his expected damages will also be $10,000 since he will pay $20,000 with probability 50%. If the

Fifth, an important question about our multiplier formula arises in cases in which the defendant is a firm: Should the damage multiplier be based on the probability that the *firm* will be found liable, or on the generally lower likelihood that the responsible *employee* will be found liable? The answer is that the firm's probability is the relevant one. Consider a situation in which a firm definitely would be found liable for a harm resulting, say, from an explosion of a chemical storage tank, but the employee whose actions led to the explosion might be difficult to identify. Since the firm will have to pay for the harm for sure, punitive damages are not needed: the firm's product price and its incentives to take precautions will be correct because it will be paying for all of the harm it causes if it pays just compensatory damages. That the particular employee who caused the explosion might not be caught does not alter this point -- his escaping responsibility does not free the firm from liability. Were the firm to face punitive damages, overdeterrence would result.[61]

Sixth, it should be observed that the award of punitive damages may itself raise the probability of suit. This effect, when applicable, must be taken into account and will tend to lower the level of punitive damages that might otherwise seem appropriate. Suppose, for example, that the probability of suit for a $100,000 harm for which a party should be liable

---

categorization is broad, including both kinds of accident, the multiplier will be 1.3333 since the probability of being found liable for an accident of some type is 75% (the average of 100% and 50%). Hence, damages when the injurer is found liable will be $13,333 (= 1.3333 x $10,000) and the injurer's expected damages for an accident will again be $10,000: comprised of a 50% chance that a type A accident will occur, in which case the injurer will be found liable for $13,333 for sure, plus a 50% chance that a type B accident will occur, in which case the injurer will be found liable for $13,333 with a 50% probability (in other words, [50% x (100% x $13,333)] + [50% x (50% x $13,333)] = $10,000). Accordingly, whichever categorization is used to determine the multiplier, the injurer's incentives will be correct, and there will be no issue of distortion of types of precaution because, by hypothesis, there is only the one type.

If there were two separate types of precaution, one reducing the frequency of type A accidents and the other reducing the frequency of type B accidents, then under the broad categorization, with a multiplier of 1.3333, a potential injurer would take an excessive degree of care to reduce type A accidents and too little care to reduce type B accidents. (We note, however, that the following point can be demonstrated: If the single multiplier used under the broad categorization is always adjusted to reflect the relative likelihoods of type A and type B accidents -- they would not necessarily be equally likely when there are different types of precaution -- then use of a single multiplier would not distort incentives to take different types of precaution. Nevertheless, making such an adjustment would be administratively difficult, to say the least.)

[61] Although punitive damages should not be imposed on the firm, the firm may want, in effect, to impose punitive internal sanctions on its employees in order to deter them from acting in ways that cause harm. Additionally, the greater the likelihood that employees would escape such sanctions, the more the firm may want to spend on its monitoring and screening efforts.

would be one-third if damages are compensatory but would rise to one-half if damages are twice compensatory damages, $200,000. Suppose also that if a suit is brought, the plaintiff will prevail for sure. If the damages multiplier is based on the one-third probability of suit, it would call for total damages of $300,000. But this would be too high a level of damages since the probability of suit increases to one-half (or greater) when damages are $300,000: expected damages would be $150,000 (or greater), far exceeding the harm. In fact, damages of $200,000 would be appropriate, since then the probability of suit would be one-half, and expected damages would be $100,000. In general, there will be a level of damages that, given the resulting probability of suit, will lead to optimal deterrence.[62] But basing punitive damages on the relatively low probability of suit that would occur if just compensatory damages were awarded will tend to lead to excessive damages.

Finally, it might seem that the analysis in this Section would virtually always call for *some* punitive damage award, because there will almost always be *some* chance of escaping liability.[63] But such a conclusion ignores a factor that we have not yet mentioned, namely, the costs associated with use of the legal system.[64] Were every case to involve the calculation of the proper multiplier of harm, a new and potentially costly-to-decide issue would be introduced into litigation. This suggests that the domain of cases in which the multiplier inquiry is made should be limited. Specifically, our formula should be applied only if the likelihood of escaping liability surpasses some threshold, for that is when the problem of dilution of deterrence will be significant, making it socially worthwhile to incur the additional litigation

---

[62] To state matters formally, let $D$ be damages and let $P(D)$ be the probability of suit, which rises as $D$ rises and which is less than one. Then the claim in the text is that there is a $D$ such that $P(D)D = H$, where $H$ is the harm. Since the function $P(D)D$ is continuous in $D$, there must exist such a $D$ between $H$ and $H/P(H)$ because $P(H)H < H$ and $P(H/P(H))[H/P(H)] > P(H)[H/P(H)] = H$ (the asserted $D$ is unique because $P(D)D$ is strictly rising in $D$). We comment further on the connection between damages and the probability of suit in Section III.E *infra*.

[63] This point is noted by David G. Owen, *Civil Punishment and the Public Good*, 56 S. CAL. L. REV. 103, 113 (1982).

[64] We discuss this consideration in more detail in Section III.E *infra*.

costs associated with calculating a corrective multiplier.[65]

## C. Consistency of Punitive Damages Law with the Basic Theory of Deterrence

We now will relate our analysis to certain important aspects of legal doctrine concerning punitive damages as well as to legislation that imposes caps on punitive damages.

As noted in the introduction, one of the two main purposes of punitive damages concerns deterrence. The courts state, for example, that punitive damages are intended "to deter the wrongdoer and others from committing similar wrongs in the future."[66] Given that achieving proper deterrence is an avowed goal of courts, *it follows inescapably from the logic of deterrence theory that courts should take the punitive damages formula presented above into explicit account*. Otherwise courts cannot responsibly weigh the proper punitive damages amount for purposes of achieving deterrence against the proper amount for purposes of achieving the other main purpose of punitive damages, punishment.[67]

However, courts' determination of punitive damages does not reflect in any reasonably clear manner the formula that achieves optimal deterrence. Although the courts do consider the magnitude of harm in assessing the proper level of punitive damages, they do not use harm as the base to be multiplied by an appropriate damage multiplier. Rather, courts take harm

---

[65] Actually, this statement oversimplifies matters. To decide when it is worthwhile to calculate a damage multiplier, one would in principle need to take into account not just the probability of escaping liability, but also the magnitude of the harm and the costs of precautions. For even if the probability of escaping liability is high, if the harm is very low or if the costs of additional precautions are very high, then it may not be worthwhile to incur the additional litigation costs to determine the proper damage multiplier.

Additionally, if the determination of punitive damages is not done on a case-by-case basis, it might be desirable to award punitive damages even when the probability of escaping liability is low. Suppose a fixed multiplier is applied in all cases in which the probability is small but positive. Then there would be no additional litigation costs associated with calculating the multiplier in each case; and if the fixed multiplier is set appropriately, incentives will be improved overall relative to what they would be if no multiplier were employed.

[66] Green Oil Co. v. Hornsby, 539 So. 2d 218, 222 (Ala. 1989). This case was endorsed by the U.S. Supreme Court in *Pacific Mutual Life Insurance Co. v. Haslip*, 499 U.S. 1, 20-24 (1991). For further support for the proposition that deterrence is one of the central purposes of punitive damages, see also the cases cited in note 5 *supra*. Indeed, in Maine, deterrence is the only justification for punitive damages. *See* Foss v. Maine Turnpike Auth., 309 A.2d 339, 345 (Me. 1973) (citing Allen v. Rossi, 146 A. 692 (Me. 1929)).

[67] How the deterrence goal and the punishment goal should be reconciled when they are in conflict is discussed in Section V *infra*. See also the Appendix to this Article.

into account in a vague way through application of the general principle that punitive damages should bear a "reasonable relationship" to compensatory damages.[68] They do not explain what this relationship should be and, even when they identify a ratio of punitive damages to compensatory damages that they find excessive, they do not supply a basis for selecting the particular ratio identified.[69]

As the reader knows, our analysis implies a simple and precise relationship between punitive damages and harm: punitive damages should equal the harm multiplied by what we

---

[68] 1 LINDA L. SCHLUETER & KENNETH R. REDDEN, PUNITIVE DAMAGES § 6.1(C) (3d ed. 1995) ("As a general rule, the punitive damages award must bear a reasonable relation to the amount of actual damages awarded."). The Supreme Court has endorsed the reasonable relationship notion in its decisions. See, e.g., BMW of N. Am., Inc. v. Gore, 116 S.Ct. 1589, 1601 (1996) ("The principle that exemplary damages must bear a 'reasonable relationship' to compensatory damages has a long pedigree," with citations dating back to 1852.); *Haslip*, 499 U.S. at 21-22 (endorsing Alabama's criteria for the post-verdict review of punitive damages, which includes "whether there is a reasonable relationship between the punitive damages award and the harm likely to result from the defendant's conduct as well as the harm that actually has occurred"; the "review ensures that punitive damages awards . . . have some understandable relationship to compensatory damages").

[69] The Supreme Court has recently presented three "guideposts" for determining whether a punitive damages award is excessive. *Gore*, 116 S.Ct. at 1598. "The second [of these,] and perhaps [the] most commonly cited indicium of an unreasonable or excessive punitive damages award is its ratio to the actual harm inflicted on the plaintiff." *Id.* at 1601. The Court, however, rejects the possibility that excessiveness can be determined by "a simple mathematical formula, even one that compares actual and potential damages to the punitive award." *Id.* at 1602 (citing TXO Prod. Corp. v. Alliance Resources Corp., 509 U.S. 443, 458 (1993)). In his concurrence in *Gore*, Justice Breyer points to problems with the reasonable relationship standard, arguing that, at least as it is interpreted by the Alabama courts, it "does little to guide a determination of what counts as a 'reasonable' relationship . . . . To find a 'reasonable relationship' between purely economic harm totaling $56,000, without significant evidence of future repetition, and a punitive award of $2 million is to empty the 'reasonable relationship test' of meaningful content." *Id.* at 1606.

Courts often make statements of the following sort: "[A]lthough there is no fixed ratio by which to determine the propriety of a punitive damage award, punitive damages should bear a reasonable relationship to the compensatory damages awarded." Little v. Stuyvesant Life Ins. Co., 136 Cal. Rptr. 653, 663 (Cal. Ct. App. 1977). The *Little* court followed that sentence with the observation that "[h]ere the ratio of punitive damages to compensatory damages is in excess of 14 to 1 and in dollar amount the punitive damage award exceeds the compensatory award by almost two and a third million dollars," *id.* at 664, and the court relied on these facts, among others, in concluding that the punitive damage award in the case was excessive, *id.*

In an important early article on punitive damages, Clarence Morris criticized approaches like that of the *Little* court:

> Courts often insist that "punitive damages must bear some relation to actual damages," and attempt to test verdicts in terms of mathematical ratios. The opinions contain statements to the effect that a verdict for punitive damages $x$ times as great as the actual damages is clearly excessive. . . .
>
> This test is probably more often a rationalization of results than a means of obtaining them. The proper ratio between actual damages and punitive damages is placed at a figure which supports the judge's view of the verdict . . . .

Clarence Morris, *Punitive Damages in Tort Cases*, 44 HARV. L. REV. 1173, 1180 (1931).

referred to as the punitive damages multiplier. If punitive damages are to achieve appropriate deterrence, the "reasonable relationship" criterion has to be interpreted *in this specific way*. Any other relationship between punitive damages and compensatory damages will lead either to inadequate deterrence or excessive deterrence.

Courts also do not pay systematic attention to the probability of escaping liability, even though this probability is the central element in determining the appropriate damage multiplier for purposes of achieving proper deterrence. Sometimes the probability of escaping liability is alluded to, but its importance with respect to deterrence rarely is recognized. For example, courts occasionally include as a consideration in determining the level of punitive damages whether the defendant has attempted to conceal or cover up his conduct.[70] Courts usually do this, however, in the context of assessing the reprehensibility of the defendant's conduct;[71] they do not appreciate that evidence of attempted concealment or cover-up should influence the calculation of the defendant's chance of escaping liability. Additionally, courts sometimes mention that the cost of litigation should be taken into account "so as to encourage plaintiffs to bring wrongdoers to trial."[72] This factor obviously is related to the injurer's chance of

---

[70] *See, e.g.*, Gamble v. Stevenson, 406 S.E.2d 350, 354 (S.C. 1991) (establishing "defendant's awareness or concealment" as one factor to consider in the post-trial review of jury awards of punitive damages); *Green Oil*, 539 So. 2d at 223 (concealment or cover-up is relevant in determining degree of reprehensibility, which influences the level of punitive damages).

[71] As just observed, see note 70 *supra*, this is true in *Green Oil*, 539 So. 2d at 223. *See also* Garnes v. Fleming Landfill, Inc., 413 S.E.2d 897, 909 (W. Va. 1991) (in determining the reprehensibility of defendant's conduct, the jury should consider "whether he attempted to conceal or cover up his actions or the harm caused by them").

[72] *Green Oil*, 539 So. 2d at 223; *see also Garnes*, 413 S.E.2d at 909 (instructing trial courts to consider the costs of litigation in reviewing punitive damages awards, because "we want to encourage plaintiffs to bring wrongdoers to trial").

In this vein, a Texas Court of Appeals affirmed an award of $4,500 in exemplary damages, in an action for invasion of privacy from telephone harassment, where actual damages were two dollars and attorney fees were $4,462.52. *See* Donnel v. Lara, 703 S.W.2d 257, 262 (Tex. App. 1985, writ ref'd n.r.e.) (superseded by statute, *see* Harkins v. Crews, 907 S.W.2d 51, 61 (Tex. App. 1995, writ denied)). The court noted plaintiffs "had obligated themselves to pay reasonable attorney fees as a necessary prerequisite for obtaining relief through the courts," *id.* and held that "the amount exemplified an accurate application of the purposes for exemplary damages--to punish and deter similar wrongs in the future." *Id.*

Likewise, under Mississippi law, juries may consider attorneys' fees in determining punitive damage awards, because one purpose of the award is to compensate the plaintiff "for service to the public in bringing the action, which should act as a deterrent of similar acts of wrongdoing to other members of the public." Mutual Life Ins. Co. v. Estate of Wesson, 517 So. 2d 521, 532 (Miss. 1987); *see also* Andrew Jackson Life Ins. Co. v. Williams,

escaping liability since one reason an injurer might not be found liable is that he is not sued (a point we noted in Section II.B above). Thus, courts occasionally refer to considerations that bear on the probability that a defendant would have escaped liability. But they rarely explain in a direct and systematic way how this probability should be used to determine the proper level of damages for deterrence purposes.[73]

Courts also generally pay insufficient attention to the potential problem of overdeterrence. Judicial opinions only infrequently mention this issue[74] and none of the lists of factors used by courts in determining punitive damages include overdeterrence as a consideration.[75] As we have emphasized, however, damages that exceed those given by the

---

566 So. 2d 1172, 1189-90 (Miss. 1990) and sources cited therein.

[73] A few courts have explicitly recognized the importance of the factor of the defendant's chance of escaping liability. *See* Kemezy v. Peters, 79 F.3d 33, 35 (7th Cir. 1996) (Posner, C.J.) ("When a tortious act is concealable, a judgment equal to the harm done by the act will underdeter. Suppose a person who goes around assaulting other people is caught only half the time. Then in comparing the costs, in the form of anticipated damages, of the assaults with the benefits to him, he will discount the costs (but not the benefits, because they are realized in every assault) by 50 percent, and so in deciding whether to commit the next assault he will not be confronted by the full social costs of his activity."); Zazú Designs v. L'Oréal, S.A., 979 F.2d 499, 508 (7th Cir. 1992) (Easterbrook, J.) ("Punitive damages are appropriate when some wrongful conduct evades detection; a multiplier then both compensates and deters."); FDIC v. W.R. Grace & Co., 877 F.2d 614, 623 (7th Cir. 1989) (Posner, J.) ("The most straightforward rationale for punitive damages . . . is that they are necessary to deter torts or crimes that are concealable. Suppose the average defrauder is brought to book only half the time. To confront him with a sanction that will make fraud worthless to him and thus deter him, it is necessary that when he is caught he be made to pay twice as much as his profits."); *see also Gore*, 116 S.Ct. at 1602 (Stevens, J.) ("A higher ratio [of punitive damages to compensatory damages] may also be justified in cases in which the injury is hard to detect . . ."). Justice Breyer's concurrence in *Gore* mentions economic theories of punitive damages that focus on ensuring that a wrongdoer pays for the total cost of the harm caused. *Id.* at 1607. He correctly interprets this theory as permitting juries "to calculate punitive damages by making a rough estimate of global harm, [and] dividing that estimate by a similarly rough estimate of the number of successful lawsuits that would likely be brought. . ." *Id.*

[74] One exception is Justice Breyer's concurrence in *Gore*, 116 S.Ct. at 1604. He observes that damages greater than the total harm caused will "'over-deter' by leading potential defendants to spend more to prevent the activity that causes the economic harm, say, through employee training, than the cost of the harm itself." *Id.* at 1607-08; *see also* Jones v. Reagan, 696 F.2d 551, 554 (7th Cir. 1983) (Posner, J.) ("[I]f considerations of deterrence are to be brought to center stage, the potential for overdeterrence must also be considered."); Roginsky v. Richardson-Merrell, Inc., 378 F.2d 832, 839-41 (2d Cir. 1967) (Friendly, J.) (noting that allowing multiple punitive damage awards for negligence in the manufacture of goods gives rise to the danger of "overkill" and "needless deterrence").

[75] For example, the *Green Oil* factors, which were endorsed by the U. S. Supreme Court, fail to mention overdeterrence. *See Green Oil*, 539 So. 2d at 223-24. Similarly, state statutes outlining the factors to be considered in awarding punitive damages do not include the danger of overdeterrence. *See* KAN. STAT. ANN. § 60-3701(b) (1995); MINN. STAT. ANN. § 549.20 subd. 3 (West 1996); MONT. CODE ANN. § 27-1-221(7)(b) (1996); OR. REV. STAT. § 30.925(2) (1995).

formula may result in wasteful precautions and the withdrawal of socially valuable products and services from the marketplace.

Not only do courts fail to correctly consider the factors that *are* relevant to proper deterrence, they also err because they consider a variety of factors that generally *are not* relevant to deterrence, including the reprehensibility of defendants' conduct and the wealth of defendants. Why these factors ordinarily should not be taken into account if the goal is to promote proper deterrence will be discussed below (in Section III), but the point we want to make here is that their consideration causes punitive damages to deviate further from the amount given by our formula.

Some aspects of legislation governing punitive damages also are inconsistent with deterrence theory. Notably, many jurisdictions have imposed caps of various kinds on punitive damages awards: an absolute ceiling (for example, $350,000 in Virginia), a maximum ratio of punitive damages to compensatory damages (three times compensatory damages in Florida), or both.[76] It should be apparent that such caps cannot be justified on deterrence grounds -- they might preclude the proper award of punitive damages. For example, suppose that the harm caused by an injurer is $100,000 and that he has only a one-in-ten chance of being found liable. Then the optimal level of punitive damages is $900,000 or nine times compensatory damages (so that total damages, including compensatory damages, are ten times the harm). This absolute amount and this ratio would exceed the caps of the majority of states that impose limits,[77] yet under the circumstances posited, a punitive damages award of this magnitude, and that has this relationship to compensatory damages, is needed for proper deterrence.

Our criticism of caps is not meant to deny that, if jury awards of punitive damages are

---

[76] *See Gore*, 116 S.Ct at 1618-19 (Ginsburg, J., dissenting) (surveying state caps on punitive damage awards in appendix to dissenting opinion).

[77] Justice Ginsburg lists sixteen states in which caps on punitive damages have been enacted or proposed. In thirteen of those states (Colo., Conn., Del., Fla., Ga., Ill., Ind., Md., Nev., N.J., N.D., Tex., and Va.), the award described in the text would exceed the cap. *See id.*

thought to be systematically excessive, caps might beneficially constrain such awards.[78] But if it is believed that punitive damages are not biased upwards, they should not be limited for the reasons given in the preceding paragraph.

### D. Punitive Damage Cases

We briefly consider here three prominent punitive damage cases in light of the deterrence principles discussed above. Our primary objective is to state what deterrence theory suggests about the appropriate level of punitive damages in these cases, given their facts and circumstances, *not* to analyze the legal doctrines that were applied or developed in them. (We have commented on the legal doctrine governing punitive damages to some extent in the previous subsection, and will discuss it extensively in Section III.)

*BMW of North America, Inc. v. Gore.*[79] In this case the plaintiff, Dr. Ira Gore, Jr., was the purchaser of a new BMW sedan from an Alabama dealer. He subsequently learned that BMW of North America had repainted part of the car because of damage to the car prior to its arrival in the United States, although BMW had not disclosed this fact. The jury awarded Dr. Gore compensatory damages of $4,000 for diminution in the value of the car due to its having been repainted, and punitive damages of $4 million. The Alabama Supreme Court reduced the punitive award to $2 million, which the U.S. Supreme Court held to be grossly excessive. Upon reconsideration by the Alabama Supreme Court, the punitive award was reduced to $50,000.[80]

Consider the probability that BMW would escape liability for having sold a repainted car as new. The determination of this probability involves two factors. One is the possibility that BMW would escape notice for having repainted a car, and the other is the likelihood that a

---

[78] *See, e.g.*, Dan Quayle, *Civil Justice Reform*, 41 AM. U. L. REV. 559, 564-65 (1992) (arguing that "the current approach to punitive damages will continue to generate disproportionately high awards in a random and capricious manner," and that one aspect of reform should be to limit the amount of punitive damages to "the full amount of compensatory damages").

[79] 116 S.Ct. 1589 (1996).

[80] *Alabama Court Slashes Punitive Award In Case Involving Repainted BMW Car*, WALL ST. J., May 12, 1987, at B10.

purchaser who did discover that his car had been repainted would sue. Dr. Gore drove the car for nine months without detecting any abnormalities regarding the paint on his car. It was only after he took his car to a detail shop that he learned that it had been repainted. It is plausible to suppose, therefore, that many purchasers of repainted cars sold as new would never discover the fact of the repainting.

Whether an owner who did discover that his car had been repainted would sue would depend on the costs to him of suit (time and out-of-pocket expense) and the amount that he could collect. If the harm is as low as the jury found in the *Gore* case, $4,000, it would seem that many owners -- or the lawyers they might hire on a contingency fee -- would not have a sufficient financial incentive to sue to collect this amount. There may have been a significant chance, therefore, that BMW would have escaped liability if damages were merely compensatory, due to victims' inadequate motive to sue.

