1991

Litigation Costs and the Economic Theory of Tort Law

Keith N. Hylton
Boston University School of Law

Litigation Costs and the Economic Theory of Tort Law

KEITH N. HYLTON*

I. INTRODUCTION

The economic theory of tort law has developed along two lines.[1] The first and more traditional is positive theory, which justifies tort doctrine.[2] The second is normative theory, which usually criticizes the operational efficiency of the tort system,[3] and is the focus of this

* Assistant Professor, Northwestern University School of Law; Research Fellow, American Bar Foundation. I thank Ron Allen, Ian Ayres, Richard Craswell, Mayer Freed, Mark Grady, Maria O'Brien Hylton, Richard Posner, Marshall Shapo, Steve Shavell, Peter Siegelman, and David Van Zandt for comments on this paper. Richard Craswell and Steve Shavell deserve an additional thanks for helpful comments on research leading to this paper. I have benefited from comments received at presentations at Northwestern University and the University of Chicago.

1. The two lines are well known in the literature. *See, e.g.*, Mark F. Grady, *A New Positive Economic Theory of Negligence*, 92 YALE L.J. 799, 799 (1983).

2. Much of what has come to be known as the positive economic theory of tort law is due to the joint work of William Landes and Richard Posner. *See* WILLIAM M. LANDES & RICHARD A. POSNER, THE ECONOMIC STRUCTURE OF TORT LAW 1 (1987). The positive economic theory of torts can be traced to Lecture III of OLIVER W. HOLMES, THE COMMON LAW (1881). A broad interpretation of positive theory would also include other works which arrive at favorable conclusions with respect to the economic efficiency of tort law. *E.g.*, STEVEN SHAVELL, ECONOMIC ANALYSIS OF ACCIDENT LAW (1987); John P. Brown, *Toward an Economic Theory of Liability*, 2 J. LEGAL STUD. 323 (1973).

3. The most important example of normative theory is GUIDO CALABRESI, THE COSTS OF ACCIDENTS: A LEGAL AND ECONOMIC ANALYSIS (1970). See also GORDON TULLOCK,

Article. This Article argues that once the dynamics of litigation are properly taken into account, all bets are off on the economic efficiency of tort law. The simple fact that litigation is a costly enterprise provides a rich source of inefficiencies with which the tort system must grapple.[4]

This Article addresses the operational efficiency of a tort system in which litigation is costly. While it may seem obvious that no system of accident law could be operationally efficient in all respects, this Article takes a broader and more pragmatic perspective. This critique does not seek simply to prove that the tort system does not solve all of the economic problems of accidents. Rather, it identifies important areas in which the system operates inefficiently, and suggests ways to change the doctrine or related rules in order to push the system in the direction of operational efficiency.

Three propositions summarize the implications of costly litigation for the economic efficiency of the tort system. First, the private incentive to bring suit generally differs from the social incentive to bring suit.[5] Second, setting damages at a level that simply compensates the victim is not necessarily optimal; indeed, optimal damages will often be higher or lower.[6] Third, and central to the argument of this paper, tort liability generally will not lead actors to exercise socially optimal precaution. Instead, they will exercise too little precaution.[7]

Two observations explain why the costliness of litigation throws

---

TRIALS ON TRIAL: THE PURE THEORY OF LEGAL PROCEDURE 24-48 (1980), which anticipates some of the arguments of this paper.

4. The focus on litigation costs is the major difference between the argument presented in this paper and earlier writings which argue that the tort system is economically inefficient. *See* CALABRESI, *supra* note 3, at 239-87; TULLOCK, *supra* note 3, at 13-69. The thesis of this Article is different from the argument that the tort system is inefficient because it does not work in practice as well as it works in theory. *See* Stephen D. Sugarman, *Doing Away with Tort Law*, 73 CAL. L. REV. 555 (1985). This Article argues that even in theory the system is operationally inefficient. For example, once institutional features such as the costliness of litigation are properly taken into account, even under the ideal assumptions of the traditional models of deterrence, the tort system can be shown to be inefficient in several important respects.

5. Steven Shavell, *The Social Versus the Private Incentive to Bring Suit in a Costly Legal System*, 11 J. LEGAL STUD. 333, 333-34 (1982).