Information was provided in *Gore* that would be useful in estimating the probability of BMW's being found liable. Among the facts established at trial was that there had been 14 new BMW cars that had been repainted in Alabama, including Dr. Gore's, and one other prior suit against BMW by an owner of one of these cars.[81] If none of the other Alabama victims of repainting were to sue,[82] the probability of detection and liability might be thought to be two-in-fourteen,[83] in which case the total damages should be seven times the $4,000 harm, or

---

[81] The other plaintiff, Dr. Thomas Yates, was awarded $4,000 in compensatory damages but nothing in punitive damages. *See Yates v. BMW of N. Am., Inc.*, 642 So. 2d 937, 939 (Ala. Civ. App.), *cert. quashed*, 642 So. 2d 937 (Ala. 1993). It should be noted that in the United States as a whole, BMW had sold 983 cars as new after repainting. *See Gore*, 116 S. Ct. at 1593. In this example, we are restricting our attention to the subset of cars sold in Alabama because the Alabama Supreme Court, in reviewing *Gore*, limited consideration in this way. Obviously, our analysis could be extended to consider BMW's conduct nationwide.

[82] In fact, there were other suits. *See BMW of N. Am., Inc. v. Gore*, 646 So.2d 619, 626 n.4 (Ala. 1994) (listing the 24 other cases brought by Gore's attorney against BMW for similar conduct). The assumption in the text that there were two suits is made for illustrative purposes.

[83] In general, the proper approach to calculating the probability of liability would be to use all available information about the likelihood of detection and suit. This might include, for example, information about the frequency of suit against BMW under similar circumstances in other states, information about how often car owners take their cars to detail shops, and information about the likelihood of lawyers taking cases with different size stakes.

$28,000.[84] Of that total, $4,000 would represent compensatory damages and $24,000 would be labeled punitive damages. By this reasoning, it is evident that the $2 million punitive award initially approved by the Alabama Supreme Court was grossly excessive, and that its subsequent reduction to $50,000 was much more reasonable.[85]

*Pacific Mutual Life Insurance Company v. Haslip.*[86] This case involved an insurance agent who misappropriated premium payments. The insurance policy in question was a group health plan sold to the municipality of Roosevelt City, Alabama. When Cleopatra Haslip, an employee of Roosevelt City, was hospitalized, she apparently did not know that the policy had lapsed due to the agent's misappropriation. The hospital and her physician sought payment from her. She and other Roosevelt City employees sued the agent and the Pacific Mutual Life Insurance Company for fraud.[87] The jury awarded her total damages of $1,040,000, of which $200,000 can be interpreted as compensatory damages and $840,000 as punitive damages.[88] The award was affirmed by the trial court, the Supreme Court of Alabama, and the U.S. Supreme Court.

The key issue in this case in terms of deterrence is whether there is a significant chance that an insurance company whose agent misappropriates premiums will escape liability for coverage that individuals expected to have. (The focus should be on the company's chance of escaping liability rather than on the agent's, for the reason we explained near the end of

---

[84] In fact, the calculation of the multiplier may be more complicated because the award of punitive damages will itself affect the probability of suit. As we said in Section II.B *supra*, total damages should be such that the probability of suit induced by that award of damages results in expected damages equal to harm. Therefore, if an award of $28,000 would induce more than two plaintiffs to sue, the proper multiplier might be less than seven.

[85] For another economically-oriented discussion of *Gore*, see Rubin, Calfee & Grady, *supra* note 12, Their discussion does not emphasize the point that we make here.

[86] 499 U.S. 1 (1991).

[87] In fact, the insurance policy that lapsed was not Pacific Mutual Life's policy, but rather a policy of another company, Union Fidelity Life Insurance Company, which the agent also was representing. *See id.* at 4-5. However, premiums for the Union policy were collected through Pacific Mutual Life's Birmingham office. *See id.* at 5. Pacific Mutual Life was sued for fraud under a theory of respondeat superior. *See id.* at 6.

[88] Although it was not entirely clear how the jury apportioned the total award between compensatory and punitive damages, the U.S. Supreme Court, based on certain evidence, presumed that not more than $200,000 of the total represented compensatory damages and not less than $840,000 represented punitive damages. *Id.* at 7 n.2.

Section II.B.) Obviously, if a policy has been invalidated because of an agent's misappropriation of premium payments, this fact will come to the attention of a person who applies for coverage under the policy. If the insurance company then does not pay the individual voluntarily, the individual probably would sue the company if the amount at stake is large enough.

In the present case, the compensatory damages were, as noted, $200,000. However, less than $4,000 of this amount represented out-of-pocket expenditures, the rest apparently consisting of non-economic losses such as emotional distress.[89] It seems reasonable to suppose that recovery of the $4,000 out-of-pocket loss is more probable than recovery of the $196,000 non-economic loss. If the likelihood of the latter recovery is sufficiently low, a lawsuit probably would not be brought. Conversely, if this likelihood is high, a lawsuit would be much more certain. A related consideration is that three other Roosevelt City employees joined Cleopatra Haslip in suing the defendants. Their awards totaled approximately $38,000.[90] The prospect of obtaining this additional amount clearly would increase the incentive to sue. On balance, therefore, while a suit seems reasonably likely in the circumstances of the *Haslip* case, there are some countervailing considerations that might justify a modest punitive damage award, to offset the chance that a lawsuit would not be brought.

*The Exxon Valdez Oil Spill Litigation.*[91] In this case the defendant's supertanker, the *Exxon Valdez*, ran aground on a reef in Prince William Sound in Alaska, resulting in the spilling of 11 million gallons of oil[92] and the pollution of over 1,000 miles of Alaskan coastline.[93] The supertanker's captain, Joseph Hazelwood, had previously been treated for

---

[89] Mrs. Haslip's out-of-pocket expenses were "less than $4,000." *Id.*

[90] The jury awarded compensatory damages for the other respondents in the following amounts: Hargrove $10,288; Craig $12,400; and Calhoun $15,290. See id. at 6.

[91] *In re* Exxon Valdez, No. A89-0095-CV (D. Alaska, judgment entered Sept. 24, 1996).

[92] *See, e.g.*, Charles McCoy, *Exxon's Corp. 's Settlement Gets Court Approval*, WALL ST. J., Oct. 9, 1991.

[93] *Fishermen Block Tankers*, Wash. Post, Aug. 22, 1993, at A9 (The *Exxon Valdez* "polluted thousands of miles of coastline.").

alcohol abuse and was found to have violated regulations governing alcohol consumption in connection with the accident at issue.[94] In the private civil litigation against Exxon stemming from the accident, the plaintiffs -- various classes of fishermen and Alaskan natives -- were awarded several hundred million dollars in compensatory damages[95] and $5 billion in punitive damages.[96] The punitive damage award was affirmed by the trial judge and is being appealed.[97]

It seems clear that in the circumstances of the *Exxon Valdez* accident, there is essentially no chance that the defendant company, the Exxon Corporation, could escape liability. An accident of this magnitude obviously would be noticed. Moreover, since the tanker was stuck on a reef, the identity of the injurer was plain. And given the substantial compensatory damages involved, in the hundreds of millions of dollars, a lawsuit certainly could be expected. Thus, according to our analysis, no punitive damages are needed, or appropriate, in the circumstances of this case because there is essentially no chance that the injurer could escape liability for compensatory damages. (In other contexts involving oil spills -- such as the intentional dumping of small amounts of waste oil that is unlikely to be detected or traced to the spiller -- some punitive damages would be appropriate.[98])

---

[94] *See* Seth Mydans, *Captain in Alaska Oil Spill Loses License for Nine Months*, N.Y. TIMES, July 26, 1990, at A12 (An administrative law judge for the Coast Guard found Hazelwood guilty of consuming alcohol within four hours of sailing; Hazelwood had pleaded no contest to the charge.); *A Question Recurs: Was Hazelwood Drunk?*, N.Y. TIMES, Feb. 25, 1990, at 29.

[95] The jury awarded $287 million as compensation for fishing losses. *See In re* Exxon Valdez, No. A89-0095-CV, 1995 WL 527988, at *5 (D. Alaska Jan. 27, 1995). The court noted that, including other verdicts and settlements, the dollar amount of harm caused by the spill was between $288.7 and $418.7 million (including the $287 million verdict). *See id.*

[96] Caleb Solomon, *Exxon Is Told to Pay $5 Billion for Valdez Spill*, WALL ST. J., Sept. 19, 1995, at A3.

[97] In June 1997, Exxon appealed the $5 billion punitive damages award entered against it in the *Exxon Valdez* case. *See Exxon Corp. Submits Brief Outlining Valdez Appeal*, WALL ST. J., June 20, 1997, at A4.

[98] This would be the case, for example, in the circumstances described in Matthew L. Wald, *Royal Caribbean Cruise Line Indicted on Charges of Dumping Oil*, N.Y. TIMES, Dec. 20, 1996, at A26 (cruise line alleged to have "routinely dumped waste oil from five of its ships for years and falsified its log books to hide its activities").

## III. DETERRENCE: EXTENSIONS OF THE BASIC THEORY

In this Section we will discuss several important doctrinal and policy-related issues in punitive damages law from the perspective of the deterrence principles developed in Section II. Most of these topics have received substantial attention in judicial opinions (such as the reprehensibility of the defendant's conduct), some have been considered primarily in a legislative context (whether the state should receive a portion of a punitive damages award), and others apparently have not been addressed in either setting (the status of the plaintiff as a customer or a third party).

### A. Reprehensibility of Conduct

The law requires that a defendant be found to have acted in a reprehensible manner -- in a way that is egregious, malicious, or undertaken with reckless disregard for the rights of others -- before punitive damages can be imposed on him.[99] If a defendant is found to have so acted, the *degree* of his reprehensibility often is treated as a key factor in determining the level of punitive damages.[100] Indeed, the U.S. Supreme Court in the *Gore* case observed that this factor is "[p]erhaps the most important indicium of the reasonableness of a punitive damages

---

[99] *See* 1 SCHLUETER & REDDEN, *supra* note 68, at 264 ("In order to receive punitive damages, a plaintiff must show that the defendant acted with malice, either actual or legal."); *see also* Masaki v. General Motors Corp., 780 P.2d 566, 570 (Haw. 1989) (punitive damages "are awarded only when the egregious nature of the defendant's conduct makes such a remedy appropriate"); Barnhouse v. Hawkeye State Bank, 406 N.W.2d 181, 184 (Iowa 1987) ("An award of punitive damages is appropriate when a party acts with actual or legal malice.") (citations omitted); Nappe v. Anschelewitz, Barr, Ansell & Bonello, 477 A.2d 1224, 1230 (N.J. 1984) ("To warrant a punitive award, the defendant's conduct must have been wantonly reckless or malicious.") (citations omitted); Hood v. Fulkerson, 699 P.2d 608, 611 (N.M. 1985) (Punitive damages "may be awarded only when the conduct of the wrongdoer may be said to be maliciously intentional, fraudulent, oppressive, or committed recklessly or with a wanton disregard to the plaintiffs' rights.").

[100] *See, e.g.*, Green Oil Co. v. Hornsby, 539 So. 2d 218, 223 (Ala. 1989) ("The degree of reprehensibility of the defendant's conduct should be considered" when "determining whether a punitive damages award is excessive or inadequate"), *endorsed in* Pacific Mutual Life Ins. Co. v. Haslip, 499 U.S. 1 (1991); Neal v. Farmers Ins. Exch., 582 P.2d 980, 990 (Cal. 1978) (Among factors to consider in assessing punitive damages is "the particular nature of the defendant's acts in light of the whole record; clearly, different acts may be of varying degrees of reprehensibility, and the more reprehensible the act, the greater the appropriate punishment, assuming all other factors are equal."); McNeill v. Allen, 534 P.2d 813, 820 (Colo. Ct. App. 1975) ("[T]he purpose of punishment and deterrence may best be served by relatively higher or lower exemplary damages according to the nature of the wrongful conduct."); Ultimate Chem. Co. v. Surface Transp. Int'l, Inc., 658 P.2d 1008, 1012 (Kan. 1983) (Among the factors to consider in assessing punitive damages are "the nature, extent, and enormity of the wrong.").

award."[101] The reprehensibility of the defendant's conduct also was one of the factors listed by the Court in the *Haslip* opinion.[102]

Should reprehensibility *per se* affect the imposition of punitive damages, given the goal of deterrence?[103] In this Section we explain why our answer is that it generally should not.[104] However, there is an important exception to this conclusion when injurers' gains do not count in social welfare, which we suggest is often the case when injurers act maliciously.[105] This exception, we will suggest, only possibly applies to individual defendants, not corporate defendants.

As the reader knows from Section II, under standard assumptions the imposition of damages equal to harm, appropriately multiplied to reflect the probability of escaping liability, achieves proper deterrence. That a defendant's conduct can be described as reprehensible is in itself irrelevant. Rather, the focus in determining punitive damages should be on the injurer's chance of escaping liability.

Making punitive damages depend on reprehensibility will distort deterrence in two ways. First, excessive damages may be imposed, when reprehensible conduct occurs in situations in which an injurer is virtually certain to be found liable. Suppose that a surgeon, through extreme negligence, fails to remove a surgical tool from the body of his patient and that this omission leads to great pain and suffering. If, because of the magnitude of the harm suffered by the patient and the unmistakable error of the surgeon, there would be a high

---

[101] BMW of N. Am., Inc. v. Gore, 116 S.Ct. 1589, 1591 (1996).

[102] See note 100 *supra*.

[103] As previously noted, we discuss the significance of the reprehensibility criterion to the punishment goal of punitive damages in Section IV *infra*.

[104] For the most part, other commentators who have considered punitive damages in terms of the deterrence goal have agreed that the reprehensibility of the defendant's conduct is not a relevant factor. *See, e.g.*, Dobbs, *supra* note 3, at 860-63; Galligan, *supra* note 3, at 62-64.

[105] Dorsey Ellis and Robert Cooter also find that the reprehensibility of a defendant's conduct is relevant to deterrence when injurers' gains do not count in social welfare -- that is, when their gains are socially illicit. *See* Cooter, *Economic Analysis*, *supra* note 12, at 86-89; Ellis, *supra* note 12, at 31-33. For more formal treatments of this issue, see SHAVELL, *supra* note 14, at 146, 159-61; Diamond, Efficiency Effects, *supra* note 12, at 8-12.

probability that the surgeon will be sued and found liable for his reprehensible conduct, extra-compensatory damages are neither needed nor appropriate. Similarly, consider a newspaper reporter who, out of reckless disregard for the truth, confuses one firm's product that is safe with another firm's product that is dangerous, substantially damaging the former firm's business reputation and profitability. Here, too, we might expect that a suit and a finding of liability would be very likely, in which case extra-compensatory damages would be excessive. Thus, *even for conduct that is reprehensible*, if there is little chance of escaping liability, compensatory damages alone will achieve appropriate deterrence, and punitive damages will result in overdeterrence.

The reader might wonder, though, how there can be overdeterrence of reprehensible acts, since society evidently has an interest in deterring such acts completely. To illustrate that overdeterrence still can occur, consider the example of the surgeon. If the magnitude of damages is very high, we can imagine that, to reduce the chance of his leaving a surgical tool in a patient, he might hire another surgeon to monitor his actions or he might dramatically increase the time he spends on each operation. Even if such responses succeed in preventing the recurrence of this event, they may be at too great a cost, especially if the likelihood of leaving a surgical tool in a patient is very low anyway. In other words, it might not be socially worthwhile for the surgeon to take the measures needed to eliminate the possibility of his being extremely negligent. Yet a level of liability in excess of that given by the damage formula would improperly encourage him to take these measures.[106]

The problem of overdeterrence also can arise in connection with the reprehensible acts of employees of corporations. Employees obviously cannot be controlled perfectly by a corporation, even though a corporation can improve its ability to prevent its employees from committing reprehensible acts by screening them before hiring them and monitoring their conduct afterwards. If damages exceed the level determined by the damage formula, however,

---

[106] Dorsey Ellis and Robert Cooter also find that the reprehensibility of a defendant's conduct is relevant to deterrence when injurers' gains do not count in social welfare -- that is, when their gains are socially illicit. *See* Cooter, *Economic Analysis*, *supra* note 12, at 86-89; Ellis, *supra* note 12, at 31-33. For more formal treatments of this issue, see SHAVELL, *supra* note 14, at 146, 159-61; Diamond, Efficiency Effects, *supra* note 12, at 8-12.

the corporation may be led to spend excessively on screening and monitoring efforts to forestall reprehensible behavior.[107] This might be true of a newspaper, for instance, if it faced punitive damages for false reporting due to extreme negligence, as in our example of the reporter who confused two firms' products. In response, the newspaper might be induced to routinely assign two reporters to every story even if doing so is not socially worthwhile given the cost of this practice and the reduction in risk of reprehensible behavior that would be accomplished thereby.

Not only can attention to reprehensibility result in the imposition of punitive damages that are excessive, such attention also may lead to the converse problem: the failure to employ punitive damages when they are needed for proper deterrence. This problem will occur if an individual engages in conduct that is harmful, though *not* reprehensible, and he is likely to escape liability. Suppose that a toxic waste disposal truck develops a leak (say from rust) that results in waste spilling onto a highway at night when no one would be likely to notice it. The driver of the truck may have performed a proper inspection before departing, and the company may have reasonable maintenance policies. Although the leak is not caused by anyone's reprehensible behavior, substantial extra-compensatory damages may be called for if the leak is discovered, in order to offset the significant likelihood that the injurer would not be identified and held responsible for the harm.

It is clear from the foregoing discussion that the stress courts place on reprehensibility of conduct in deciding on punitive damages cannot be justified on grounds of deterrence. A minor qualification of this point is that, as we observed in Section II.C, courts treat attempts by the defendant to conceal wrongdoing as a factor that enhances reprehensibility, and thus punitive damages. This makes rough sense because such behavior clearly reduces the probability of liability. But, as suggested above, the link that courts make between this

---

[107] For example, in the illustration of a firm's screening decision in note 24 *supra*, if damages exceed the amount determined by the damage formula, the firm might be led to spend $4,000 per applicant on screening, which would be socially excessive given the assumed benefit from this level of screening. The point that the imposition of punitive damages may lead corporations to spend excessively to forestall reprehensible behavior by their employees is made by Daniel R. Fischel & Alan O. Sykes, *Corporate Crime*, 25 J. LEGAL STUD. 319, 348 (1996).

behavior and punitive damages is vague in nature.[108] We believe that it would be preferable to use evidence of concealment directly to aid in the determination of the chance that the defendant might have escaped liability, rather than as a factor in determining reprehensibility.

Finally, although the reprehensibility of a defendant's conduct should not be used *per se* as a basis for imposing punitive damages to achieve proper deterrence, such conduct may sometimes provide useful information about the defendant's chance of escaping liability. Everything else equal, the lower the chance of being found liable, the lower will be an individual's level of care. Therefore, low care may suggest a low probability of liability,[109] and thus the appropriateness of a higher level of punitive damages according to our formula.

Let us now turn to the important exception to our general conclusion about reprehensibility, which, as noted above, arises if injurers' gains are not counted in social welfare. Suppose that a person, out of spite, punches another individual; his purpose is to cause harm to the victim. Society might well treat the pleasure the injurer obtains from this act as *socially illicit*, not to be counted in social welfare.[110] If so, the act should be deterred completely -- it produces no social gain, only harm. To accomplish this, damages must exceed

---

[108] *See* text accompanying notes 70-73. For example, in *Green Oil*, because reference to concealment and cover-up occurs in the context of the reprehensibility criterion, it is difficult to infer how such behavior should affect punitive damages. In part, this is because there is no discussion of how the degree of reprehensibility should affect the level of punitive damages. But more specifically, there is no guidance concerning how evidence of concealment/cover-up should affect the degree of reprehensibility, including how much weight concealment/cover-up should be given in relation to the other factors mentioned that bear on the degree of reprehensibility (such as the duration of the conduct). Thus, while *Green Oil* clearly suggests that concealment/cover-up should be a basis for raising the level of punitive damages, the extent to which this should be done is unclear.

[109] For example, consider a firm that believes that any pollution that it generates will be very difficult to detect. Such a firm might not invest in any pollution control equipment -- and thus its conduct would be considered reprehensible. In contrast, an otherwise identical firm that believes that its pollution will be detected with a high probability would make reasonable investments in pollution control equipment. Consequently, if a court does not have direct information about a firm's chance of escaping liability, the court might be able to infer from a firm's level of investment in pollution control equipment whether the firm faced a low or a high chance of escaping liability.

[110] We believe that the notion of socially illicit utility reflects how people often would characterize the utility that individuals derive from certain reprehensible acts (like rape). But because there is no theoretical basis for determining which categories of utility are socially illicit, we are not suggesting below that particular categories of utility are necessarily socially illicit. Which categories of utility from wrongful conduct are socially illicit is an empirical question -- determined by what society does or does not want to count in social welfare. The notion of socially illicit utility has been considered by some commentators in the context of punitive damages. *See* note 105 *supra*.

the injurer's utility from committing the act. Since the injurer's illicit utility could be greater than the harm suffered by the victim, the level of damages needed for proper deterrence might be in excess of harm.[111] In other words, punitive damages might be socially desirable even if there is no chance that the injurer could have escaped liability.

When are the benefits from harmful conduct likely to be considered socially illicit? We suggest that this will tend to be the case when the injurer's utility *derives from causing harm itself*, as when a person punches another out of spite or defames another to see him suffer. The injurer benefits *because* the victim suffers harm. Situations with this characteristic fit under the general rubric of maliciousness and would be considered reprehensible. Thus, *some reprehensible conduct -- malicious conduct -- could give rise to gains that are not counted in social welfare, in which case punitive damages may be justified even in the absence of a chance of escaping liability,* for the reasons discussed in the previous paragraph.

But many acts that are reprehensible do not seem to be associated with socially illicit utility; they are not undertaken with malice. Consider a person who drives at 60 miles per hour through a residential area in order to arrive at work on time and causes a fatal accident. We would call this act reprehensible due to the driver's wanton disregard for the safety of others. Yet because the purpose of the act is not to cause harm, but rather to arrive at work on time, a perfectly legitimate objective, it does not appear that the utility from the act would be classified as socially illicit. In general, we surmise that reprehensible acts that are not undertaken with the objective of causing harm, but rather that happen to cause it as a highly likely byproduct, usually are not associated with socially illicit utility. Thus, for these kinds of acts, there is not a reason to impose punitive damages unless the injurer has a significant chance of escaping liability -- our usual conclusion.[112]

---

[111] For example, suppose an individual obtains a utility gain worth $2,000 to him from maliciously hitting someone in the nose, and that the harm to the victim is equivalent to $500 (say this would fully compensate the victim for his pain and medical costs). Then, even if the injurer would be found liable for sure, punitive damages of at least $1,500 would be required to deter him. (Clearly, if there is a chance that he can escape liability, the punitive damages amount would have to be higher.)

[112] We do not mean to suggest that an injurer's gain necessarily counts in social welfare if his conduct is undertaken without malice. It may well be the case that the utility from certain types of *non-malicious* reprehensible conduct also would be treated as socially illicit. Consider, for example, a person who gets pleasure from "joyriding"

Note that because the goal of corporations is to make a profit, not to cause harm to others, their gains presumably do count in social welfare. Hence, by the foregoing reasoning, if a corporation engages in conduct labeled as reprehensible, this fact *per se* should not affect the level of its damages. Rather, its damages should be based on the harm it caused and the chance that it might have escaped liability, with punitive damages awarded only if the latter chance is significant.

In summary, we believe that the reprehensibility of a defendant's conduct generally should be ignored for the purpose of determining optimal damages for deterrence. The notable exception to this conclusion arises when the defendant is an individual whose conduct is motivated by malice and whose gains consequently are not included in social welfare.

### B. Wealth of Defendants

The courts often state that a defendant's financial condition is a relevant factor in setting a punitive damages award, with the understanding that higher punitive damages may be appropriate for defendants with higher wealth.[113] The defendant's wealth also frequently is included in jury instructions as a factor that is permissible for jurors to take into account in determining the level of punitive damages.[114] Not surprisingly, plaintiffs tend to emphasize

---

on city streets (driving on them at high speed for fun). Our conclusions regarding the appropriateness of punitive damages for malicious conduct would apply to any category of conduct in which an injurer's gain does not count in social welfare.

[113] *See, e.g.*, Green Oil Co. v. Hornsby, 539 So. 2d 218, 222 (Ala. 1989) ("The defendant's financial position is . . . a consideration essential to a post-judgment critique of a punitive damages award.") (citation omitted). The Supreme Court endorsed the *Green Oil* approach of using the defendant's financial position as one factor in determining whether an award of punitive damages is excessive or inadequate in *Pacific Mutual Life Ins. Co. v. Haslip*. 499 U.S. 1, 22 (1991). For other state court decisions endorsing the consideration of wealth, see, for example, Neal v. Farmers Ins. Exch., 582 P.2d 980, 990 (Cal. 1978) ("Also to be considered is the wealth of the particular defendant; obviously, the function of deterrence . . . will not be served if the wealth of the defendant allows him to absorb the award with little or no discomfort."); Ultimate Chem. Co. v. Surface Transp. Int'l, Inc., 658 P.2d 1008, 1012 (Kan. 1983) (among the factors to consider in assessing punitive damages is "defendant's financial condition.").

[114] *See, e.g.*, ARK. MODEL JURY INSTRUCTIONS § 2217 (Supp. 1986); CAL. JURY INSTRUCTIONS: CIVIL, § 14.71 (8th ed. 1995); WIS. JURY INSTRUCTIONS: CIVIL § 1707.1 (1995).

this factor when defendants are wealthy, especially when they are large corporations.[115]

Should defendants with greater wealth pay higher punitive damages? Our main conclusion in this Section is that, from the perspective of achieving proper deterrence, a defendant's wealth generally is inappropriate to consider when the defendant is a corporation. We also conclude that the wealth criterion frequently is inappropriate to consider when the defendant is an individual, although we discuss certain circumstances in which an individual's wealth should be taken into account in imposing punitive damages.[116]

We explained in Section II that if damages equal harm multiplied by a factor reflecting the chance of escaping liability, then defendants, including corporations, would be induced to take optimal precautions and to participate in risky activities to the proper extent. It follows from this basic conclusion that if damages are raised above the magnitude given by our formula for corporations that are relatively wealthy, then those corporations will be led to choose excessive precautions, will undesirably curtail their activities, and will set prices above the proper level, chilling consumption of their products. In the extreme, such corporations might even withdraw their products from the marketplace despite the value of the products to society.

An additional point reinforces the conclusion that corporate wealth should not influence punitive damages: Imposing punitive damages on the basis of corporate wealth effectively

---

[115] For example, in the *Exxon Valdez* case, Exxon's wealth was "virtually the exclusive focus of plaintiffs' Phase III [punitive damages] case." *In re* Exxon Valdez, No. A89-0095-CV, 1995 WL 527988, at *7 (D. Alaska Jan. 27, 1995). The evidence of Exxon's wealth introduced by the plaintiffs included Exxon's 1990 Annual Report, which claimed that "Exxon's consistently strong earnings performance has enabled the company to achieve and maintain a position of extraordinary financial strength and flexibility. For example, over the past ten years, Exxon's internal cash generation from operations amounted to more than $100 billion." *Id.* at *8 n.16.

[116] Our conclusions about punitive damages and wealth are similar to those of other economically-oriented writers on punitive damages. *See, e.g.*, Abraham & Jeffries, *supra* note 3, at 415 ("In our view, the defendant's wealth is irrelevant to the goal of deterring socially undesirable conduct ..."); Chapman & Trebilcock, *supra* note 12, at 824 ("In the case of economic wrongs, the conventional economic theory of deterrence . . . suggests no role for corporate wealth in structuring an optimal deterrence regime . . . ."); Cooter, *Deterrence*, *supra* note 12, at 1176-77 (1989) (the total assets or wealth of the defendant "is typically inappropriate to deterrence of economically self-interested decisionmakers" (footnote omitted)); Galligan, *supra* note 3, at 65 ("Considering the defendant's wealth has simply no articulable efficiency justification."); Gary T. Schwartz, *Deterrence and Punishment in the Common Law of Punitive Damages: A Comment*, 56 S. CAL. L. REV. 133, 140 (1982) ("The 'wealth of the defendant' bears no obvious relationship to deterrence goals . . . .").

imposes a tax on corporate size and success, thereby discouraging growth and development. This effect can be important in industries in which liability costs are a significant component of total cost (such as in the pharmaceutical and general aviation aircraft industries).[117] Of course, retarding the natural growth of corporations can have adverse consequences, notably, that society forgoes economies of scale in production and in research and development. It also may mean that the risk of harm increases, because small firms may not have enough at stake to make it worthwhile to them to spend a socially proper amount on precautions.

Our discussion of the inappropriateness of taking corporate wealth into account presumes that all corporations -- large and small -- will, if required to pay for the harms they cause, tend to balance correctly the costs of precautions against the resulting reduction in harm. An argument sometimes is made, however, that because bigger corporations are more bureaucratic, they will not adequately respond to liability risks unless the damages imposed on them are especially high. According to this argument, higher damages are needed against large corporations to attract the attention of senior management.[118] This view is mistaken, as we now discuss.

Although it is true that large corporations typically have complicated organizational structures, with senior management at some remove from the level of operations, it does not follow that large corporations will tend to be inappropriately attentive to the reduction of risk. If the cost of a precaution is less than the damages incurred by not taking it, a large firm will want someone employed by it to recognize that circumstance and take the precaution -- because the firm's goal is to maximize profits. A large grocery chain, for example, will want some employee at each of its stores to inspect that store's floor after being mopped to ensure that it is safe. The company will delegate this responsibility to an employee

---

[117] *See* note 29 *supra*.

[118] Sometimes this is expressed as a need to "send a message to headquarters." *See, e.g.*, Browning-Ferris Indus., Inc. v. Kelco Disposal, Inc., 492 U.S. 257, 261 (1989) (Kelco's attorney "urged the jury to return an award of punitive damages, asking the jurors to 'deliver a message to Houston [BFI's headquarters]'") (quoting trial transcript). The converse view that it is not necessary to impose higher damages when corporations are not large is reflected in Green Oil Co. v. Hornsby, 539 So. 2d 218, 221 (Ala. 1989) (The trial court reduced the jury's award of $150,000 in punitive damages to $25,000, noting that "it doesn't take a large verdict to be heard and felt a few miles down the road . . . by two local individuals.").

low in the corporate hierarchy, such as an assistant store manager. That this task does not receive the attention of top management, as it might in the case of a firm consisting of only one or two grocery stores, does not mean that the task will be neglected or attended to inadequately. As long as a corporation -- large or small -- expects to have to pay for the harms it causes, it will have a socially appropriate incentive to reduce the harms.[119]

Now consider the question of the relevance of wealth for the imposition of punitive damages on individuals. Again, the general arguments we made in Section II imply that punitive damages should not depend on an individual's wealth; rather, punitive damages should depend only on the level of harm and the chance of escaping liability, so that, applying the damage multiplier formula, expected damages equals harm. However, there are two qualifications to this conclusion that imply that wealth might be relevant in certain circumstances.

The first concerns risk aversion and the unavailability of insurance against punitive damages. We noted in Section II that if potential injurers are risk averse and do not have access to liability insurance, then appropriate deterrence will be accomplished with a lower level of damages than if they are risk neutral. Further, the more risk averse an individual is, the lower the optimal level of damages. Assuming that poor individuals are more risk averse than rich individuals,[120] this implies that the optimal level of punitive damages is lower for poorer individuals. Equivalently, punitive damages should be higher for wealthier individuals. However, even for the wealthiest individuals, punitive damages should not exceed the level determined by our formula.[121] The relevance of these observations, we reiterate, is limited to

---

[119] Although we have just emphasized the point that large corporations will take appropriate steps to reduce risk through the delegation of risk-reduction responsibilities, we are not claiming that large corporations will necessarily take the *same* precautions that small corporations do. They may take different precautions (perhaps greater, perhaps not) as a result of their different organizational and decisionmaking structures. But the precautions taken will still be socially appropriate because of the basic principle that parties will behave socially responsibly if they are made to pay for the harms their actions cause.

[120] This means, for example, that a poor person would be more averse to a 50% chance of losing $100 than a rich person.

[121] This is because, as has been explained, the bearing of risk by uninsured risk-averse individuals makes it socially desirable to *reduce* damages from the level implied by our formula. Thus, while damages should rise with wealth for the reasons just discussed, the highest level of damages -- imposed on the wealthiest individuals, who are

situations in which insurance against punitive damages is not available.

The second circumstance in which the level of an individual's wealth may be relevant to the calculation of punitive damages is if the individual's gain from committing the harmful act is socially illicit, a possibility we discussed in Section III.A. We explained there that punitive damages may be needed to offset the illicit benefits. To accomplish this, punitive damages generally will have to be higher the higher the wealth of an individual, because the value of money tends to decline with wealth.[122] For example, to offset the utility a rich person would obtain from slandering someone he disliked, we might need to impose $10,000 in punitive damages, whereas to deter a person with only modest assets, $1,000 in punitive damages might suffice.

We believe that the point of the preceding paragraph underlies the common intuition that punitive damages should be linked to wealth. However, it is important to recognize the very limited scope of this point. It applies only to individuals whose benefit from causing harm is socially illicit, which we generally associate with conduct whose goal is to cause harm. Otherwise, the conclusion of the previous paragraph does not apply to individuals. Moreover, the conclusion does not carry over to firms because firms are motivated by profits; their objective is not to cause harm.

### C. Potential Harm

In reviewing the appropriateness of a punitive damages award, some courts have considered not only the harm that actually occurred, but also the harm that might have occurred -- the potential harm.[123] According to these courts, the higher the potential harm, the

---

presumed to be the least risk averse -- should still not exceed the level called for by our formula.

[122] That the value of a dollar declines with the level of wealth is a standard assumption of economists, reflecting the view that individuals first fulfill their most important needs and desires, and then spend on successively less important things. *See, e.g.*, PINDYCK & RUBINFELD, *supra* note 39, at 144-45.

[123] *See, e.g.*, Bemer Aviation, Inc. v. Hughes Helicopter, Inc., 621 F. Supp. 290, 300 (E.D. Pa. 1985), *aff'd*, 802 F.2d 445 (3d Cir. 1986) (In assessing an award of punitive damages, juries may consider "the potential harm that a defendant's conduct poses."); *Green Oil*, 539 So. 2d at 223 (*see infra* note 124); Levine v. Knowles, 197 So. 2d 329, 331 (Fla. Dist. Ct. App. 1967) ("The seriousness of the probable result of the defendant's conduct . . . is the yardstick for determining the advisability of discouraging such behavior in the future, rather than the seriousness

higher the level of punitive damages that can be justified.[124] This idea was endorsed by the U.S. Supreme Court in the *Haslip* and *TXO* opinions.[125] It also served as a basis for the upholding by the trial court of the $5 billion punitive damages verdict in the *Exxon Valdez* oil spill litigation; the court noted that whereas 11 million gallons of oil spilled, there were another 45 million gallons in the *Exxon Valdez* that could have spilled, making the potential harm much higher.[126]

We conclude, however, that a policy of taking potential harm into account in the determination of punitive damages is undesirable given the goal of deterrence. To explain the reasoning behind this conclusion, it will be expositionally convenient first to discuss why damages should be based on actual rather than potential harm when there is no chance of escaping liability and the issue of punitive damages does not arise.

Consider an example in which an injurer's act will result in either a low or a high level of harm and the injurer does not know in advance which level will occur. Let the two levels of harm be $1 million and $5 million, which occur with equal probability. This example raises the issue of potential harm since, when a $1 million harm occurs, the harm could have been $5 million. Note that the injurer's act entails an expected harm of $3 million (= (.5 x $1 million) + (.5 x $5 million)). To achieve proper deterrence, therefore, the injurer's expected damages should equal $3 million.

---

of the damage actually caused.") (footnote omitted).

[124] The first of the seven *Green Oil* factors for evaluating punitive damages awards states:

> (1) Punitive damages should bear a reasonable relationship to the harm that is likely to occur from the defendant's conduct as well as to the harm that actually has occurred. If the actual or likely harm is slight, the damages should be relatively small. If grievous, the damages should be much greater.

*Green Oil*, 539 So. 2d at 223 (quoting Aetna Life Ins. Co. v. Lavoie, 505 So. 2d 1050, 1062 (Ala. 1987) (Houston, J., concurring)).

[125] The *Green Oil* factors are endorsed in *Pacific Mutual Life Insurance Co. v. Haslip*, 499 U.S. 1, 21-22 (1991); *see also* TXO Prod. Corp v. Alliance Resources Corp., 509 U.S. 443, 460 (1993) ("It is appropriate to consider the magnitude of the potential harm that the defendant's conduct would have caused . . .").

[126] "The evidence established that the *Exxon Valdez* spilled 11,000,000 gallons of crude oil, approximately one-fifth of its cargo. Had the remaining 45,000,000 gallons of oil spilled, the disaster and harm would have been many times greater." *In re* Exxon Valdez, No. A89-0095-CV, 1995 WL 527988, at *6 (D. Alaska Jan. 27, 1995) (Holland, J.).

Given our assumption that injurers will be found liable when they cause harm, observe that *deterrence will be optimal if damages are always set equal to actual harm.* For when an injurer engages in the harmful activity, he will expect to have to pay $1 million in damages half of the time and $5 million in damages half of the time. Hence, his average damage payment will be $3 million or, stated differently, the expected value of his damage payment is $3 million (= (.5 x $1 million) + (.5 x $5 million)). This is, as noted in the previous paragraph, the amount needed for proper deterrence.

Nevertheless, if actual harm turns out to be low, one might wonder why basing damages on actual harm does not result in inadequate deterrence. In our example, if the harm is $1 million and the injurer is made to pay this amount, he will be paying relatively little compared to the $5 million harm that his act might have caused. (The difference between actual harm and potential harm could be much greater -- indeed, a person may act very dangerously but cause *no* harm and pay *no* damages if damages are based on actual harm.) The reason why this is not a problem, however, should be apparent: when a potential injurer chooses whether to engage in a harmful act, he *does not know* what the harm -- and therefore what his damages -- will be. The injurer in our example cannot predict whether his damages will be $1 million or $5 million. Consequently, he will decide whether to commit the harmful act on the basis of having to pay the average or expected damage amount, which is $3 million. It would be a mistake, therefore, to think that he will be inadequately deterred if the actual harm turns out in a particular case to be $1 million and he only has to pay this amount. The possibility that the harm and his damages might have been $5 million also will influence his behavior, in an appropriate way.

Now suppose that instead of basing damages solely on actual harm, courts take potential harm into account by raising damages when the actual harm is unusually low -- because it could have been much higher -- but do not lower damages when the actual harm is high. (This is a plausible interpretation of how the potential harm factor is used in practice, at least in the context of punitive damages.[127]) The point we want to emphasize is

---

[127] For example, Judge Holland observed that more oil might have spilled from the *Exxon Valdez*, *see* note 126 *supra*, but he did not mention the possibility that less might have spilled.

that such a policy imparts a systematic upward bias to the level of damages and results in injurers bearing damages in excess of actual harm. Suppose in the example that damages are raised to $3 million when the actual harm is $1 million -- on the ground that the former amount is the average harm -- but that damages are not lowered when the actual harm is $5 million. Then the injurer will pay on average $4 million ($3 million half of the time and $5 million half of the time), even though the average harm is $3 million. Making injurers pay damages in excess of harm will have the undesirable consequences associated with overdeterrence that we have discussed previously.

There is, however, another way in which potential harm could be taken into account that would not cause damages to systematically exceed harm. Specifically, suppose that damages are set equal to the average, or expected, harm *regardless* of whether the actual harm is below or above this amount. Thus, in the example, damages would be set equal to $3 million regardless of whether the actual harm is $1 million or $5 million. In effect, under this policy recognition is given both to the fact that when actual harm is low, it could have been higher, and to the fact that when actual harm is high, it could have been lower. Such a policy would result in proper deterrence because, on average, the injurer will be paying for the harm he causes: each time he commits the harmful act he pays $3 million and causes, on average, harm of $3 million. Whether this policy would be employed in practice is questionable, however. One might be skeptical that, when the actual harm is high, juries and courts would reduce damages because the harm might have been lower.[128]

Even if a policy of basing damages on average or expected harm were applied consistently, there is a strong argument on administrative grounds in favor of relying solely on actual harm. For the courts to be able to calculate *expected* harm, they would have to determine each level of harm that could have occurred and its probability of occurrence. This

---

[128] To illustrate, imagine the response to the Union Carbide Company if, in the Bhopal disaster in which 4,000 people were killed and thousands of others were injured, *see* Kenneth J. Cooper, *Slums Sprawl in Shadow of Bhopal Gas Leak*, WASH. POST, June 27, 1996, at A19, Union Carbide had argued that damages should be reduced because a gas leak of the kind that occurred ordinarily would be expected to kill and injure a much smaller number of people. *See also* note 127 *supra*.

ordinarily would be far more difficult to establish than the harm that actually did occur.[129] Additionally, the more open-ended scope of an inquiry into expected harm seems likely to lead to more disputes between the parties, for it is easier to disagree about what might have happened, and the odds of it happening, than about what actually did happen. These observations suggest that the courts will bear greater administrative expense, and the parties will bear greater litigation costs, if the goal is to calculate expected harm rather than to ascertain just the actual harm in the instant case.[130]

To summarize, there are two reasons why potential harm generally should not be taken into account in determining damages. First, we envision that the potential harm factor would be used in practice to raise damages when harm is low but not to lower damages when harm is high, thus causing overdeterrence. Second, even if potential harm were considered in the theoretically correct way -- by always setting damages equal to average or expected harm -- such a policy would require an inquiry into what might have occurred and is therefore likely to increase the public and private costs of resolving legal disputes.[131]

As noted at the beginning of this Section, we assumed for expositional simplicity that the injurer is found liable for sure. The arguments for basing damages on actual harm rather than on potential harm or expected harm are essentially the same if the injurer might escape

---

[129] In some cases, however, it might be easier to calculate expected or average harm than the particular harm in the case at hand. For example, it might be easier to determine the average value of the contents of a house that burned down than to ascertain the actual value of the contents (since much of the contents may have been consumed in the fire).

[130] This conclusion also would apply if courts attempt to take potential harm into account in some other, less sophisticated, way than by calculating the expected harm.

[131] We have discussed what we believe to be the main arguments bearing on the desirability of taking potential harm into account in calculating damages, but there are other considerations, some of which reinforce our conclusions and some of which do not. Among these additional points, we note two. First, supporting our conclusion, is the point that if damages are based on potential harm rather than actual harm, then the incentive to limit actual harm will be dulled. For example, an oil company would have less of an incentive to curtail the leakage of oil from a grounded tanker if damages are based on the total amount of oil in the vessel rather than on the amount that actually leaks. The second point, which favors basing punitive damages on potential harm, is that such a policy can reduce the dilution of incentives due to the judgment proof problem. Specifically, because consideration of potential harm in determining damages means that parties will pay larger amounts than otherwise when actual harm is low, their not being able to pay higher damages when actual harm is high (due to the judgment proof problem) will tend to be counteracted.

liability. The discussion then would be framed in terms of a comparison between actual harm, appropriately inflated to make up for the chance of escaping liability, and potential harm or expected harm, also so inflated. Our point about the courts' tendency to apply the potential harm factor in a way that leads to overdeterrence still holds, as does the point about the greater administrative complexity of determining expected harm. Thus, when the injurer has a significant chance of escaping liability and punitive damages therefore are needed to achieve proper deterrence, such damages generally should be an appropriate multiple of actual harm, not of expected harm or potential harm. The potential harm factor in punitive damages law generally should be ignored.

### D. Gain of Defendants

When punitive damages are imposed, their level sometimes is influenced by application of the principle that the defendant should not gain from his wrongful conduct.[132] If setting damages equal to harm would not remove the defendant's gain, then the argument is that damages should include a sufficient punitive component to offset his gain. The notion that the defendant's gain should be a factor in calculating punitive damages was endorsed by the U.S. Supreme Court in the *Haslip* case.[133]

Does it make sense in terms of deterrence to ensure that the defendant's gain is disgorged, or should damages be based solely on harm (abstracting from the issue of the

---

[132] The third *Green Oil* factor states:

> (3) If the wrongful conduct was profitable to the defendant, the punitive damages should remove the profit and should be in excess of the profit, so that the defendant recognizes a loss.