6. A. Mitchell Polinsky & Daniel L. Rubinfeld, *The Welfare Implications of Costly Litigation for the Level of Liability*, 17 J. LEGAL STUD. 151, 152 (1988). See also Keith N. Hylton, *The Influence of Litigation Costs on Deterrence Under Strict Liability and Under Negligence*, 10 INT'L REV. L. & ECON. 161 app. B. at 170-71 (1990), showing that an upper bound can be placed on optimal damages. The upper bound is discussed in the text accompanying notes 70-72.

7. Hylton, *supra* note 6, at 163; Keith N. Hylton, *Costly Litigation and Legal Error under Negligence*, 6 J.L. ECON. & ORG. 433, 443 (1990).

the economics of tort law off the tracks of positive theory. First, because litigation is costly, not every victim will find it profitable to bring suit.[8] Some victims will bear their losses without seeking compensation through the tort system.[9] The victims who choose to bring suit will do so on the basis of an arbitrary standard: whether the anticipated damage award exceeds the cost of litigating. Second, because litigation is costly, the probability of winning a lawsuit becomes an important consideration in the decision to bring suit. This, however, implies that the deterrence properties of the tort system will depend on litigation prospects. This is true in the real world, but is ignored in positive theory.

Because it lies at the heart of this argument, it may help to explain in simple terms why costly litigation implies that the tort system fails to compel actors to exercise socially optimal precaution. For example, in a strict liability regime, potential injurers have incentives to exercise socially optimal precaution if liability effectively internalizes all of the social costs of failing to take care. However, in a system requiring litigation in order to internalize social costs, two types of costs will not be internalized: (1) the losses suffered by victims who choose not to bring suit because the cost of bringing suit exceeds the anticipated damage award, and (2) the litigation costs of victims who do bring suit. Strict liability, therefore, fails to provide incentives to exercise optimal precaution.

Under negligence, because the cost of litigation is just as much a cost that results from failing to take care as the loss suffered by an accident victim, an optimal due care standard would require a poten-

---

8. There seems to be a great deal of evidence in support of the claim that not every potential plaintiff sues. *See, e.g.*, Richard L. Abel, *The Real Tort Crisis—Too Few Claims*, 48 OHIO ST. L.J. 443, 448-52 (1987) (reviewing studies and anecdotal evidence on the litigiousness of tort victims). However, there does not appear to be any empirical study that examines the extent to which this can be explained by the costliness of litigation.

Contingent fee arrangements have no effect on the claim that not every victim will find it profitable to bring suit. Whether the victim pays the lawyer up front or gives the lawyer part of the damage award, suit will be brought only if the anticipated damage award exceeds the cost of litigating. For example, suppose the anticipated damage award is $10,000 and the cost of bringing suit is $11,000. Under a fee arrangement that takes the lawyer's compensation out of the damage award, suit obviously will seem to be profitable from the victim's perspective. But suit will not be profitable to the lawyer, and therefore no profit-oriented attorney will bring the claim.

9. This behavior results in considerable externalization of losses under the tort system that prevents the tort system from providing adequate incentives to take care. The problem of externalization under the tort system was the central source of inefficiency identified in CALABRESI, *supra* note 3, at 28. However, Calabresi attributes externalization to insufficient subcategorization for insurance pricing purposes, the transfer of accident costs to groups other than responsible parties, and inadequate information. *Id.* at 144-145. Calabresi's work does not analyze the influence of litigation costs as a source of externalization.

tial injurer to exercise a higher level of precaution than is required by the Hand formula.[10] Nevertheless, negligence liability is incapable of compelling potential injurers to exercise the optimal level of precaution. In order to provide incentives for a potential injurer to exercise a level of precaution beyond that required by the Hand formula, the losses suffered by victims should be shifted to the injurer. But this cannot happen under traditional negligence rules. The losses that are shifted to injurers under the Hand formula are the losses suffered by victims who choose to bring suit—a fraction of the number of relevant victims—discounted by the probability that damages will be awarded because the court errs in applying the negligence test.[11] Since this is only a fraction of the external losses, the negligence rule fails to provide incentives for injurers to exercise optimal precaution.[12]

One might argue that if litigation costs are relatively small, then at best this Article presents an interesting theoretical sidelight. However, litigation costs are not small. A tort victim's cost of litigating consumes roughly thirty percent of the average damage award.[13]

This argument is not entirely pessimistic. Discovering a source

10. *See* Hylton, *supra* note 7, at 445. This Article uses the traditional test for negligence stated by Judge Learned Hand in United States v. Carroll Towing Co., 159 F.2d 169 (2d Cir. 1947). Under the Hand formula, negligence is a failure to take care when the cost of care is less than the probability of the accident multiplied by the loss if the accident occurs. Using the language of *Carroll Towing*, commentators sometimes describe the test as the "BPL" criterion. If B is the cost of taking care, P is the probability of the accident, and L is the loss if the accident occurs, an injurer should take care as long as $B < PL$. This Article proposes that if C is the total expected cost of litigating (plaintiff's plus defendant's costs), then the correct criterion would require the injurer to take care whenever $B < P(L+C)$.