*Green Oil*, 539 So. 2d 218, 223 (Ala. 1989) (quoting Aetna Life Ins. Co. v. Lavoie, 505 So. 2d 1050, 1062 (Ala. 1987) (Houston, J., concurring)); *see also* Estate of Hartz v. Nelson, 437 N.W.2d 749, 755-56 (Minn. Ct. App. 1989) (noting under MINN. STAT. § 540.20 punitive damages should be measured, in part, by "the profitability of misconduct to the defendant"); Tindall v. Konitz Contracting, Inc. 783 P.2d 1376, 1382-83 (Mont. 1989) (noting that under MONT. CODE ANN. § 27-1-221(7)(b) the awarder of punitive damages must consider "the profitability of the defendant's wrongdoing, if applicable").

[133] Pacific Mutual Life Ins. Co. v. Haslip, 499 U.S. 1, 21-22 (1991) (favorably citing the third *Green Oil* factor).

chance of escaping liability[134])? We conclude in this Section that setting damages equal to harm generally results in proper deterrence even when the harm is less than the defendant's gain; a policy of removing the defendant's gain may result in overdeterrence. An exception arises, however, when the defendant's gain is socially illicit, in which case extracting the defendant's gain is desirable.[135]

The question whether punitive damages should be imposed in order to remove the defendant's gain arises only when his gain exceeds the victim's harm (otherwise, compensatory damages would eliminate the gain). One reason gain could exceed harm is that the level of harm is uncertain and, by chance, turns out to be low -- and less than the injurer's gain.[136] This could occur even though the *expected harm* exceeds the gain. For example, suppose that a firm would save $100,000, and thus would gain that amount, by not purchasing a safety device; that the expected harm from failing to purchase the device is $1 million; but that only a $10,000 harm occurs. Here the firm's $100,000 gain exceeds the unusually low harm of $10,000.

Using this example, we first want to show that basing damages on harm *will* accomplish proper deterrence even though the defendant's gain exceeds the harm: At the time the firm decides whether to buy the $100,000 safety device, it does not know what the harm

---

[134] The arguments that we make in this section do not depend on whether a defendant might escape liability. See note 140 *infra*. Thus, for expositional convenience, we assume here that defendants never escape liability.

[135] Although there has been some scholarly discussion regarding whether to base liability on harm or gain, the points developed in this literature differ from the points that we present here. Our focus, as the reader will see, is on a measure of damages equal to the greater of gain or harm. Previous literature studies the measure of damages equal to the gain and compares it to the measure of damages equal to the harm. *See, e.g.*, Richard S. Gruner, *Just Punishment & Adequate Deterrence for Organizational Misconduct: Scaling Economic Penalties Under the New Corporate Sentencing Guidelines*, 66 S. CAL. L. REV. 225 (1992); A. Mitchell Polinsky & Steven Shavell, *Should Liability Be Based on the Harm to the Victim or the Gain to the Injurer?*, 10 J. LAW ECON. & ORG. 427 (1994); Jeffrey S. Parker, *Criminal Sentencing Policy for Organizations: The Unifying Approach of Optimal Penalties*, 26 AM. CRIM. L. REV. 513 (1989); Donald Wittman, *Liability for Harm or Restitution for Benefit?*, 13 J. LEGAL STUD. 57 (1984); Donald Wittman, *Should Compensation be Based on Costs or Benefits?*, 5 INT'L REV. LAW & ECON. 173 (1985).

[136] If both harm and gain were certain, then it is unlikely that the gain would exceed the harm in a punitive damages case: If the gain is known to exceed the harm, the act would be likely to be regarded as socially desirable, or at least not one calling for imposition of punitive damages. Hence, we consider the possibility here that harm is uncertain, and below that gain is uncertain (see text at note 137 *infra*).

will be. If damages always are set equal to harm and the expected harm is $1 million, then the firm's *expected* damages will be $1 million, which of course will induce it to spend $100,000 on the safety device. It is not necessary to impose punitive damages just because the harm turns out to be unusually low, and below the injurer's gain. (Note that the point of this paragraph is analogous to the point we made about potential harm in Section III.C.)

If the standard policy of imposing damages equal to harm appropriately deters, is there a disadvantage of imposing higher damages -- namely, to remove the defendant's gain -- when gain turns out to exceed harm? The answer is in the affirmative, because overdeterrence may result. Specifically, if damages are set equal to harm when the harm exceeds the injurer's gain, but damages are set equal to gain when gain exceeds harm, then expected damages will exceed the expected harm, resulting in the usual problems of excessive liability.

If damages are set so as to remove gains, an additional reason that overdeterrence may result is that the basis for measuring the injurer's gains might be interpreted too expansively. In the example above, the firm's gains might be construed to be its profits from the entire line of activity that gave rise to the accident (say the profits from manufacturing automobiles at a particular plant), rather than just the saving from not taking the particular precaution (say not purchasing a $100,000 instrument to test the integrity of the automobiles' brakes). If gains are erroneously measured in this way, a policy of setting damages equal to gain will be even more likely to result in excessive liability.

Another circumstance in which a defendant's gain might exceed the victim's harm is when the level of gain is uncertain[137] and, by happenstance, turns out to be high even though the expected gain is low and less than harm.[138] For example, suppose that the cost of re-engineering an assembly line to make production safer ordinarily is $200,000, but unforeseen complications could raise the cost to $800,000. If the assembly line is modified, harm of $500,000 will be avoided. Thus, while the expected gain from foregoing the safety

---

[137] For simplicity, we assume now that harm is certain.

[138] If the expected gain exceeded the harm, imposition of punitive damages would be unlikely.

improvements is less than the $500,000 harm,[139] the actual gain could be $800,000 and greater than the harm. The analysis of this situation is similar to that when the uncertainty concerned harm; again it can be shown that setting damages equal to harm will create appropriate deterrence and that imposing punitive damages to remove injurers' gains will tend to result in overdeterrence.

Finally, consider the possibility that the injurer's gain is socially illicit, as when a person acts out of malice. We noted in Section III.A that it is desirable to deter an injurer whose gain is illicit even if his gain exceeds the victim's harm, and that this implies that punitive damages may be needed to offset the injurer's gain. Thus, in the case of socially illicit utility, the notion of using punitive damages to ensure that the defendant's gain is removed is justifiable with respect to the goal of deterrence. (As we also observed previously, this justification for removing the defendant's gain does not apply to corporations or to individuals whose actions are non-malicious.)

In sum, then, our conclusion is that removing the defendant's gain is potentially appropriate and necessary only when the defendant is an individual who acted maliciously and obtained a socially illicit gain. Otherwise, including when the defendant is a firm, the usual policy of setting damages equal to harm is desirable in terms of deterrence, and imposing damages so as to remove gains will tend to cause overdeterrence.[140]

### E. Litigation Costs

Several courts have suggested that the plaintiff's litigation costs should be a factor in

---

[139] This statement will be true if the $200,000 cost of re-engineering the assembly line is more likely than the $800,000 cost, which is what we mean by saying that the cost ordinarily is the lower amount. For then the expected cost will be less than $500,000, which is equivalent to saying that the expected gain from not re-engineering the assembly line is less than $500,000.

[140] We suggested in note 134 *supra* that the arguments in this section apply without substantial modification to situations when defendants might escape liability. To illustrate, consider the initial point that imposing damages equal to harm, whatever its magnitude, will cause expected damages to equal expected harm. The analogue of this point when defendants can escape liability is that imposing damages according to our multiplier formula will result in expected damages equal to expected harm, essentially for the reasons given in the text.

the determination of punitive damages,[141] and, as noted previously, some have stated that such costs should be included as a component of punitive damages so as to encourage victims to sue injurers.[142]

Should litigation costs bear on the calculation of punitive damages in order to achieve proper deterrence? Our answer in this Section emphasizes two points. The first is that litigation costs may cause the probability of suit to be low and thus justify a punitive damages award according to the damage formula presented in Section II.B. The second point is that punitive damages generally should not be augmented for the purpose of inducing suits that otherwise might not be brought due to the cost of litigation. Raising the probability of suit usually is unnecessary to achieve proper deterrence, and encouraging suits has the disadvantage of increasing the litigation costs borne by society. Indeed, we argue that a policy adopted in many states of *decoupling* punitive damages -- giving the plaintiff only a fraction of the punitive damages judgment paid by the defendant, with the remainder going to the state -- may be desirable because it can reduce the volume of litigation without compromising deterrence.[143]

The first point, that litigation costs may be relevant to the calculation of punitive damages because they influence the probability of suit, and therefore the chance of escaping liability, is one that we have made previously. We observed that if litigation costs are significant relative to the expected gain from suit, the probability of suit may be small, and

---

[141] *See, e.g.*, Ultimate Chem. Co. v. Surface Transp. Int'l, Inc., 658 P.2d 1008, 1012 (Kan. 1983) ("A jury may also consider . . . the probable litigation expenses."); Fischer v. Johns-Manville Corp., 512 A.2d 466, 482 (N.J. 1986) (Among the factors to be considered in determining the size of a punitive damages award is "the plaintiff's litigation expenses.").

[142] *See* note 72 *supra.*

[143] A number of articles on the economics of litigation are relevant to the conclusions that we reach in this Section. On the general topic of how to structure the legal system given that litigation is costly, see A. Mitchell Polinsky & Yeon-Koo Che, *Decoupling Liability: Optimal Incentives for Care and Litigation*, 22 RAND J. ECON. 562 (1991); A. Mitchell Polinsky & Daniel L. Rubinfeld, *The Welfare Implications of Costly Litigation for the Level of Liability*, 17 J. LEGAL STUD. 151 (1988); Steven Shavell, *The Fundamental Divergence between the Private and the Social Motive to Use the Legal System*, 26 J. LEGAL STUD. 575 (1997) [hereinafter Shavell, *Divergence*]; Steven Shavell, *The Social versus the Private Incentive to Bring Suit in a Costly Legal System*, 11 J. LEGAL STUD. 333 (1982). *See also* Friedman, *supra* note 12 (suggesting that punitive damages may beneficially lower litigation costs by discouraging harmful behavior); Kahan & Tuckman, *supra* note 12 (specifically addressing decoupling punitive damages).

this fact may justify imposing punitive damages on the injurer. For example, we suggested that in the circumstances of the *Gore* case, litigation costs may have led to a low likelihood of suit because the harm to the plaintiff was found to be only $4,000. However, in other circumstances, like those in the *Exxon Valdez* case, litigation costs are likely to be insignificant in relation to the expected gain from suit, so that the probability of suit may be presumed to be very high. Then, consideration of litigation costs does not provide a basis for imposing punitive damages.

Note that when punitive damages are justified because of litigation costs, such costs should not be simply added to punitive damages, as some courts have suggested doing. Proper punitive damages are determined by the multiplier formula, which calls for a level of punitive damages that generally differs from litigation costs. For instance, suppose harm is $10,000 and the plaintiff's litigation costs are $5,000 and lead to a 25 percent likelihood of suit. Then total damages should be four times the harm, or $40,000, and punitive damages should be $30,000, not the $5,000 amount corresponding to litigation costs.

Let us now turn to our second point. Because the punitive damage formula is *designed* to achieve appropriate deterrence when suit does not always occur, it is not necessary to raise punitive damages awards for the specific purpose of raising the probability of suit (provided that at least some suits are brought).[144] If suit occurs only half of the time because of the discouraging effect of litigation costs, then total damages according to our formula would be twice the harm, and deterrence will be appropriate; there is no need to increase punitive damages to make suit occur more frequently. Moreover, encouraging lawsuits would increase social costs. Obviously, the greater the number of suits, the higher are the legal costs borne

---

[144] If no suits are brought, then deterrence obviously cannot be achieved. But if there is a positive probability of suit, deterrence will be optimal if damages are set according to our formula. There is a reason, however, that the probability of suit not be too low. According to our formula, punitive damages are higher the lower is the probability of suit. Consequently, if the probability of suit is relatively low, injurers may not have assets sufficient to pay the proper punitive damages amount. In other words, the assets of injurers impose an implicit constraint on how low the probability of suit can be without compromising deterrence. Assuming this minimum likelihood of suit is attained, our point is that there is no reason to elevate it by raising punitive damages.

by the parties and the administrative costs borne by the state.[145] Raising damages to induce suits also will cause parties to spend more litigating each suit.[146] Thus, awarding punitive damages to spur suit is socially undesirable, other things being equal.

The tendency of higher damage awards to increase litigation costs lends appeal to the policy of decoupling punitive damages. As noted, this means awarding the plaintiff only a part of the punitive damages judgment paid by the defendant, with the remainder going to the state.[147] Use of decoupling allows society to mitigate the effect that punitive damages awards encourage unnecessary litigation, but decoupling does not dilute deterrence because defendants' damage payments are unaffected.[148]

Several states have adopted statutes that decouple punitive damages.[149] For example, in Iowa 25 percent of the punitive damages amount paid by the defendant is given to the plaintiff, and in Florida 65 percent is given to the plaintiff.[150] For the reasons provided in the previous paragraph, there is much to recommend decoupling schemes of this sort.

In summary, the main justification for considering litigation costs is in connection with

---

[145] Indeed, because litigation is costly, the full social harm due to an accident is the direct harm *plus* the costs associated with use of the legal system. Hence, for the injurer to have correct incentives, he should, in principle, pay damages equal to the direct harm plus these additional costs. *See* Shavell, *Divergence*, *supra* note 143. As noted previously, *see* note 18 *supra*, we have ignored this refinement in the text.

[146] This intuitively plausible proposition has been confirmed in research undertaken by scholars at the Institute for Civil Justice at the RAND Corporation. *See* JAMES S. KAKALIK ET AL., VARIATION IN ASBESTOS LITIGATION COMPENSATION AND EXPENSES (Institute for Civil Justice, RAND Corp., No. R-3132-ICJ, 1984).

[147] When we use the term "decoupling," we presume that the defendant pays more than the plaintiff receives, even though, as a logical matter, the plaintiff could be awarded more than the defendant pays.

[148] Because some of the damages paid by defendants go to the state, plaintiffs' incentives to sue will lessen. However, the punitive damages amount determined by our formula will automatically rise to reflect any decrease in the probability of suit. Thus, the expected damages borne by defendants will not decline if our formula for punitive damages is applied.

[149] *See generally* BMW of North Am., Inc. v. Gore, 116 S. Ct. 1589 app. at 1619 (1996) (appendix of Ginsburg, J., dissenting) (listing state provisions that allocate a portion of punitive damages awards to state agencies). Although the statutes that describe these allocation arrangements do not use the term "decoupling," this term is employed in some of the economic literature analyzing litigation. *See, e.g.*, Polinsky & Che, *supra* note 143.

[150] In Iowa, the plaintiff receives 25% of the punitive damages award under certain circumstances. *See* IOWA CODE ANN. § 668A.1(2)b) (West, WESTLAW through 1995 Reg. Sess.). In Florida, the plaintiff receives 65% of the punitive damages award. *See* FLA. STAT. ANN. § 768.73(2) (West, WESTLAW through 1996 2d Reg. Sess.).

estimating the chance that a defendant might have escaped liability because he would not be sued. Punitive damages should be awarded to make up for the chance of escaping liability for this reason, but not as a general matter to encourage the bringing of lawsuits. Decoupling punitive damages may allow proper deterrence to be achieved without inducing needless litigation.

### F. Related Private Litigation

A defendant sometimes may be the subject of multiple suits because he engages in the same type of harmful conduct repeatedly, or because he commits a single harmful act that injures many individuals. The circumstances of the *Gore* case exemplify the former possibility: a car manufacturer that engages in the practice of repainting damaged cars and selling them as new may be sued by different purchasers of these cars. A case involving, say, the dumping of toxic waste that infiltrates an aquifer illustrates the latter possibility: the dumping, a single act, may give rise to suits by many different parties who have been harmed. When there have been prior judgments against a defendant for the same conduct, the U.S. Supreme Court has endorsed the notion that these judgments should be taken into account in mitigation of a punitive damages award against the defendant.[151]

In this Section we discuss the application of our punitive damages formula when multiple suits may be brought, and we observe that the formula's implications generally comport with the view that punitive damages should be lowered in light of other private judgments against a defendant. We also note that multiple punitive damages claims against a

---

[151] The seventh *Green Oil* factor states, "If there have been other civil actions against the same defendant, based on the same conduct, this should be taken into account in mitigation of the punitive damages award." Green Oil Co. v. Hornsby, 539 So. 2d 218, 224 (Ala. 1989) (quoting Aetna Life Ins. Co. v. Lavoie, 505 So. 2d 1050, 1062 (Ala. 1987) (Houston, J., concurring), and endorsed in Pacific Mutual Life Ins. Co. v. Haslip, 499 U.S. 1, 22 (1991)). This factor is echoed in the Restatement:

> Another factor that may affect the amount of punitive damages is the existence of multiple claims by numerous persons affected by the wrongdoer's conduct. It seems appropriate to take into consideration both the punitive damages that have been awarded in prior suits and those that may be granted in the future, with greater weight being given to the prior awards.

RESTATEMENT (SECOND) OF TORTS § 908 cmt. e (1979). However, when the harm that originates from the defendant's conduct is repetitive, as in *Gore*, it is not clear whether prior judgments against the defendant would be taken into account. *See* BMW of N. Am., Inc. v. Gore, 116 S. Ct. 1589, 1607 (1996) (Breyer, J., concurring).

defendant for the same or related conduct may result in his paying for more than the harm he caused, and we discuss a mechanism -- a punitive damages escrow account -- that can be used to reduce this risk.

Let us first consider the proper level of punitive damages when multiple suits may be brought because of repeated harmful conduct, as in *Gore*. Whether there have been prior suits may be relevant to assessing the probability of suit, and thus to the determination of the punitive damages multiplier. If few (or no) suits have been brought even though harm had occurred in the past, that fact suggests that the likelihood of suit is low, implying that the multiplier should be high. Conversely, if a large number of prior suits have occurred, the usual inference would be that the likelihood of suit is significant, and therefore that the multiplier should be low. Note that these points mean that punitive damages in a particular case should be mitigated on the basis of the number of prior suits.[152]

That few (or no) suits have been brought prior to the instant case does not necessarily mean, however, that there will be a paucity of litigation in the future. For example, publicity about the current suit may attract future suits, or an award of punitive damages in the current suit may stimulate litigation.

Because of the difficulty of predicting the amount of future litigation, courts might mistakenly believe that relatively few suits will be brought, and therefore perceive a greater need for punitive damages than is appropriate. (While courts also could incorrectly expect that many suits will occur in the future, this error does not give rise to the problem we are about to discuss.[153]) If such a mistake occurs, a defendant may be made to pay more than the harm that he caused. For example, in circumstances like those in *Gore*, suppose that the court in which the first case is filed believes that only ten percent of similarly harmed car purchasers will sue

---

[152] It is not just the number of prior suits that matters to the punitive damages multiplier in a case. The magnitude of the awards in prior suits matters as well. The higher the prior awards, everything else being equal, the lower should be punitive damages in the case in question, since the goal of deterrence is to make the injurer's payments equal to the total harm.

[153] If courts overestimate the likelihood of suit in the future, punitive damages will be lower than they should be in the case at hand. But this problem can be corrected by raising punitive damages in future cases. *See also* note 156 *infra*.

in the future. The punitive damages formula then would imply that the court should impose total damages on the manufacturer equal to ten times the current plaintiff's harm. If, however, the truth is that much more than ten percent of the other victims actually will bring suit, the car manufacturer may ultimately pay for more than the harm that it has caused, because of the excessive initial award of punitive damages.

A way to avoid the problem of excessive damages when there are multiple suits is to use a *punitive damages escrow account*. Under this approach, punitive damages would be paid into an escrow account rather than immediately to the plaintiff.[154] If, over time, more suits are brought than were anticipated, the damage awards to the plaintiffs can be financed from the escrow account rather than charged to the defendant. In this way, the defendant will not be made to pay more in total than the harm done. If, at some natural termination date,[155] funds remain in the escrow account, they can be distributed to plaintiffs whose punitive damages awards had been placed in escrow.[156]

Finally, let us turn briefly to the situation in which multiple suits arise because the defendant has committed a single harmful act that injured many individuals (as in the example involving the dumping of toxic waste). Here, our points are analogous to those discussed above. Whether there have been prior suits again may be relevant to evaluating the probability of suit: a greater number of prior suits should raise the estimated likelihood of suit, reduce the punitive damages multiplier, and thereby lead a court to impose lower punitive damages. Similarly, a punitive damages escrow account can be used to avoid imposing excessive damages on the defendant.

---

[154] *See* Margaret I. Lyle, Note, *Mass Tort Claims and the Corporate Tortfeasor: Bankruptcy Reorganization and Legislative Compensation Versus the Common-Law Tort System*, 61 TEX. L. REV. 1297, 1348 (1983) (suggesting that a court might "order[] an equitable stay on collection of a punitive damage award for a number of years if it seems likely that the collection of too many of these awards early in the litigation of a mass tort might deprive later plaintiffs of compensatory damages").

[155] For example, a termination date might be the expiration of the statute of limitations period for the bringing of suits.

[156] For example, a termination date might be the expiration of the statute of limitations period for the bringing of suits.

### G. Related Public Penalties

Another question of interest is whether public penalties that may be imposed for the type of wrongful conduct at issue in a private suit should affect the determination of punitive damages in that suit. Courts have answered this question in two ways. First, some have stated that punitive damages should be reduced to reflect any public penalties that have been paid by the defendant for the same conduct.[157] Second, the U.S. Supreme Court has argued in *Gore* that the level of punitive damages should reflect the level of public penalties that *could be* imposed for comparable misconduct -- the higher the possible public sanctions, the higher should be punitive damages.[158]

How do these positions relate to our conclusions about punitive damages and deterrence? In this Section we observe that the view that punitive damages should be mitigated if public sanctions have been paid has a straightforward justification. However, we suggest that the view taken in *Gore*, that public penalties should serve as a benchmark for punitive damages, is problematic.

To begin, there is an obvious basis for subtracting any public penalties already incurred by the defendant from the level of punitive damages that otherwise would be appropriate: this

---

[157] The sixth *Green Oil* factor states: "If criminal sanctions have been imposed on the defendant for his conduct, this should be taken into account in mitigation of the punitive damages award." Green Oil Co. v. Hornsby, 539 So. 2d 218, 223-24 (Ala. 1989) (quoting Aetna Life Ins. Co. v. Lavoie, 505 So. 2d 1050, 1062 (Ala. 1987) (Houston, J., concurring), and endorsed in Pacific Mutual Life Ins. Co. v. Haslip, 499 U.S. 1, 22 (1991)).

[158] In *Gore*, the Court expressed this as follows: "Comparing the punitive damages award and the civil or criminal penalties that could be imposed for comparable misconduct provides a third indicium of excessiveness." BMW of N. Am., Inc. v. Gore, 116 S. Ct. 1589, 1603 (1996).

Lower courts have applied this standard. For example, in *Lee v. Edwards*, 101 F.3d 805 (2d Cir. 1996), the court found that punitive damages of $200,000 awarded by the jury in a § 1983 claim were excessive. In reducing the award to $75,000 (if plaintiff agreed to remittitur), the court noted that the defendant's conduct "could have exposed him to a charge of making a false statement"; if convicted, the defendant "would have faced imprisonment of up to one year and/or a fine of up to $2,000." *Id.* at 811. The court noted that although imprisonment was "a serious sanction," "the maximum fine of $2,000 gives little warning that the offense could entail a $200,000 civil award." *Id.* The award of $75,000 was justified in part because the defendant was a police officer and on "notice as to the gravity of misconduct under color of his official authority," *id.*, and in part by reference to other criteria. Also, in *Management Computer Servs., Inc. v. Hawkins, Ash, Baptie & Co.*, 557 N.W.2d 67 (Wis. 1996), involving the unauthorized copying of computer software, the court reduced a punitive damages award from $1.75 million to $650,000. As part of its *Gore* analysis, the court noted that the defendant's wrongdoing resulted in damages of $65,000 to the plaintiff, and that "the potential criminal penalty for copying computer programs if the damage is greater than $2,500 is a fine not exceeding $10,000." *Id.* at 82-83.

is necessary to ensure that the total payment made by the defendant is the proper amount for purposes of deterrence. As the reader knows, the defendant's total payment should be such that his expected payment equals the harm done. If punitive damages are not reduced from the amount implied by our formula to reflect penalties borne by the defendant, the defendant's combined private and public payments would result in his expected payments exceeding the harm done.[159]

Now consider the use of public penalties as a benchmark for setting punitive damages, the approach suggested in *Gore*. This role for public penalties makes sense only if their level conveys information relevant to determining the proper amount of punitive damages. The question naturally arises, therefore, whether the level of public penalties implies something about, among other things, the chance of escaping liability. Ostensibly, the answer is yes. For example, suppose that significant public penalties are imposed on restaurants for food poisoning because food poisoning often will not lead to suit. Then, conceivably, a court reviewing a punitive damages judgment in a case against a restaurant for food poisoning might use this information, inferred from the magnitude of the public penalties, to justify the award.[160]

Nevertheless, we are skeptical about whether the information that usually can be inferred from public penalties will be very useful, given the information that courts already will have about a case. In the course of a trial, a court will typically obtain information about

---

[159] This statement presumes for simplicity that the outcomes of the private and the public suits are identical -- they either both succeed or they both fail. Otherwise, the proper adjustment of the amount implied by our formula does not necessarily involve simply subtracting the amount paid as a public penalty. To illustrate, consider the following example. Suppose that the harm suffered by the victim is $5,000 and that the injurer has a 10 percent chance of being found liable as a result of a private suit. Suppose also that the government will impose a $1,000 fine on the injurer with certainty (it is not essential to this example that the fine is certain, only that it is imposed with a higher probability than the private plaintiff's probability of prevailing against the defendant). Let $A$ be the amount awarded to the private plaintiff if he prevails, with $A$ set such that the defendant's combined expected public and private payments equal the harm caused. In other words, $A$ is set such that $\$1{,}000 + .1A = \$5{,}000$. Solving for $A$ yields $40,000 as the proper private award. If instead our formula were applied and the public penalty were simply subtracted from the amount implied by our formula, the private award would be $49,000: the $50,000 award implied by our formula (= $5,000/.1) less the $1,000 public penalty.

[160] The point of this paragraph -- that public penalties can serve as useful guidelines for the setting of punitive damages -- is discussed by Cooter, *Deterrence*, *supra* note 12, at 1179-80.

the likelihood of escaping liability that is particular to that case. For example, a court might learn whether or not it would be easy to link harm from food poisoning to the defendant's restaurant (the type of poisoning may make identification straightforward or difficult), or whether or not the magnitude of the harm from the poisoning would be sufficient to induce suit (the poisoning may or may not result in expensive hospitalization and substantial lost wages). By contrast, the information implicit in public sanctions for food poisoning reflects, one presumes, only the average likelihood of liability over the range of cases of food poisoning.

There are further difficulties involved in inferring useful information from the level of public penalties. Such penalties are influenced in part by political factors -- interest group pressures, logrolling, and the like. Consequently, it would be hard for courts to determine in any precise way what legislators thought about the likelihood of escaping liability when they set the level of public sanctions. Another complication is that public penalties may themselves be influenced by the possibility of punitive damages awards in private suits: public penalties might be low *because* legislators believed that punitive damages awards would create effective deterrence. If the courts then constrain such awards on the ground that public penalties for comparable conduct are low, deterrence will tend to be inadequate, due to a kind of circularity -- the legislature relying on the courts and the courts relying on the legislature.[161] Because of the possibility of such circularity, the information inferred by the courts from the level of public penalties may be misleading.

For the foregoing reasons, we believe that courts generally should not use public sanctions as a benchmark in setting punitive damages. Such sanctions should be used, however, as an offset: any public penalties paid for the same conduct at issue should reduce the magnitude of punitive damages calculated according to our formula.

### H. Tax Treatment of Punitive Damages

If punitive damages are imposed on a defendant as a result of his engaging in some

---

[161] A different type of circularity could result in excessive deterrence. Suppose the legislature sets public penalties at high levels because it expects inadequate use of punitive damages. If the courts then impose substantial punitive damages because public penalties are high, deterrence could be excessive.

business or other income-earning activity, such damages generally can be deducted from taxable income, just as can compensatory damages in those circumstances.[162] But neither punitive nor compensatory damages are deductible if they are incurred as a result of the defendant's engaging in a non-business or personal activity.[163]

We explain here why these policies are desirable, given the goal of creating appropriate deterrence.[164] In the business context, the essence of the argument for the deductibility of punitive damages is that, were they not deductible, overdeterrence would result because the punitive damages component of liability would be more significant than it should be. Conversely, in the non-business or personal context, if punitive damages were deductible, underdeterrence would result.[165]

To see why damages should be deductible in a business context, consider a simple example in which a harm of $10,000 can be prevented by taking a precaution. It is socially desirable that the precaution is taken if it costs less than $10,000, and not taken if it costs more than $10,000. We first will show that if the injurer always will be sued and have to pay $10,000 in compensatory damages, then precautions will be taken precisely when they should be if the damage payment is deductible. Suppose that the precaution costs $8,000 and it is not taken. Then the defendant will pay damages of $10,000, and if these are deductible at, say, a 40 percent tax rate, the defendant will bear after-tax damages of $6,000. If the defendant does take the precaution, he pays $8,000 for it, but since this is a deductible expense, the after-tax

---

[162] *See* 2 STUART M. SPEISER ET AL., THE AMERICAN LAW OF TORTS § 8:64 (Supp. 1996); ROBERT W. WOOD, TAXATION OF DAMAGE AWARDS AND SETTLEMENT PAYMENTS ¶ 6.292 (1991 & Supp. 1996).

[163] *See* WOOD, *supra* note 162, ¶ 6.6

[164] Because our focus is on properly deterring potential injurers, we do not consider the tax treatment of *plaintiffs' receipts* of punitive damages awards. However, whether punitive damages are taxable income to the recipient may affect a plaintiff's incentive to sue. In this indirect way, the tax treatment of the receipt of punitive damages might affect deterrence.

[165] *See* I. P. L. Png & Eric M. Zolt, *Efficient Deterrence and the Tax Treatment of Monetary Sanctions*, 9 INT'L REV. L. & ECON. 209 (1989) (noting the different tax treatments of monetary sanctions, and arguing that to avoid overdeterrence, monetary sanctions should be deductible or the amount of the sanction should be adjusted to account for the offender's tax rate); Eric M. Zolt, *Deterrence Via Taxation: A Critical Analysis of Tax Penalty Provisions*, 37 UCLA L. REV. 343, 364-70 (1989) (expanding on the analysis presented in the Png & Zolt article).

cost of the precaution is $4,800. Hence he will take the precaution. Alternatively, if the precaution costs more than $10,000, the defendant is better off paying damages of $10,000 and deducting this amount than spending more on the precaution. Thus, the defendant will act optimally.

To put the point differently, because the defendant is able to deduct all of his expenses, whether damages or precautions, he will want to act so as to minimize his after-tax costs and thus will choose the precaution if and only if its cost after taxes is less than the damages the defendant would bear after taxes. Since the tax rate is the same whether applied to deducting precaution costs or damages, the defendant's behavior is equivalent to his choosing the precaution if and only if its cost is less than the damages (putting tax considerations aside), which is the behavior that is desired.

If damages were not deductible, a business actor might take precautions even when they cost more than the harm. Consider a firm in a 40 percent tax bracket that is deciding whether to take a precaution costing $15,000 that would prevent a harm of $10,000. The firm will take the precaution if damages are not deductible: the after-tax cost of the precaution is $9,000, which is less than the after-tax cost of compensatory damages, $10,000. This decision is socially undesirable because the cost to society of the precaution is $15,000, while the benefit to society is $10,000.

The reason that the firm is led to take a socially-wasteful precaution is that damages are not deductible, but the precaution cost is, so that the effective cost of the damages to the firm is heightened. Indeed, with a tax rate of 40 percent, damages appear to be one-and-two-thirds as important to eliminate as they would be if they were deductible, meaning that the firm would be willing to spend up to $16,667 to eliminate a $10,000 harm.[166]

---

[166] To see this, let $C$ be the cost of the precaution, $H$ the harm, and $T$ the tax rate. Ideally, the precaution should be taken if $C < H$. If both the precaution cost and damages are deductible, the firm will take the precaution if $(1 - T)C < (1 - T)H$, which is equivalent to $C < H$. But if damages are not deductible, the firm will take the precaution if $(1 - T)C < H$, which is equivalent to $C < H/(1 - T)$. Thus, if $H$ = $10,000 and $T$ = .4, the firm would be willing to spend up to $10,000/(1 - .4) =$16,667 to eliminate the harm. In other words, for every dollar of harm, the firm would be induced to spend up to $1/(1 - .4) = $1.67 to eliminate it. It is in this sense that damages appear to be one-and-two-thirds more important to eliminate than they would be if they were deductible.

Note that the logic of our discussion in the text implies that if damages were not deductible but were reduced appropriately, then the firm could be induced to take optimal precautions. If the tax rate is 40% and the harm is

The explanation that we have provided for the desirability of allowing compensatory damages to be deducted in a business context applies equally to punitive damages. Punitive damages are just another form of damages that are intended to make the expected damages of injurers equal to the harm they do. If punitive damages were not deductible, but precaution costs are deductible, overdeterrence would result for essentially the same reason as that discussed above: the non-deductibility of punitive damages would make causing harm more costly to business actors than expenditures to prevent harm, so that such actors would be induced to spend too much to reduce harm.[167]

The importance of allowing punitive damages to be deductible in a business context may be substantial because the tax rate for corporations and other business actors is relatively high. As we noted, at a tax rate of 40 percent,[168] if damages were not deductible a potential injurer might be induced to spend up to one-and-two-thirds the harm to prevent it.[169] Thus, if

---

$10,000, then reducing damages to $6,000 and not allowing damages to be deductible is equivalent to keeping damages at $10,000 but allowing their deductibility. Reducing damages by the precise amount that would be necessary to avoid distortions is functionally equivalent to allowing deductibility. But a policy of allowing deductibility may be preferable on administrative grounds since it obviates the need to determine the defendant's marginal tax bracket in order to properly calculate damages.

[167] For instance, in the example that we considered in which the harm is $10,000, suppose that the defendant is caught one time out of three, so that a third of the time he pays $30,000, consisting of $10,000 in compensatory damages and $20,000 in punitive damages. If the punitive damages component is not deductible and the defendant is in a 40% tax bracket, his after-tax liability cost if a judgment is rendered against him is $26,000 (an after-tax cost of $6,000 associated with the compensatory damages component and an after-tax cost of $20,000 associated with the punitive damages component). Because there is a one-third chance that he will bear this amount, his expected after-tax liability cost is $8,667 (that is, $26,000 divided by 3). Since his precaution expenditures are deductible, and he is in a 40% tax bracket, he would be willing to spend up to one-and-two-thirds of this amount to avoid this liability cost. In other words, he would be willing to spend up to $14,445 (= (5/3) x $8,667). But the harm is only $10,000, so there could be many instances in which he will be induced to spend substantially more than $10,000 to avoid imposing a harm of $10,000 -- a socially wasteful outcome caused by the non-deductibility of punitive damages.

[168] Under the federal tax code, taxable corporate income in excess of $75,000, but not in excess of $10 million, is taxed at a 34% rate. *See* 26 U.S.C. § 11(b)(1)(C) (1997). Taxable corporate income in excess of $10 million is taxed at a 35% rate. *See id.* § 11(b)(1)(D). The use of a 40% tax rate as an illustration in the text is reasonable in light of the additional state income taxes that corporations often would have to pay. *See, e.g.*, CAL. REV. & TAX. CODE § 23151(a) (West 1996) (8.84% of net income for 1997 and beyond); 35 ILL. COMP. STAT. 5/201(b)(7) & (d) (West 1997) (7.3% of net income); N.J. STAT. ANN. § 54:10A-5 (West 1996) (9% of net income); N.Y. TAX LAW § 210 (McKinney 1996) (9% of net income).

[169] The explanation why the injurer would be induced to spend up to one-and-two-thirds of the harm now is complicated by the fact that the fraction of total damages that is accounted for by punitive damages -- and therefore the portion of total damages that is not deductible -- depends on the chance of escaping liability. In note 167 *supra*, we showed that if there is a one-in-three chance of catching the injurer, he would be induced to spend $14,445 to

the deductibility of punitive damages in a business setting were disallowed, the overdeterrence that would result could be significant.

The explanation why punitive damages should not be deductible in a non-business or personal context is the mirror image of why they should be deductible in a business context. In a non-business or personal setting, the cost of precautions is not deductible. Hence, if punitive damages also are not deductible, a potential injurer will properly balance the cost of precautions against the reduction in harm from taking the precaution. If punitive damages were deductible, a potential injurer would not give sufficient weight to the reduction in harm from taking precautions, resulting in underdeterrence.

### I. Insurability of Punitive Damages

Policies regarding the insurability of punitive damages vary among states. Many allow punitive damages to be covered by liability insurance, but some do not.[170]

Should punitive damages be insurable? The basic answer to this question is in the affirmative, although we qualify this conclusion somewhat below. The reason why it generally is desirable to allow insurance for punitive damages is best understood by recognizing that punitive damages are, according to our theory, a way to make defendants pay for the harm they do when they have a chance of escaping liability. Thus, the question of whether punitive

---

prevent a harm of $10,000 -- that is, 1.44 times the harm. By similar logic, it can be demonstrated that if the chance of detection is sufficiently small -- so that nearly all of the damages paid by the injurer are punitive damages and not deductible -- the injurer will be induced to spend up to 1.67 times the harm to prevent it.

[170] A majority of jurisdictions follow the approach exemplified in *Lazenby v. Universal Underwriters Insurance Co.*, 383 S.W.2d 1 (Tenn. 1964), under which punitive damages are insurable. In general, however, insureds cannot indemnify themselves against punitive damages assessed for intentional misconduct. *See, e.g.*, Harrell v. Travelers Indem. Co., 567 P.2d 1013 (Or. 1977). A minority of jurisdictions follow the approach taken in *Northwestern National Casualty Co. v. McNulty*, 307 F.2d 432 (5th Cir. 1962), under which insurance coverage of punitive damages is disallowed because it would violate public policy (such coverage would permit wrongdoers to escape punishment and also would compromise deterrence). Even these jurisdictions, however, generally allow insurance coverage of punitive damages in cases of vicarious liability. *See, e.g.*, Ohio Cas. Ins. Co. v. Welfare Fin. Co., 75 F.2d 58 (8th Cir. 1934). For further discussion of the insurability of punitive damages, and summaries of the relevant law among the states, see generally ROBERT G. SCHLOERB, RICHARD L. BLATT, ROBERT W. HAMMESFAHR & LORI S. NUGENT, PUNITIVE DAMAGES: A GUIDE TO THE INSURABILITY OF PUNITIVE DAMAGES IN THE UNITED STATES AND ITS TERRITORIES (1988). General discussion of the insurability of punitive damages also can be found in 2 SPEISER, *supra* note 162, § 8:54 and 2 SCHLUETER & REDDEN, *supra* note 68, § 17.2.

damages should be insurable is essentially the same as the question of whether compensatory damages should be insurable.[171]

Of course, compensatory damages *are* insurable, but what reason can be given for allowing them to be? Consider this question when liability is strict and harm is entirely monetary. In this case, allowing the purchase of liability insurance is socially desirable. Liability insurance raises the well-being of potential injurers, for that is why they choose to buy it, and the availability of such insurance does not affect the welfare of victims, who will be fully compensated anyway. Even if the purchase of liability insurance causes injurers to take less care and thereby increases the frequency of accidents, victims will not be affected because they are fully compensated.[172]

However, if losses are nonmonetary, victims might not be fully compensated, or if the negligence rule applies, they might not be compensated at all.[173] Consequently, their welfare would be adversely affected if liability insurance leads to an increase in the frequency of accidents. Nevertheless, it can be shown that such insurance often is socially desirable even then, because the value of the insurance to insureds may exceed the loss of welfare to victims.[174] The victims' loss is mitigated because liability insurers have a financial incentive to structure coverage and premiums to control risks. For example, insurers may make insurance premiums depend on an insured's history of claims, may only offer partial coverage, or may require that the insured take certain steps to reduce risks. As a result, the purchase of liability insurance may not significantly reduce the insured's incentives to exercise precautions.

---

[171] For economically oriented discussion of the question of the social desirability of liability insurance for punitive damages, see Chapman & Trebilcock, *supra* note 12, at 821-22; Cooter, *Deterrence*, *supra* note 12, at 1182-85; Ellis, *supra* note 12, at 71-76; Priest, *Insurability*, *supra* note 12. *See also* note 172 *infra*.

[172] The argument in this paragraph is based on Shavell, *supra* note 41, which first formally analyzed the social desirability of liability insurance. *See also* SHAVELL, supra note 14, at 206-27. It is shown in these references that, in the standard model of accidents, both injurers and victims will be made better off if the sale of liability insurance is permitted.

[173] Victims might not be fully compensated if their losses are nonmonetary because, among other reasons, this may be impossible (the loss may be of a person's life). Victims may not be compensated at all under the negligence rule because the injurer may not have been negligent.

[174] See SHAVELL, *supra* note 14, at 251-52.

Moreover, if liability insurance were disallowed, not only would potential injurers' well-being decrease, they also might forgo engaging in some socially beneficial activities that pose liability risks. For example, surgeons might refuse to perform certain operations, or general aviation aircraft companies might decline to make planes. Even when services and products continue to be offered, prices would rise to cover the liability risks that such providers would now have to bear directly, and the resulting price increase often will exceed that which would have resulted from the purchase of liability insurance. Both of these consequences of disallowing liability insurance -- the possible withdrawal from certain activities and the increased price of other activities -- hurt consumers of the affected products and services. This consideration lends support to the case for allowing the sale of liability insurance.

A major complication in the preceding discussion concerning the desirability of liability insurance arises due to the judgment proof problem.[175] If injurers can avoid having to pay for some of the harm they cause because their assets are limited, they will have a reduced incentive to take precautions and control their participation in risky activities. (In the extreme, an injurer with no assets would have no liability-related incentive to reduce risks.) The sale of liability insurance could either worsen or ameliorate this problem. If insurers are substantially unable to control the risky behavior of insureds because insurers cannot easily observe insureds' risk-taking behavior and link policy features, such as premium rates, to insureds' behavior, then liability insurance would tend to exacerbate the judgment proof problem. It might be beneficial then to forbid the sale of such insurance. Conversely, if insurers can relatively easily observe and control the behavior of insureds, liability insurance could lessen the judgment proof problem. For instance, insurers might require that a restaurant install fire extinguishers and a sprinkler system to reduce fire risks. In this case, not only might it be desirable to allow the sale of liability insurance, it might even be beneficial to require its

---

[175] Before discussing this complication, we want to observe that a common view is that, if there is a chance that injurers will be unable to pay for harm, liability insurance is socially desirable because it enhances the ability of victims to collect damages from injurers. We do not consider this reason for liability insurance to be a strong one; as we have noted previously, *see* note 15 *supra,* we believe that first-party insurance is a superior way to provide compensation to victims because it is administratively cheaper than the tort system.

purchase.[176]

We have now explained why, with some qualifications, it generally is desirable to allow potential injurers to purchase liability insurance for compensatory damages. Because punitive damages are, by our formula, a substitute for compensatory damages when injurers can escape liability, essentially the same arguments imply that it generally is desirable also to allow the sale of insurance for punitive damages, as many states already do.

### J. Third Party Versus Consumer Victims

In our analysis of punitive damages, we have been assuming implicitly that the parties harmed by the injurer are "third parties" -- that is, parties who have no market or contractual relationship with the defendant. This was the case, for example, with respect to the fishermen and Alaskan natives whose livelihood was affected by the *Exxon Valdez* oil spill. In many situations, however, the victims are customers of the defendant. This was true, for instance, in *Gore*, where the plaintiff was a purchaser of a BMW.[177]

Courts devote little attention when determining punitive damages to whether the plaintiff was a third party or a consumer.[178] The list of factors in *Haslip*, for example, does not include this distinction, nor does any other similar list or authoritative source of which we

---

[176] The conclusions of this paragraph are based on an economic analysis of the judgment proof problem and insurance in Steven Shavell, *The Judgment Proof Problem*, 6 INT'L REV. LAW & ECON. 45 (1986), which is distilled in SHAVELL, *supra* note 14, at 240-43.

[177] Other cases in which customers were harmed include *Sears, Roebuck & Co. v. Harris*, 630 So. 2d 1018 (Ala. 1994) (customer harmed from self-installed gas water heater); *Moore v. Jewel Tea Co.*, 253 N.E.2d 636 (Ill. App. Ct. 1969) (customer harmed when unopened can of drain cleaner exploded), *aff'd*, 263 N.E.2d 103 (Ill. 1970); and *Leibeck v. McDonald's Corp.*, No. CV-93-2419 (N.M. Dist. Ct. 1994) (customer burned by hot coffee).

[178] This difference is not emphasized, for example, in any of the following U.S. Supreme Court cases on punitive damages: BMW of North Am., Inc. v. Gore, 116 S. Ct. 1589 (1996); Honda Motor Co., Ltd. v. Oberg 512 U.S. 415 (1994); TXO Production Corp. v. Alliance Resources Corp., 509 U.S. 443 (1993), Pacific Mut. Life Ins. Co. v. Haslip, 499 U.S. 1 (1991); Browning-Ferris Indus. v. Kelco Disposal, Inc., 492 U.S. 257 (1989). In *Gore*, however, the market relationship between the plaintiff and the defendant was mentioned as a factor that would lessen the need for affirmative disclosure requirements "because the self-interest of those involved in the automobile trade in developing and maintaining the good will of their customers will motivate them to make voluntary disclosures or to refrain from selling cars that do not comply with self-imposed standards." *Gore*, 116 S. Ct. at 1596.

are aware.[179]

However, the status of victims as third parties or consumers is important to consider, for when victims are consumers, the need for punitive damages is lessened. The reason is that, when individuals might be harmed by the products (or services) they buy, producers will tend to be concerned that customers may not be willing to pay as much for the products or that they may stop purchasing the products altogether. Given that producers have this market-based incentive to be attentive to the risk of harm to their customers, the need for liability in general, and for punitive damages in particular, to control their behavior is diminished.[180] Obviously, this market mechanism cannot operate if the victims are not customers of the defendant -- that is, if they are third parties.[181]

The extent to which market forces reduce the need for liability as a deterrent depends on how much customers know about product or service hazards. In some circumstances, customers will not be able to effectively discipline firms due to their lack of knowledge of such risks.[182] Because travelers probably would not know much about the chance of suffering food poisoning from eating at a family-owned restaurant at a turnpike stop, the restaurant would not be likely to fear loss of clientele if food poisoning occurs; thus the threat of liability, including

---

[179] The sections on punitive damages in the RESTATEMENT (SECOND) OF TORTS §§ 908-909 do not mention this distinction explicitly, nor do 1-2 SCHLUETER & REDDEN, *supra* note 68, or JAMES D. GHIARDI & JOHN J. KIRCHER, PUNITIVE DAMAGES LAW AND PRACTICE (1985). This distinction also is absent from cases that provide factors for the jury to consider in determining the amount of punitive damages. *See, e.g.*, Fischer v. Johns-Manville Corp., 512 A.2d 466, 481-82 (N.J. 1986); Estate of Hartz v. Nelson, 437 N.W.2d 749, 755-56 (Minn. Ct. App. 1989).

[180] To the extent that liability is not needed to promote product safety, imposing liability would be redundant and the costs associated with litigation would be socially wasteful.

[181] Craswell, *supra* note 12, discusses the role of damage multipliers in market relationships, although he does not emphasize the point that we do -- that producers have a market-based incentive to be attentive to the risk of harm to their customers.

[182] For the view that consumers are not well informed about product risks, see Howard Latin, *"Good" Warnings, Bad Products, and Cognitive Limitations*, 41 UCLA L. REV. 1193, 1234 (1994) (arguing that knowledge of "the great majority of product risks cannot be available to product users").

punitive damages, might be desirable to induce the restaurant to reduce this risk.[183]

But in many settings, consumer information about the dangers of products and services is relatively good.[184] This may be because the risks have a fairly obvious character; because they have been publicized by the media;[185] or because the customers are repeat purchasers and have learned about them from experience. In such circumstances, the threat of liability would be relatively unimportant in controlling risk. Indeed, if consumer information about risk were perfect, there would be no need at all for liability to improve product safety: consumers would reduce their willingness to pay for a firm's product or service by precisely the amount of the expected harm that the product or service exposed them to, which in turn would cause firms to invest in any cost-justified precaution.

Our conclusion, therefore, is that in deciding on punitive damages, courts should take into account whether the victims are third parties or customers and, if the latter, whether market forces are likely to lead sellers to properly reduce risk. A cautious approach in imposing punitive damages should be adopted when consumers are relatively well informed about the risk of the seller's product or service.[186]

---

[183] Daughety and Reinganum study the role of punitive damages in reducing product risks when consumers do not have direct information about the risks. *See* Daughety & Reinganum, Punitive Damages, *supra* note 12; Daughety & Reinganum, Settlement, *supra* note 12.

[184] *See* Patricia M. Danzon, *Comments on Landes and Posner: A Positive Economic Analysis of Products Liability*, 14 J. LEGAL STUD. 569, 572 (1985) (arguing that the cost of obtaining information about product hazards could be low in many circumstances and that the value of such information is high for consumer goods that are purchased repeatedly, durable consumer goods, and producer goods); Alan Schwartz, *Proposals for Products Liability Reform: A Theoretical Synthesis*, 97 YALE L.J. 353, 380 (1988) (arguing that "evidence drawn from surveys and actual market behavior more strongly supports the view that consumers are informed than the view that they are ignorant"). Even if consumer information about the risks of products and services is not widespread, markets may work reasonably well if a sufficiently large fraction of the population of potential consumers is well informed. *See* Alan Schwartz & Louis L. Wilde, *Intervening in Markets on the Basis of Imperfect Information: A Legal and Economic Analysis*, 127 U. PA. L. REV. 630, 637-39 (1979).

[185] For example, we would expect problems with automobiles (as in the *Gore* case, *see* text accompanying notes 79-80 *supra*) to come to the attention of consumers through stories in newspapers, evaluations in *Consumer Reports*, and the like.

[186] A similar conclusion applies when the victims of accidents are employees rather than customers. To the degree that employees are aware of workplace risks, they will insist on higher wages (or they may seek employment elsewhere). Thus, market forces will tend to induce employers to increase workplace safety even in the absence of liability.

## K. Breach of Contract

Although we have been discussing the imposition of punitive damages in situations governed by tort law, punitive damages sometimes can be levied in contractual disputes as well.[187] Indeed, there seems to be an increasing tendency to employ such damages in the area of contract[188] -- for example, in employment termination and insurance litigation.[189]

The argument for awarding punitive damages in contract cases, however, is weak. As we will explain, this is because, on one hand, contracting parties themselves should, at least in principle, be allowed to include a provision for an appropriately high penalty for breach of contract if it serves their mutual purposes.[190] On the other hand, if the courts award punitive damages when the parties to a contract have not provided for extra-compensatory damages for breach, the value of the contract to them tends to be reduced.[191]

Parties may find it worthwhile to include a term in their contract that provides for

---

[187] While punitive damages traditionally are not awarded in contract cases, exceptions often are made when the wrongful conduct is also considered to be a tort. *See, e.g.*, 1 SCHLUETER & REDDEN, *supra* note 68, § 7.3.

[188] *Id.* § 7.0 ("[O]ver the last twenty years, the courts have broken down the traditional doctrinal barriers between contracts and torts. The result is a growing list of exceptions to the general rule and a growing recognition of punitive damages within the law of contracts."); *see also* Mark Pennington, *Punitive Damages for Breach of Contract: A Core Sample from the Decisions of the Last Ten Years*, 42 ARK. L. REV. 31 (1989) (noting departures from the traditional rule against awarding punitive damages for breach of contract); John A. Sebert, Jr., *Punitive and Nonpecuniary Damages in Actions Based Upon Contract: Toward Achieving the Objective of Full Compensation*, 33 UCLA L. REV. 1565 (1986) (noting trend toward allowing punitive damages in contract cases).

[189] *See* 2 SCHLUETER & REDDEN, *supra* note 68, § 13.3(B) (discussing punitive damages in the employment context); *id.* §§ 17.3(A) & 17.4(A) (discussing punitive damages in insurance cases with respect to contract theories).

[190] In practice, however, such provisions -- liquidated damage clauses -- are not enforced by courts if they are determined to exceed compensatory damages. *See* RESTATEMENT (SECOND) OF CONTRACTS § 356 (1981) (a liquidated damages clause will be enforced if the contract specifies "an amount that is reasonable in the light of the anticipated or actual loss caused by the breach and the difficulties of proof of loss," but "[a] term fixing unreasonably large liquidated damages is unenforceable on grounds of public policy as a penalty."); *see also* U.C.C. § 2-718(1) (West 1996).

[191] Any term imposed on the parties that they could have freely chosen, but did not, presumptively makes them worse off. This statement applies to the level of the damage award.

On the general economic role of punitive damages in breach of contract disputes, see, for example, Daniel A. Farber, *Reassessing the Economic Efficiency of Compensatory Damages for Breach of Contract*, 66 VA. L. REV. 1443 (1980), Barry Perlstein, *Crossing the Contract-Tort Boundary: An Economic Argument for the Imposition of Extracompensatory Damages for Opportunistic Breach of Contract*, 58 BROOK. L. REV. 877 (1992), Alan Schwartz, *The Myth That Promisees Prefer Supercompensatory Remedies: An Analysis of Contracting for Damage Measures*, 100 YALE L.J. 369, 370-72, 394-405 (1990).

extra-compensatory damages -- damages exceeding the value of performance (so-called expectation damages). They will do this when extra-compensatory damages are necessary to induce the promisor to perform adequately. For example, suppose that a company contracts with a city to replace its burned-out street lights. It is plausible that there would be a contractual provision stating that, for any light that the city discovers was not replaced on a timely basis (such as one week after burning out), the contractor will have to pay the city a penalty of $250, which is much more than the loss due to the absence of the light. The reason we can imagine that the parties would include this provision is that they both recognize that the city will not discover most of the lights that burn out and that are not repaired on a timely basis. They jointly realize that a penalty for breach in excess of the loss from breach will give the repair company a stronger and more appropriate motive to search for and replace burned-out street lights than would compensatory damages. Because the city will be willing to pay more to the company for its better service, both parties to the contract can benefit from the provision of a penalty clause.[192]

Note that the circumstances in this example are analogous to those in the tort settings in which we have said that punitive damages are desirable -- namely, there is a probability that a

---

[192] To illustrate, suppose that: (a) each light will definitely fail at some time during the year; (b) the value to the city of timely repair of a light is $50; (c) the cost to the contractor of assuring timely repair is $25; (d) the cost to the contractor of less-than-timely repairs is $5 (the contractor does not have to check lights as frequently); (e) the likelihood of the city's detecting breach (failing to repair a light on a timely basis) is 20%; (f) the city's payment to the contractor is $35 per light per year; and (g) the penalty for breach is $50 (the value of timely repair).

Note that the contractor will not be induced to spend $25 (instead of $5) to assure timely repair. This is because his expected penalty payment per light will be only $10 (= 20% x $50) if he does not spend the extra $20 to assure timely repair. Thus, the contractor's profit per light will be $20 -- the city's payment of $35 less the $5 cost of repairs, and less the $10 expected penalty payment. The city's total cost per light is $75 -- its $35 payment to the contractor, plus its $50 loss of value as a result of the contractor's failure to repair the light on a timely basis, less the $10 it receives in expected penalty payments.

We want to show that *both* parties can benefit if the penalty for breach is raised to $250. Under this penalty, the contractor will be led to spend the extra $20 to assure timely repair, since now his expected penalty payment per light will be $50 (= 20% x $250) if he does not assure timely repair. (Consequently, there will be no breach and no penalties actually paid.) Because timely repair is assured, the city should be willing to pay the contractor more. Suppose the payment is raised to $55 per light. Then the contractor's profits per light will be $30 ($55 less the $25 cost of repair), so he will be better off (his profits had been $20 per light). The city's cost per light now will be $55 (just the fee), so it will be better off too (its cost had been $75).

The source of the mutual benefit for the parties is that use of the $250 penalty induces the contractor to create a $50 benefit for the city at only an extra $20 cost to itself. This enables the city to make a payment sufficiently higher to make the contractor better off and still leave itself better off.

party will not be found liable when he does harm (in the present context, by committing a breach). Thus, the role of the penalty for breach in the example resembles the role of punitive damages in tort situations, to make up for the chance of escaping liability.[193]

Parties may want extra-compensatory damages to be paid for breach in two contexts in which a breaching party may escape liability. One is when the breached-against party does not automatically observe whether performance has occurred; this was the situation in the street light example.[194] The other situation is when the breached-against party does know that performance has been deficient but may not be able to prove this in court or would not have a financial incentive to sue. For instance, if an insurance company fabricates a reason for not paying a small claim, the insured may not sue due to the uncertainty of success and the cost of a lawsuit. But because insureds would in principle be willing to pay higher premiums if an insurance company can be deterred from acting in this way, the insurer may benefit from an agreement to pay extra-compensatory damages when it is found liable for falsely denying a small claim.

Sometimes parties may not have included extra-compensatory damages in their contract even though such damages are desirable, because the parties were ill-informed or because they faced high contracting costs. If courts can identify these circumstances, it then would be desirable for courts to impose punitive damages for breach.

In many circumstances, however, parties will not want damages for breach of contract to exceed the compensatory level because the fact of breach is obvious, the nature of the breach is such that it easily can be proven in court, and the amount at stake is large enough to justify a lawsuit. To the degree that courts impose punitive damages in these situations, such

---

[193] We should distinguish the present discussion from that in the previous section, which concerned injured parties who were customers. There we assumed that buyers' knowledge of product risks might be able to induce sellers to take appropriate precautions. Here we are *not* making an analogous assumption. In our present example, we did not consider the possibility that the contractor would repair street lights on a timely basis due to a concern that his business reputation might otherwise suffer. Rather, we assumed (implicitly) that the prospect of damages for breach of contract is needed to motivate the contractor to repair lights on a timely basis.

[194] More generally, whenever a party buys a large quantity of a product and does not inspect every unit to determine whether the product complies with the specifications in the contract, there can be a problem of detecting a breach.

damages will result in excessive and expensive performance (the analogue of overdeterrence), and thereby lower the welfare of the contracting parties.

In conclusion, therefore, courts should be cautious about awarding punitive damages for breach of contract. This point is worth noting because the law governing the imposition of punitive damages for breach of contract does not restrict their award to cases in which the likelihood of escaping liability for breach is substantial.[195]

### L. Components of Harm Not Included in Compensatory Damages

It often is suggested that punitive damages should be awarded to compensate plaintiffs for non-economic and other losses that would not otherwise be incorporated into compensatory damages.[196] Many courts have endorsed this justification for punitive damages.[197]

Although we recognize that awarding punitive damages as a substitute for a missing component of harm has a potential rationale in terms of assuring proper deterrence, we suggest in this section that remedies for missing components of harm are best pursued through revision of the rules for calculating compensatory damages.

---

[195] For example, courts have awarded punitive damages in contract cases where there is a "special relationship between the parties." 1 SCHLUETER & REDDEN, *supra* note 68, § 7.3(A). These relationships include: "bank and depositor, employer and employee, franchiser and franchisee, lawyer and client, public utility and customer, and security broker and customer." *Id.* (citations omitted). The courts argue that punitive damages are appropriate because one party has greater bargaining power. *See id.* But this superior position, in and of itself, does not suggest that the party with the upper hand will escape liability for a breach of contract.

[196] *See* Chapman & Trebilcock, *supra* note 12, at 768-69 (suggesting that punitive damages serve as a means of compensating for dignitary loss); Dorsey D. Ellis, Jr., *Punitive Damages in Iowa Law: A Critical Assessment*, 66 IOWA L. REV. 1003, 1006, 1010 (1981) (noting the use of punitive damages as compensation for a variety of nonpecuniary harms); Ellis, *supra* note 12, at 3 (noting that compensating victims for otherwise uncompensable losses and paying the plaintiff's legal fees are reasons often cited by legal commentators and courts for imposing punitive damages); Galligan, *supra* note 3 (emphasizing that compensatory damages generally should be augmented to reflect otherwise missing elements of harm); David G. Owen, *Punitive Damages in Products Liability Litigation*, 74 MICH. L. REV. 1257, 1295-96 (1976) ("[P]unitive damages do indeed play an important -- even if usually residual -- compensatory role.").

[197] *See* SCHLUETER & REDDEN, *supra* note 68, § 2.2. For example, in Connecticut "exemplary damages cannot exceed plaintiff's expenses, and therefore, in fact and effect are considered compensatory." *Id.* (citing Doroszka v. Lavine, 150 A. 692 (Conn. 1930) and Craney v. Donovan, 102 A. 640 (Conn. 1917)). In Michigan, "exemplary damages are granted to compensate the plaintiff and not to punish the defendant." *Id.* (citing Oppenhuizen v. Wennersten, 139 N.W.2d 765 (Mich. 1966)). In New Hampshire, exemplary damages "are also considered as compensation for the mental injury to the plaintiff." *Id.* (citing Fay v. Parker, 53 NH 342 (1872)).

As the reader knows, our basic analysis of deterrence implies that injurers should have to pay for the entire harm they cause; this is necessary to induce them to take appropriate precautions and so that prices and participation in risky activities are proper. Thus, if there is a component of harm that otherwise would be omitted, a policy of including it in the form of punitive damages would seem to be beneficial.[198]

Notwithstanding this point, there is a problem with employing punitive damages as a substitute for missing components of compensatory damages. Namely, a component of harm might be excluded from compensatory damages because of the difficulties and expense that would be encountered in its estimation. Consider, for example, the pain and suffering experienced by the friends of a person who dies. Were this category of harm included in compensatory awards, the number of claimants in cases of wrongful death could become quite large, and the cost of litigation would also increase as a result of parties contesting the magnitude of their psychological losses. It may well be best, then, for the law to exclude from compensatory damages many such speculative, difficult-to-determine elements of harm, even though these elements are real and their omission does undesirably dilute deterrence.[199]

If a component of loss is excluded from compensatory damages for these reasons, arguably it should be excluded from punitive damages as well. The disadvantages of attempting to ascertain the missing component of harm would not be lessened just because it is calculated under a different head of damages. It will be no easier to determine the pain and suffering due to the death of a friend just because this loss is imported into punitive damages. Indeed, the accuracy of measurement of this loss would be expected to be worse because the calculation of punitive damages is not disciplined by the procedures and evidentiary

---

[198] More precisely, the principle would be to include a multiple of the missing component, with the multiplier determined by the defendant's chance of escaping liability.

[199] The costs of estimating such elements of harm on a case-by-case basis could be largely avoided, however, if courts were to use a table listing standard values of the missing components. This would be an inexpensive (essentially costless) method of including missing components of damages, and it would be preferable to excluding them.

requirements common to the determination of compensatory damages.[200]

Of course, if a component of loss should have been included in compensatory damages, despite the costs of doing so, then the natural response is to rectify the mistake by incorporating it in compensatory damages. If instead the component of loss is included as part of punitive damages, not only will it be less accurately measured for the reason noted in the previous paragraph, there will be an additional problem: the component will be omitted in the large majority of cases, those in which only compensatory damages are awarded.

### M. Economic Loss versus Personal Injury

Several legal opinions express the view that the level of punitive damages should depend on whether the plaintiff's harm involved personal injury or was entirely economic.[201] For example, much was made of this distinction in *Gore*.[202]

Does it make sense for punitive damages to be influenced by whether the harm consists of a personal injury as opposed to an economic loss? The answer is basically no. Assuming that the amount that courts award in compensatory damages in personal injury cases is

---

[200] Juries are given broad discretion over the award of punitive damages, with relatively little guidance from the courts. *See* 1 GHIARDI & KIRCHER, *supra* note 179, § 5.38 ("It is a generally accepted rule that once the court determines that the evidence merits submitting the punitive damages issue to the jury, it is entirely within the discretion of the jury to determine whether damages should be awarded at all and to determine the amount which should be awarded."); 1 SCHLUETER & REDDEN, *supra* note 68, § 6.1(A) (referring to the jury as "less restricted in awarding punitive damages than in awarding compensatory damages").

[201] *See, e.g.*, Lightning v. Roadway Express, Inc., 60 F.3d 1551, 1559 (11th Cir. 1995) ("In determining the reasonableness of an award of punitive damages, courts should consider whether . . . the misconduct caused personal injury or merely damage to property . . . ."); Eisert v. Greenberg Roofing & Sheet Metal Co., 314 N.W.2d 226 (Minn. 1982) (The nature of the plaintiff's injury "may reasonably be taken into account in deciding where punitive damages will be allowed. Where that injury is limited to property damage, the public interest in punishment and deterrence is largely satisfied by the plaintiff's recovery of compensatory damages."); 1 SCHLUETER & REDDEN, *supra* note 68, § 9.5(A), at 536 ("[C]ourts distinguish between whether property damage or personal injury was the result of the defendant's wrongdoing.").

[202] In *BMW of North America v. Gore*, 116 S. Ct. 1589 (1996), the Court contrasted the "purely economic" harm inflicted by the defendant with instances of "reckless disregard for the health and safety of others," implying that the latter acts are more reprehensible and therefore should be subject to higher punitive damages. *Id.* at 1599.

proper,[203] the formula that we have advanced for the determination of punitive damages should apply without modification: the level of compensatory damages for the personal injury should be multiplied by the inverse of the probability of being found liable.

We recognize, however, that the level of compensatory damages awards in personal injury cases may be too low in practice to accomplish proper deterrence.[204] For example, it has been calculated that in wrongful death cases, the amount that an injurer should pay is between $3 million and $6 million,[205] whereas actual awards are usually substantially lower.[206] If compensatory damages are too low in personal injury cases, they should be raised appropriately.[207] Punitive damages should not be awarded to correct for inadequate compensatory damages, for reasons analogous to those discussed in the previous section.

---

[203] By the proper amount of compensatory damages in this context, we mean the amount of damages that induces a potential injurer to take optimal precautions to prevent personal injury. This amount might not compensate the victim for an injury -- indeed that might be impossible (as in the case of loss of life). Thus, the term "compensatory damages" may be a misnomer when applied to personal injuries, but we employ it because it is used to describe the usual level of damages.

[204] Notably, compensatory damages in wrongful death cases are generally calculated as a survivor's financial loss. *See* CHARLES T. MCCORMICK, HANDBOOK ON THE LAW OF DAMAGES §§ 93-106 (1935). This amount usually will not lead to proper deterrence. For example, if a child or a non-working spouse is killed, the financial loss will be low, but the event will be one for which expensive preventive measures are justified to reduce risk. Such measures may not be taken if damages are based solely on the financial loss. For discussion of this and related points about the distinction between optimal compensation for non-monetary losses and optimal deterrence, see Philip J. Cook & Daniel A. Graham, *The Demand for Insurance and Protection: The Case of Irreplaceable Commodities*, 91 Q.J. Econ. 143 (1977); Michael Spence, *Consumer Misperceptions, Product Failure, and Product Liability*, 64 REV. ECON. STUD. 561 (1977); SHAVELL, *supra* note 14, at 228-35.

[205] *See* VISCUSI, *supra* note 29, at 108 (using the observed risk-dollar tradeoff of blue-collar workers to calculate that the implicit value of life is between $3 and $6 million); Michael J. Moore & W. Kip Viscusi, *The Quality-Adjusted Value of Life*, 26 ECON. INQUIRY 369, 386 (1988) (finding an implicit value of life of $6 million).

[206] *See, e.g.*, JAMES S. KAKALIK ET AL., COSTS AND COMPENSATION PAID IN AVIATION ACCIDENT LITIGATION (Institute for Civil Justice, RAND Corp., No. R-3421-ICJ, 1988). In this study of 25 major airline accidents occurring between 1970 and 1984, *id.* at 4, the authors calculated that the average compensation for airline accident deaths was $321,300 from 1970-1976, and $408,500 from 1977-1982 (measured in constant 1986 dollars), *id.* at 20. *See also* Randall R. Bovbjerg, Frank A. Sloan & James F. Blumstein, *Valuing Life and Limb in Tort: Scheduling "Pain and Suffering,"* 83 NW. U. L. REV. 908, 920-23 (1989) (surveying Florida and Kansas jury verdicts and finding the median loss of life award to be $620,000, and the mean loss of life award to be $1,224,000).

[207] In saying this, we are presuming that problems of implementation, or considerations of cost, do not subvert this recommendation. If they do, then a table of standard values for different types of personal injuries could be used instead. *See* note 199 *supra*.

## N. Externalization of Risk through Independent Contractors

An effect of imposing liability that we have not yet discussed is what we will call *externalization of risk*. By this we mean the ability of potential injurers to avoid liability by hiring independent contractors to undertake risky tasks that they would otherwise perform themselves. The motive to externalize risks results to some extent from the threat of compensatory damages alone, but it is accentuated if punitive damages are awarded. To our knowledge, courts do not consider this factor in the determination of punitive damages.[208]

In this section, we discuss two socially undesirable consequences of externalization of risk. First, the number of accidents that occurs tends to be higher because the independent contractors who are engaged generally do not operate as safely as the firms hiring them. Second, society sacrifices the economic benefits that would have accrued if the firms had carried out certain tasks themselves instead of having them performed by independent contractors. We conclude that when the factor of externalization of risk is relevant, it argues for caution in the use of punitive damages, other things being equal.

Let us amplify on these points. It is well recognized that firms often can avoid liability by hiring independent contractors to undertake risky tasks.[209] For example, a firm that is transporting its toxic waste in its own trucks may be able to hire another company to transport the waste and thereby avoid liability for spills.[210]

---

[208] For example, the issue of externalizing risk is not mentioned in any of the recent major U.S. Supreme Court cases discussing punitive damages. *See, e.g.*, BMW of N. Am., Inc. v. Gore, 116 S. Ct. 1589 (1996); Honda Motor Co., Ltd. v. Oberg, 512 U.S. 415 (1994); TXO Prod. Corp. v. Alliance Resources Corp. 509 U.S. 443 (1993); Pacific Mutual Life Ins. Co. v. Haslip, 499 U.S. 1 (1991); Browning-Ferris Indus., Inc. v. Kelco Disposal, Inc., 492 U.S. 257 (1989).

[209] *See* RESTATEMENT (SECOND) TORTS § 409 (stating the "employer of an independent contractor is not liable for physical harm caused to another by an act or omission of the contractor or his servants"); PROSSER AND KEETON ON THE LAW OF TORTS § 71, at 509 (5th ed. 1984) ("For the torts of an independent contractor, . . . it has long been said to be the general rule that there is no vicarious liability upon the employer.").

[210] However, this strategy may not always work. For example, in *Kenney v. Scientific, Inc.*, 497 A.2d 1310 (NJ 1985), the court held the defendant liable for an independent contractor's transportation of toxic waste. The court stated, "A company which creates the Frankenstein monster of abnormally dangerous waste should not expect to be relieved of accountability for the depredations of its creature merely because the company entrusts the monster's care to another, even an independent contractor." *Id.* at 1321. Notwithstanding such exceptions, we will address situations in which it is possible to shift liability by hiring an independent contractor.

However, the incentive of a firm to transfer liability to an independent contractor is more complicated than may at first appear. This is because, if an independent contractor assumes liability by undertaking risky tasks for a firm, the contractor will charge the firm more for performing these tasks. An independent contractor that is hired to haul a firm's toxic waste clearly will charge the firm an amount reflecting its expected damages for spillage of the waste. Thus, the firm might in the end pay for the accident risks it creates even though it hires an independent contractor. It is apparent, therefore, that the firm will want to hire an independent contractor only if the contractor would charge the firm less for assuming liability than the firm would have borne itself.

In fact, an independent contractor might be willing to charge a firm less. The reason is that an independent contractor might not have assets sufficient to cover the full liability it may incur, so that its effective expected damages are lower than the firm's -- assuming that the firm does have the assets necessary to pay the full judgment it would have faced, or at least more assets than the contractor. For example, suppose potential damages are $10 million and the risk of liability is 5 percent. If a firm with assets of $10 million undertakes the risky task itself, its expected damages would be $500,000 (= 5% x $10 million). However, if the firm hires an independent contractor with assets of only $1 million, the contractor's expected damages would be $50,000 (= 5% x $1 million). Hence, the independent contractor only needs to add $50,000 to the price it charges the firm to be compensated for its expected damages. Accordingly, the firm could in effect reduce its expected damages from $500,000 to $50,000 by hiring the independent contractor.[211]

We have explained why firms might benefit if they externalize their liability risks to independent contractors with assets less than the harm resulting from accidents. It should be

---

[211] The point of this example holds even if the firm does not have assets sufficient to pay for the full $10 million in damages; as long as its assets exceed those of the independent contractor, there still is a potential gain to the firm from externalizing the risk to the contractor. For instance, suppose that the firm has assets of only $5 million. Its expected liability then would be $250,000 (= 5% x $5 million), which still exceeds the independent contractor's expected liability of $50,000. Thus, the firm will continue to have an incentive to externalize its risk. (The general condition for when it will be advantageous to the firm to use an independent contractor to externalize its risk is twofold: the contractor's assets must be less than the firm's; and there must be a positive probability that the judgment will exceed the contractor's assets.)

emphasized that because firms secure an advantage from dealing with such contractors, they will seek them out and favor them over contractors with greater assets, other things being equal.[212]

The externalization of risk to potentially judgment proof contractors has an important implication. These contractors will tend to conduct their activities with less care than actors with more at stake. In the example above, the independent contractor with assets of only $1 million clearly will not have as great an incentive to invest in precautions as the firm that could pay $10 million.[213] Therefore, *the frequency of accidents will increase as a result of the externalization of risk.*[214]

An increase in the number of accidents is not the only socially undesirable consequence of the externalization of risk. The reason that firms undertake certain tasks themselves rather than have them done by independent contractors is to gain economic advantages, and these advantages will be lost if firms hire independent contractors to avoid liability. In our example, the advantages that the firm would lose if it hires an independent contractor might include its

---

[212] There is evidence in the oil industry that is consistent with the view that firms have an incentive to externalize some of their liability risks to less well capitalized independent contractors. For example, after the *Exxon Valdez* oil spill, the Shell Oil Company shifted some responsibility for the transport of oil from its own tanker fleet to vessels owned by independent contractors. *See* William J. Cook, *An easy way out of this mess*, U.S. NEWS & WORLD REP., June 25, 1990; Caleb Solomon & Joann S. Lublin, *Tanker Fire Raises Serious Questions About Liabilities in Oil Spills Off U.S.*, WALL ST. J., June 12, 1990, at A3. Such contractors, who might own just a few supertankers (or even only one), generally are vastly smaller than the major oil companies. *See* Eric Nalder, *Oil Firms Trying to Shield Assets from Liability for Costly Spills*, SEATTLE TIMES, Sept. 26, 1991, at A6 (stating the assets of independent transporters "are about 25 times smaller than the holdings of the oil giants"). One can think of many other industries, including those involved in the transport or disposal of hazardous materials, in which risks can be externalized to independent contractors that are much smaller than the firms that hire them.

[213] For instance, consider a safety device that would reduce the magnitude of harm from an accident from $10 million to $5 million. This device would be of value to the firm if the firm is exposed to liability, since the device would reduce the firm's expected damages by half; but the device would be of no value to the independent contractor, for the contractor is only capable of paying $1 million. Similarly, consider a safety device that reduces the likelihood of a $10 million accident by one percent. This will be worth $100,000 to the firm (= 1% x $10 million), so the firm will pay up to $100,000 for the device, but it will be worth only $10,000 to the independent contractor (= 1% x $1 million), so that the contractor might not buy the device when the firm would have.

[214] In the oil industry, for example, it is plausible that tankers owned by independent contractors are more prone to accidents than tankers owned by large oil companies. *Compare* Cook, *supra* note 212 (describing independent tankers as "clunkers" operating under "lax standards" with "badly trained crews"), *with* Daniel Southerland, *Mobilizing The Fleet: Oil Giant Hopes Emphasis on Tanker Safety Also Will Produce Profits*, WASH. POST, June 23, 1996, at H1 (emphasizing the high safety standards maintained by Mobil Oil's shipping subsidiary).

ability (given its superior knowledge of its own situation) to purchase a type of truck better suited for transport of its particular type of waste and the opportunity to schedule waste disposal more efficiently.[215]

The two problems caused by the externalization of risk -- the increased number of accidents and the loss of economic efficiencies -- are exacerbated by the imposition of punitive damages, for this increases the desire of firms to externalize their risks. Moreover, it should be noted that the judicial tendency to impose higher punitive damages on wealthier firms (a practice we discussed above[216]) has the perverse consequence of increasing the incentive of such firms to externalize risks despite their being more likely to take appropriate precautions (because they are less likely to be judgment proof). Our conclusion, therefore, is that the increased externalization of risk induced by punitive damages argues for conservatism in their imposition.

### O. Encouraging Market Transactions

In some circumstances it is possible for a potential injurer to communicate with a potential victim before causing harm. This is the case, for example, when a firm deliberately infringes on another's copyright or when an individual regularly trespasses on someone's property. If prior communication is possible, a potential injurer could negotiate in advance with the potential victim to purchase the right to engage in the harm-creating conduct, rather than first causing the harm and then paying damages. The firm contemplating a copyright violation could secure a license to use the copyrighted material, or the trespasser could obtain an easement. Obviously, the greater the level of damages that would be imposed on an injurer who causes harm without having purchased the right to engage in the harm-creating conduct, the greater the incentive to purchase the right. In this sense, punitive damages can be said to

---

[215] For a discussion of the advantages of performing tasks within a firm rather than delegating them to independent contractors, see, for example, Ronald H. Coase, *The Nature of the Firm*, 4 ECONOMICA 386 (1937), OLIVER HART, FIRMS, CONTRACTS, AND FINANCIAL STRUCTURE 1-92, and JEAN TIROLE, THE THEORY OF INDUSTRIAL ORGANIZATION 15-60 (1988).

[216] *See* text accompanying notes 116-119.

encourage market transactions. To our knowledge, this effect of imposing punitive damages rarely is mentioned by courts.[217]

In this section, we explain why it may be desirable to impose punitive damages to encourage market transactions. The reason in essence is that inducing potential injurers to bargain may better lead them to take harm into account and may avoid causing parties to engage in wasteful efforts to try to take and protect property. In addition, market exchange may be cheaper than litigation. A qualification to this discussion, however, is that imposing punitive damages when the parties are *not* able to easily bargain may overdeter injurers. Note that the rationale for punitive damages discussed in this section does not presume that a party who causes harm is able to escape liability with positive probability. In other words, the present rationale is independent of the escaping-liability rationale for punitive damages that has been the focus of our Article.[218]

To elaborate on the preceding points, suppose that compensatory damages alone are employed and that they are underestimated. A potential injurer then might cause harm when doing so is socially undesirable -- because the benefit to the injurer might be less than the harm done, but greater than the low estimate of compensatory damages.[219] In general, as we observed in Section II.A, an excessive amount of harm will be caused if damages are too low.

There may be additional undesirable repercussions, similar to those associated with the theft of property, when compensatory damages are underestimated. If injurers can take property from victims without having to pay for the full value of the property, potential injurers will devote effort to identifying and taking such property, and potential victims will

---

[217] The only reference to this effect of which we are aware is in Kemezy v. Peters, 79 F.3d 33, 34 (7th Cir. 1996) (Posner, C.J.) (noting in dictum that "[p]unitive damages are necessary in some cases to make sure that people channel transactions through the market when the costs of voluntary transactions are low").

[218] The point that inducing market transactions may better lead potential injurers to take harm into account was made by Guido Calabresi and A. Douglas Melamed, *Property Rules, Liability Rules, and Inalienability: One View of the Cathedral*, 85 HARV. L. REV. 1089 (1972). This point has been developed by others. *See, e.g.*, Biggar, *supra* note 12; Haddock, McChesney & Spiegel, *supra* note 12; Louis Kaplow & Steven Shavell, *Property Rules Versus Liability Rules: An Economic Analysis*, 109 HARV. L. REV. 713 (1996); Landes & Posner, *supra* note 12.

[219] In these circumstances, a potential victim would have an incentive to pay a potential injurer not to cause harm. Such a payment might not occur, however. For example, if there are many potential injurers, paying one not to cause harm would not forestall others from causing harm. *See generally* Kaplow & Shavell, *supra* note 218.

expend effort to prevent their property from being taken. Copyright violators, for example, will devote resources to copying others' protected material, and copyright owners will take steps to try to keep this from happening. Such efforts are socially wasteful.

The foregoing problems -- an excessive amount of harm and wasteful efforts to take and protect property -- can be avoided if punitive damages are imposed. If such damages are set so that total damages substantially exceed the value of someone's property, a potential taker of that property would be induced to bargain with the owner rather than to take the property -- it would be cheaper to pay an agreed upon price than to pay damages. Consequently, property will be exchanged only if the buyer values it more than the property owner, and the incentive to take and protect property whose value might be underestimated by compensatory damages would be eliminated.

Another possible reason to employ punitive damages to encourage market transactions concerns administrative costs. If compensatory damages are used alone, exchange often will be mediated through the legal system by the bringing of a lawsuit; the cost of exchange then will be the cost of litigation (frequently reduced because of settlement). But if punitive damages are used in addition, exchange will be much more likely to occur through voluntary transactions, whose costs may be much lower than those incurred in litigation.

The arguments that we have now discussed in favor of using punitive damages to promote market exchange obviously do not apply if bargaining between parties is not possible or if there are substantial impediments to bargaining. Suppose, for instance, that a hiker lost in the mountains discovers an unoccupied cabin. The benefit he would obtain from using the cabin and consuming the food in it presumably would exceed the loss borne by the cabin's owner. But because there is no opportunity for the hiker to bargain with the owner, the effect of punitive damages might be to discourage the hiker from using the cabin. Hence, when parties cannot bargain, it may be better just to employ compensatory damages (despite the possibility of errors in estimation); punitive damages would tend to overdeter injurers' conduct. Even if bargaining is feasible, there may be other impediments to efficient exchange -- such as bargaining failures due to strategic behavior -- that also could justify

relying solely on compensatory damages.[220]

The conclusion from this discussion, therefore, is that punitive damages may sometimes have appeal when it is possible for a potential injurer to communicate with a potential victim before causing harm, in order to encourage market transactions. As we noted above, this rationale for punitive damages, when applicable, is independent of the escaping-liability rationale.[221]

---

[220] For example, suppose a seller holds out for a high price and ultimately refuses to sell to a potential buyer who places a much greater value on the item at issue. Then it may be better to set damages equal to harm and allow the "buyer" to take the item and pay damages than to encourage bargaining. This point was emphasized by Calabresi and Melamed, *supra* note 218, at 1106-07.

[221] If there is a probability of escaping liability, then the punitive damages amount that is appropriate for the purpose of encouraging market transactions should itself be inflated according to the multiplier formula that we developed in Section II.B *supra*.

## IV. PUNISHMENT

By the punishment objective we refer to the imposition of appropriate sanctions on blameworthy parties.[222] We equate blameworthiness with the reprehensibility of a party's conduct, that is, with its maliciousness or the extent to which it reflects disregard for the well-being of others.[223] We assume that the punishment objective derives ultimately from the desire of individuals to have blameworthy parties punished (although our essential conclusions do not depend on this assumption[224]).[225]

When the defendant is an individual, the connection between the imposition of punitive damages and the accomplishment of the punishment objective is conceptually straightforward: if, after assessing the blameworthiness of an individual's act, appropriate punitive damages are levied, the punishment objective is achieved.[226]

However, when the defendant is a firm, the relationship between punitive damages and the punishment objective is more complex. In this regard, we will develop three points. The first is that there are different ways of viewing the objective of punishment: the goal may be to punish firms as *entities*, that is, independently of whether blameworthy individuals within the

---

[222] *See generally* GEORGE FLETCHER, RETHINKING CRIMINAL LAW § 6.3.2, at 417 (1978) ("[T]he offender is duty-bound to suffer punishment, for his offense creates an imbalance of benefits and burdens in the society as a whole." (citing H. MORRIS, ON GUILT AND INNOCENCE 34-36 (1976)); SIR WALTER MOBERLY, THE ETHICS OF PUNISHMENT 95 (1968) (Retributive theory essentially claims "that punishment should serve both to express and to deepen the horror with which certain types of action ought to be regarded."); HERBERT L. PACKER, THE LIMITS OF THE CRIMINAL SANCTION 37-39 (1968) ("The retributive view [of punishment] rests on the idea that it is right for the wicked to be punished: because man is responsible for his actions, he ought to receive his just deserts."); C. L. TEN, CRIME, GUILT, AND PUNISHMENT: A PHILOSOPHICAL INTRODUCTION 2 (1987) ("Punishment involves the infliction of some unpleasantness on the offender . . . . made to express disapproval or condemnation of the offender's conduct which is a breach of what is regarded as a desirable and obligatory standard of conduct.").

[223] It is not necessary for our purposes to settle on a more refined definition of blameworthiness or culpability.

[224] It will be evident that, for the most part, our arguments also will hold if the view is that there is a correct level of punishment that is philosophically motivated, not derived from the level that raises individual welfare.

[225] The punishment objective is, to be more precise, to raise social welfare by punishing culpable parties. Punishing such parties raises social welfare by increasing the utility of individuals who desire punishment.

[226] A qualification to this statement concerns liability insurance. If the punished party is insured, the degree to which he is punished depends on the extent to which his coverage is incomplete (due, for example, to deductibles or coinsurance) and the possibility that his premiums will rise in the future.

firms are penalized; or the goal may be to punish firms only as a means of punishing culpable *individuals* in the firms. Our second point is that there are reasons why the imposition of punitive damages on firms may not lead to the punishment of blameworthy individuals within them; thus, the goal of punishing blameworthy employees may not be well promoted by imposing punitive damages on firms. The final point is that the imposition of punitive damages on firms often penalizes individuals who are unlikely to be considered culpable, namely shareholders and customers. We conclude that, to the extent that the goal is to punish culpable individuals within firms, and not firms as entities, the utility of punitive damages in achieving the punishment objective is significantly attenuated.

Consider the possibility that the punishment objective might be furthered because people obtain satisfaction directly from the punishment of a blameworthy firm as an organization, without reference to whether anyone within the firm behaved inappropriately or is punished.[227] We find this conception of the punishment goal unappealing both because it requires a definition of blameworthiness of firms that is divorced from human behavior, and because it necessitates believing that people would, after reflecting about the matter, want to impose a penalty on what ultimately is an artificial legal construct. The notion that individuals would want to punish firms *per se* strikes us as not entirely different from the idea that individuals would want to punish inanimate objects for causing harm (such as trees that fall on people).

Notwithstanding these reservations, it is possible that individuals do want to personify firms and punish them as entities, and the reader can make up his or her mind about the

---

[227] A number of authors have discussed the punishment rationale in relation to corporations as entities. *See, e.g.*, Albert W. Alschuler, Comment, *Ancient Law and the Punishment of Corporations: Of Frankpledge and Deodand*, 71 B.U. L. REV. 307, 312 (1991) (stating that corporate punishment is a necessary vent for societal anger); *Developments in the Law–Corporate Crime: Regulating Corporate Behavior Through Criminal Sanctions*, 92 HARV. L. REV. 1227, 1237 (1979) ("Even though deterrence clearly plays a critical role in the justification of corporate criminal sanctions, the argument that retribution cannot be involved is unconvincing."); V. S. Khanna, *Corporate Criminal Liability: What Purpose Does It Serve?*, 109 HARV. L. REV. 1477, 1494 n.91 (describing "the notion that society has a retributive need so great that it must punish nonhuman entities and label them criminal" as "implausible"); Christopher D. Stone, *The Place of Enterprise Liability in the Control of Corporate Conduct*, 90 YALE L.J. 1, 27 (1980) ("Corporate penalties often impose losses in circumstances when no one appears blameworthy . . . . True, we allow some of these same innocents to suffer . . . when corporate agents wrongfully break a contract or commit a tort . . . . But it seems one thing to make a blameless investor help absorb ordinary damages . . . and quite another thing to reduce his investments further by imposing a penalty.").

importance of this way of defining the punishment objective. To the extent that it is important, the imposition of punitive damages on a blameworthy firm directly promotes the punishment objective, much the same as it does when the defendant is a culpable individual.[228]

Now consider the alternative reason for punishing firms -- to punish blameworthy individuals within them. Supposing that this is the purpose of punishment, we turn to our second point, about the extent to which the imposition of punitive damages on firms will actually result in the punishment of blameworthy employees. Because firms clearly have an interest in discouraging culpable conduct by their employees that could give rise to punitive damages, they can be expected to seek to control such conduct through the use of internal sanctions such as demotion or dismissal. However, two considerations suggest that the imposition of punitive damages on firms will have a smaller effect on the punishment of blameworthy employees than might at first be supposed.

First, culpable employees may not be punished by firms because the firms may have difficulty identifying them. Such individuals may be able to obfuscate their role in decisionmaking or conceal their behavior in a variety of ways. For example, an employee responsible for checking a safety valve on a tank storing dangerous chemicals that subsequently explodes due to a defective valve may claim that he performed the inspection even if he did not, and may place a false entry in his record book attesting to the inspection. Or a manager whose judgement is impaired by alcohol and who gives oral instructions to a subordinate that are wrong and that lead to an accident, may deny ever having told the subordinate to do what the subordinate did.

Second, even if culpable individuals within a firm can be identified and punished by the firm, imposing punitive damages on firms often will have little or no *marginal effect* on their punishment. That is, the internal sanction imposed on such employees may not be much (if at all) greater as a result of the firm's bearing both punitive and compensatory damages than if the firm had borne compensatory damages alone. When a firm incurs high compensatory

---

[228] Because the punishment objective is, according to our approach, derived from the desire of individuals to punish culpable parties, the importance of the view that corporations *per se* should be punished is an empirical matter -- dependent on how many individuals hold this position and how strongly.

damages because of the blameworthy conduct of an identifiable employee, it may want to levy whatever sanctions on him that it can; imposing punitive damages on the firm then would not result in additional punishment of the employee.[229]

The preceding discussion presupposed that there exist culpable employees in the firm to punish. But in some situations there may not be any. If there is a significant delay between misconduct and the manifestation of harm and litigation (as was the case, for instance, in connection with the use of asbestos in products), blameworthy individuals may have changed jobs, retired, or died.[230] Also, because decisions in firms often are made by many individuals, it may be that no one individual has the requisite knowledge of risk and of the consequences of his behavior to be considered culpable. One person may decide to put a toxic liquid in a storage tank believing that the tank can never leak; and another person may leave the tank in a state in which a leak can occur, thinking that the liquid in the tank is not toxic, so that a leak would not cause harm. Here, each decision considered by itself may not be blameworthy because, by hypothesis, each person believes that what he is doing does not create a risk of a harmful accident.[231]

Let us now consider the third point, concerning how imposing punitive damages on firms often penalizes the firms' shareholders and customers. Shareholders, as residual claimants on a firm's profits, obviously will be made worse off when punitive damages are

---

[229] It may be worth elaborating on this point. If employees are risk neutral (which we assume here for simplicity), a firm would seek to make an employee pay for any damages he caused the firm to bear; this would make the employee's incentives to prevent harm correct from the firm's perspective. Thus, if damages are $1,000, the employee would pay this amount to the firm. But since an employee's assets are limited, the firm's ability to punish an employee will be exhausted as soon as the judgment against the firm exceeds the employee's assets. Thus, if the employee's assets are $10,000, there will be no marginal effect of higher damages once the total damages exceed $10,000. (The situation just described ignores other possible responses of the firm to higher damages. For example, firms might increase their efforts to detect employee misconduct and thereby increase the expected punishment of employees. But our basic point would still apply.)

[230] However, individuals who have changed jobs or retired might be subject to punishment by their former employer. The employer might be able to sue them for acts done when employed.

[231] Of course, some other person within the firm may have been responsible for directing the flow of information, but that person also may not have acted culpably.

imposed on a firm.[232] Indeed, they usually can be expected to bear a major fraction of the burden of punitive damages.[233] Given that shareholders are punished by punitive damages, the question whether they are blameworthy must be considered. If a shareholder owns a significant fraction of a firm's stock, participated actively in the firm's decisions and acted egregiously, then his position would be much like that of a blameworthy employee with decisionmaking power; each would be culpable. But if a shareholder owns a minuscule fraction of the stock of the firm, and was a passive investor with no direct involvement in the firm's decisionmaking processes, then his degree of blameworthiness is small, if there is any blameworthiness at all.[234]

A firm's customers also will suffer from the imposition of punitive damages on the firm if such damages cause the prices of the firm's products or services to rise. Firms may regard punitive damages as an additional cost of doing business -- a cost that, with a positive probability, will be borne by them beyond their ordinary costs. To cover the added cost of punitive damages, firms will have to raise their prices, which will cause the welfare of their customers to decline.[235] It seems clear, however, that customers would not ordinarily be considered blameworthy, because they do not exert direct control over the actions of firms that pose risks to other persons. Consequently, to the extent that customers pay higher prices as a

---

[232] There are other stakeholders of the firm who also will suffer, such as unsecured creditors who are less likely to be repaid if the firm's assets are diminished due to the payment of punitive damages.

[233] This is because blameworthy employees themselves generally will bear only a small part of a punitive damages judgment (due to their limited assets), leaving shareholders and customers to absorb the remainder. Under certain circumstances, moreover, customers will not bear much of the burden of punitive damages. *See* note 235 *infra*.

[234] If one does believe that each shareholder is slightly blameworthy, then the fact that each bears a small portion of punitive damages might be thought to be desirable.

[235] If, however, a particular firm bears punitive damages for a reason not generally applicable to other firms in its industry, these other firms would not have a reason to raise their prices. Consequently, the firm paying the punitive damages would not be able to raise the price of its products (consumers would purchase from the other firms).

result of the imposition of punitive damages on firms, innocent parties have been penalized.[236]

We can summarize our discussion of the punishment of firms as follows. The idea that a firm should be punished *per se* -- without reference to the punishment of individuals within it -- is a possible view but one that we find problematic. Another view is that the punishment goal is promoted *only* by punishing blameworthy individuals within firms. We have explained, however, why the imposition of punitive damages on firms often may not result in the punishment (or any *additional* punishment) of blameworthy employees, so the use of such damages might not advance the punishment goal very much. Moreover, punitive damages often will penalize shareholders and customers, parties who are not likely to be considered blameworthy. This adverse consequence of punitive damages must be weighed against the beneficial effects of such damages in furthering the punishment goal.

Having now addressed punitive damages and punishment in general terms, we want to consider briefly how the reprehensibility of the defendant's conduct and the wealth of the defendant should influence punitive damages with respect to the punishment objective.[237] Regarding reprehensibility, we merely observe that the punishment objective will, by definition, be met if sanctions are imposed on those who have acted reprehensibly. Hence, determination of the reprehensibility of the defendant's conduct is intrinsic to satisfaction of the punishment objective, and the law's focus on reprehensibility obviously makes sense given this objective. In the case of firms, however, the connection between reprehensibility and punishment may be attenuated for reasons discussed above -- the imposition of punitive damages on a firm may not result in the punishment of individuals within the firm who acted reprehensibly.

---

[236] For a complementary discussion of the effects of corporate sanctions on a firm's shareholders, bondholders, employees, and customers, see John C. Coffee, Jr., *"No Soul to Damn: No Body to Kick": An Unscandalized Inquiry into the Problem of Corporate Punishment*, 79 MICH. L. REV. 3, 401-02 (1981) (observing that "the costs of [corporate] deterrence tend to spill over onto parties who cannot be characterized as culpable.")

[237] We will not, however, re-examine the other topics in Section III in relation to punishment. For the most part, what can be said about these topics is clear. Consider, for example, the issue of the tax deductibility of punitive damages. Allowing punitive damages to be deductible would reduce their sting, and so might be undesirable in terms of accomplishing the punishment objective. Likewise, allowing liability insurance coverage against punitive damages also might be undesirable from the perspective of punishment.

Concerning defendants' wealth and the appropriate level of damages from the perspective of punishment, first consider the situation when defendants are individuals. In this context, the common belief that punitive damages should be higher the higher is the defendant's wealth can be justified. The punishment goal is furthered if a proper punishment is imposed on a culpable individual, which we interpret to mean reducing the individual's utility by particular amount. To accomplish this, it generally is necessary to assess a higher penalty if the individual is wealthy than if he is poor, because money is worth less to him if he is wealthy.[238]

When the defendant is a firm, the relevance of the defendant's wealth may depend on whether the punishment goal is viewed in terms of punishing the firm as an entity or punishing culpable individuals within the firm. Under the first view, the firm's wealth might be thought to be relevant to the proper level of damages for punishment purposes.[239] Under the second view, however, the firm's wealth generally would not be relevant: the level of damages needed to induce a firm to punish its culpable employees ordinarily would not depend on its wealth. A $100 million firm and a $10 million firm would both be expected to impose the same sanction on an employee for misconduct that resulted in a punitive damages award of a given amount. The reason is that, as we have said, rational firms will develop a policy of punishing employee misbehavior in order to lower their liability expenses. This policy should depend on variables other than the firm's wealth -- notably, the damages that the firm will bear as a result of employee misbehavior.[240] To the extent that the internal sanctions that firms impose on culpable employees do not depend on the firm's wealth, the punishment objective will not be advanced by making punitive damages depend on its wealth.

---

[238] We mentioned this point in note 120 *supra*. How much the penalty must rise will be determined by the rate at which the marginal utility of money declines with wealth for the individual. In particular, there is no reason to believe that the proper penalty for purposes of punishment would be proportional to wealth.

[239] We say "might" because the first view is not well articulated, and therefore it is unclear what this view would imply about the proper relationship between the level of punitive damages and a corporation's wealth.

[240] The reasoning behind this statement is essentially that used in Section III.B *supra* when we explained why a firm's wealth will not affect its incentive to invest in safety precautions.

## V. CONCLUSION

In this Article we have discussed the two fundamental purposes of punitive damages -- deterrence and punishment -- and have come to conclusions regarding each objective that we now briefly review.

Our central conclusion about punitive damages and deterrence was conceptually simple. Punitive damages should be imposed when deterrence otherwise would be inadequate due to the possibility that injurers would escape liability. In particular, punitive damages should be set at a level such that the expected damages of defendants equal the harm they have caused, for then their damage payments will, in an average sense, equal the harm. This implies a simple formula for calculating punitive damages, according to which harm is multiplied by a factor reflecting the likelihood of escaping liability. If punitive damages are calculated according to this multiplier formula, precautions will tend to be optimal -- neither inadequate nor excessive -- as will product prices and the incentive to participate in risky activities. These conclusions about punitive damages, and the importance of the role of the defendant's chance of escaping liability, flow from the standard and well-accepted theory of deterrence. We also discussed a deterrence rationale for punitive damages that is not based on the possibility of escaping liability: that punitive damages may be needed to offset the socially illicit utility that individuals obtain from committing malicious acts. This rationale, as we noted, does not apply to firms.

The theory of deterrence not only yields a multiplier formula for computing punitive damages, it also provides guidance regarding a range of important doctrinal and policy-related issues concerning punitive damages. Notably, we discussed the point that the reprehensibility of a party's conduct generally should not be a factor in the assessment of punitive damages (except in the case of an individual's malicious act), as well as the point that the wealth of a defendant usually should not influence punitive damages (subject to the same exception).

A corollary of our analysis is that imposing punitive damages when they are not justified on deterrence grounds generally has socially detrimental consequences. These can take the form of excessive precautionary measures and inappropriate discouragement of participation in socially beneficial activities. In the case of firms, the latter effect may

manifest itself as undesirably high prices and the withdrawal of products from markets.

With respect to the punishment objective, we observed that the connection between punitive damages and punishment is relatively straightforward if the defendant is an individual or if the defendant is a firm and the goal is to punish firms as entities (although we found this latter goal problematic). We came to a different conclusion, however, when the defendant is a firm and the objective is to punish culpable employees. Because the imposition of punitive damages on firms may not result in the punishment of blameworthy employees, but often will penalize shareholders and customers -- parties who are not likely to be considered blameworthy -- the ability of punitive damages to advance the punishment goal in the case of firms is limited.

We did not comment on how the level of punitive damages should be determined when the objectives of deterrence and punishment have different implications for the proper measure of punitive damages. It is evident that the best level of punitive damages should be a *compromise* between the levels that are optimal when each objective is considered independently.[241] (The quantities of punitive damages that are separately optimal with respect to the two objectives should not be added to each other.[242]) The weights to be used in the determination of the compromise will reflect the relative importance accorded to the goals of punishment and of deterrence.

Whatever are the weights that policymakers, judges, or juries place on these two goals, we hope that the conceptual framework developed in this Article will be of aid to them in determining the appropriate amount of punitive damages.

---

[241] For example, suppose that the level of damages that is best with respect to deterrence is $1 million and that the amount that is best in terms of punishment is $2 million. The optimal amount, taking account of both objectives, must be between $1 and $2 million. As damages are increased from $1 million to $2 million, deterrence worsens due to overdeterrence, but punishment is better promoted. It is optimal to stop raising damages when the marginal social loss from overdeterrence begins to outweigh the marginal social gain from better punishment. (To complete the explanation, observe that optimal damages cannot be less than $1 million or more than $2 million: were damages less than $1 million, that would be worse than damages of $1 million with respect to both the deterrence and the punishment goals; and were damages in excess of $2 million, that would be worse than damages of $2 million with respect to both goals.) For further discussion regarding the choice of the level of punitive damages that best balances the deterrence and punishment goals, see Diamond, Punishment and Efficiency, *supra* note 12.

[242] In essence, this is because the punitive damages amount that is proper for deterrence purposes also punishes, and the punitive damages amount that is proper for punishment purposes also deters.

## APPENDIX: MODEL JURY INSTRUCTIONS

The following model jury instructions encapsulate many of the conclusions of our Article. Three sets of instructions are presented: for individuals who have not committed malicious acts; for individuals who have committed malicious acts; and for firms. With regard to the instructions for firms, we have assumed that the goal of punishment is to penalize blameworthy employees, not to punish firms as entities. (It would be easy to modify the instructions to reflect a different assumption.)

* * *

### For Individuals Who Have Not Committed Malicious Acts

These instructions apply to defendants who have not committed malicious acts. An act is malicious only if it was done for the *purpose* of causing harm.

In considering the imposition of punitive damages on the defendant, you should determine three dollar amounts: (I) an amount to accomplish deterrence; (II) an amount to accomplish punishment; (III) a final amount -- your punitive damages award -- in between the first two amounts.

### I. Deterrence

A. Punitive damages fulfill the deterrence objective to the extent that they serve as a message and warning to the defendant and to other similarly situated individuals to take appropriate steps to prevent harm in the future. But punitive damages will not fulfill the deterrence objective if they cause individuals to take wasteful steps to prevent harm, or if they cause individuals to refrain from engaging in socially desirable activities.

B. Your principal task is to estimate the likelihood that the defendant might have escaped having to pay for the harm for which he or she should be responsible. Thus, for example, if the harm was noticeable and likely to lead to a lawsuit, your estimate of the likelihood of escaping liability would be relatively low. But if the harm might not have been attributed to the defendant, or if the defendant tried to conceal his or her harmful conduct, your estimate of the likelihood of escaping liability would be relatively high.

C. You should use the Table below to determine the punitive damages multiplier that corresponds to your estimated probability of escaping liability. Then multiply the compensatory damages amount by your punitive damages multiplier. The resulting number is the *base punitive damages amount*.

D. The base punitive damages amount *should be lowered* if the defendant has paid other private judgments or settlements, or public penalties, for the harm at issue in the present

case. If the defendant has made such payments, the base punitive damages amount should be lowered by the amount of these payments.

E. The base punitive damages amount *should not be adjusted* because of any of the following considerations:

(1) reprehensibility of the defendant's conduct;

(2) net worth or income of the defendant;

(3) potential harm, that is, the harm that might have been caused by the defendant's conduct;

(4) gain or profit that the defendant might have obtained from his or her harmful conduct;

(5) litigation costs borne by the plaintiff;

(6) components of harm that you did not include in compensatory damages;

(7) whether the harm included personal injury.

**II. Punishment**

A. Punitive damages fulfill the punishment objective to the extent that they penalize blameworthy defendants for reprehensible behavior. You should determine the amount of punitive damages that you believe will accomplish proper punishment.

B. In considering punishment, keep in mind that the defendant's payment of compensatory damages already punishes the defendant to some extent. The amount of punitive damages that you believe will accomplish proper punishment should be what you think must be added to compensatory damages to accomplish the punishment objective, if any additional damages are necessary.

**III. Determination of Punitive Damages**

Punitive damages should be an amount *in between* the amount that you found appropriate for deterrence purposes and the amount that you found appropriate for punishment purposes. If you attach greater importance to the deterrence objective, punitive damages should be closer to the amount that you found that best promotes deterrence. If you attach greater importance to the punishment objective, punitive damages should be closer to the amount that you found that best promotes punishment.

* * *

**For Individuals Who Have Committed Malicious Acts**

These instructions apply to defendants who have committed malicious acts. An act is malicious if it was done for the *purpose* of causing harm.

In considering the imposition of punitive damages on the defendant, you should determine three dollar amounts: (I) an amount to accomplish deterrence; (II) an amount to accomplish punishment; (III) a final amount -- your punitive damages award -- in between the first two amounts.

**I. Deterrence**

A. Punitive damages fulfill the deterrence objective to the extent that they serve as a message and warning to the defendant and to other similarly situated individuals not to commit malicious acts in the future. But punitive damages will not fulfill the deterrence objective if they cause individuals to take wasteful steps to avoid possible liability, or if they cause individuals to refrain from engaging in socially desirable activities.

B. Your principal task is to estimate the likelihood that the defendant might have escaped having to pay for the harm for which he or she should be responsible. Thus, for example, if the harm was noticeable and likely to lead to a lawsuit, your estimate of the likelihood of escaping liability would be relatively low. But if the harm might not have been attributed to the defendant, or if the defendant tried to conceal his or her harmful conduct, your estimate of the likelihood of escaping liability would be relatively high.

C. You should use the Table below to determine the punitive damages multiplier that corresponds to your estimated probability of escaping liability. Then determine the amount that you believe is equivalent to the gain that the defendant obtained from his or her conduct. Multiply this amount by your punitive damages multiplier. The resulting number is the *base punitive damages amount*.

D. The base punitive damages amount *should be lowered* if the defendant has paid other private judgments or settlements, or public penalties, for the harm at issue in the present case. If the defendant has made such payments, the base punitive damages amount should be lowered by the amount of these payments.

E. The base punitive damages amount *should not be adjusted* because of any of the following considerations:

(1) potential harm, that is, the harm that might have been caused by the defendant's conduct;

(2) litigation costs borne by the plaintiff;

(3) components of harm that you did not include in compensatory damages.

**II. Punishment**

A. Punitive damages fulfill the punishment objective to the extent that they penalize blameworthy defendants for reprehensible behavior. You should determine the amount of punitive damages that you believe will accomplish proper punishment.

B. In considering punishment, keep in mind that the defendant's payment of compensatory damages already punishes the defendant to some extent. The amount of punitive damages that you believe will accomplish proper punishment should be what you think must be added to compensatory damages to accomplish the punishment objective, if any additional damages are necessary.

**III. Determination of Punitive Damages**

Punitive damages should be an amount *in between* the amount that you found appropriate for deterrence purposes and the amount that you found appropriate for punishment purposes. If you attach greater importance to the deterrence objective, punitive damages should be closer to the amount that you found that best promotes deterrence. If you attach greater importance to the punishment objective, punitive damages should be closer to the amount that you found that best promotes punishment.

* * *

**For Firms**

In considering the imposition of punitive damages on the defendant, you should determine three dollar amounts: (I) an amount to accomplish deterrence; (II) an amount to accomplish punishment; (III) a final amount -- your punitive damages award -- in between the first two amounts.

**I. Deterrence**

A. Punitive damages fulfill the deterrence objective to the extent that they serve as a message and warning to the defendant and to other similarly situated firms to take appropriate steps to prevent harm in the future. But punitive damages will not fulfill the deterrence objective if they cause firms to take wasteful steps to prevent harm; or if they cause the prices of products and services to rise excessively; or if they cause firms to withdraw socially valuable products or services from the market.

B. To achieve the deterrence objective, your principal task is to estimate the likelihood that the defendant might have escaped having to pay for the harm for which it should be responsible. Thus, for example, if the harm was noticeable and likely to lead to a lawsuit, your estimate of the likelihood of escaping liability would be relatively low. But if the harm might not have been attributed to the defendant, or if the defendant tried to conceal its harmful conduct, your estimate of the likelihood of escaping liability would be relatively high.

C. You should use the Table below to determine the punitive damages multiplier that corresponds to your estimated probability of escaping liability. Then multiply the compensatory damages amount by your punitive damages multiplier. The resulting number is the *base punitive damages amount*.

D. The base punitive damages amount *should be lowered* if the defendant has paid other private judgments or settlements, or public penalties, for the harm at issue in the present case. If the defendant has made such payments, the base punitive damages amount should be lowered by the amount of these payments.

E. The base punitive damages amount also *may be lowered* if the plaintiff was a customer of the defendant. If the plaintiff was a customer and you believe that customers are, or will become, aware of accidents of the type at issue in this case, the base punitive damages amount should be lowered. The more knowledgeable customers are, the more the base punitive damages amount should be lowered.

F. The base punitive damages amount *should not be adjusted* because of any of the following considerations:

(1) reprehensibility of the defendant's conduct;

(2) net worth, revenues, or profits of the defendant;

(3) potential harm, that is, the harm that might have been caused by the defendant's conduct;

(4) gain or profit that the defendant might have obtained from its harmful conduct;

(5) litigation costs borne by the plaintiff;

(6) components of harm that you did not include in compensatory damages;

(7) whether the harm included personal injury.

## II. Punishment

A. Punitive damages fulfill the punishment objective to the extent that they cause defendants to penalize their *blameworthy employees* for reprehensible behavior.

B. In considering punishment, you should keep in mind that the defendant's payment of compensatory damages already may lead to the punishment of blameworthy employees to some extent.

C. In considering how well the imposition of punitive damages will fulfill the punishment objective, you should also bear the following in mind:

(1) the extent to which you believe blameworthy employees can be identified and penalized by the defendant. The easier this is, the higher should be the level of punitive damages.

(2) the extent to which you believe that innocent parties will suffer as a result of the imposition of punitive damages on the defendant; such parties might include shareholders as well as customers, who may have to pay higher prices for the defendant's products or services. The more likely it is that innocent parties will be punished, the lower should be the level of punitive damages.

D. In the light of these considerations, you should determine the amount of punitive damages that you believe will accomplish proper punishment.

## III. Determination of Punitive Damages

Punitive damages should be an amount *in between* the amount that you found appropriate for deterrence purposes and the amount that you found appropriate for punishment purposes. If you attach greater importance to the deterrence objective, punitive damages should be closer to the amount that you found that best promotes deterrence. If you attach greater importance to the punishment objective, punitive damages should be closer to the amount that you found that best promotes punishment.

TABLE

| Probability of Escaping Liability | Punitive Damages Multiplier |
|---|---|
| 0% | 0 |
| 10% | .11 |
| 20% | .25 |
| 30% | .43 |
| 40% | .67 |
| 50% | 1.00 |
| 60% | 1.50 |
| 70% | 2.33 |
| 80% | 4.00 |
| 90% | 9.00 |

*Note*: The multipliers in the table are derived as follows. Let $P$ be the probability of being found liable; thus, the probability of escaping liability is $1 - P$. The multiplier then equals $(1 - P)/P$.