11. This assumes that the injurer complied with the due-care standard. Therefore, damages would be awarded only if the court errs in applying the negligence test.

12. In the section examining the efficiency properties of the negligence rule, this Article does not discuss comparative negligence regimes. However, the criticisms made of the negligence rule extend in a straightforward way to any regime, such as comparative negligence, in which a finding of negligence is necessary. *See* David Haddock & Christopher Curran, *An Economic Theory of Comparative Negligence*, 14 J. LEGAL STUD. 49, 50, 59, 66 (1985) (noting that the efficiency properties of the negligence rule extend to comparative negligence regimes); *see also* Robert D. Cooter & Thomas S. Ulen, *An Economic Case for Comparative Negligence*, 61 N.Y.U. L. REV. 1067, 1070-71, 1100-01 (1986); Samuel A. Rea, Jr., *The Economics of Comparative Negligence*, 7 INT'L REV. L. & ECON. 149, 150, 160 (1987); Daniel L. Rubinfeld, *The Efficiency of Comparative Negligence*, 16 J. LEGAL STUD. 375, 376 (1987).

13. *See* COMMITTEE FOR ECONOMIC DEVELOPMENT, WHO SHOULD BE LIABLE? A GUIDE TO POLICY FOR DEALING WITH RISK 49-53 (1989); JAMES S. KAKALIK & NICHOLAS M. PACE, COSTS AND COMPENSATION PAID IN TORT LITIGATION vii-viii (RAND Corp. Institute for Civil Justice No. R-3391-ICJ, 1986). Excluding compensation costs, the total expenditure for all tort litigation in the United States terminating in 1985 (866,000 cases) was between $16 and $19 billion. *See also* ANDREW SCHOTTER & JANUSZ ORDOVER, C.V. STARR CENTER FOR APPLIED ECONOMICS, NEW YORK UNIVERSITY, THE COST OF THE TORT SYSTEM (1986) (stating that the administrative costs of the tort system were $15 to $20 billion in 1984); David M. Trubek et al., *The Costs of Ordinary Litigation*, 31 UCLA L. REV. 72 (1983).

of inefficiency often results in the simultaneous discovery of its solution. For example, under strict liability a requirement that the defendant pay the litigation costs of a successful plaintiff may, under certain circumstances, ameliorate the underdeterrence problem.[14] Under negligence, underdeterrence may be solved by modifying the negligence test to take litigation costs into account.[15]

This argument is presented as a critique of the implications of positive economic theory. One should note at the outset, however, that positive and normative theory are, in an important sense, completely different projects. One explains tort doctrine; the other prescribes an operationally ideal tort system. It is quite possible for the two approaches to conflict and yet be entirely defensible within their own spheres. When the two do conflict, however, the appropriate goal should be operational efficiency, for positive theory is useless if it has no implications for the operational efficiency of the tort system.

Part II of this Article summarizes the traditional positive economic analysis of the tort system. Part III discusses the implications of costly litigation, and presents the case for the three propositions discussed above. Part IV extends this discussion by examining the effects of legal error. As one might expect, introducing legal uncertainty into the analysis complicates the picture of the tort system generated in Part III, but its basic propositions remain unaffected. Part V presents an informal discussion of solutions to the identified deterrence problems. Part VI concludes that the contemporary theoretical model of the tort system is not operationally efficient, and future analysis will remain unsatisfying without consideration of the costliness of litigation.

II. BASIC PROPOSITIONS OF THE POSITIVE ECONOMIC THEORY OF TORT LAW

In the economic theory of tort law, two descriptions of the potential injurer's behavior are of particular concern. One is the instantaneous level of care exercised by the actor, for example, whether the actor drives at reasonable speeds or looks both ways when crossing the street. The other is the potential injurer's amount of involvement in a risky activity—how often the actor drives or crosses busy intersections.[16]

14. *See infra* Part V.

15. *See infra* Part V.

16. The distinction is usually put in terms of "care" and "activity" levels. The first formal presentation of this distinction was presented in Steven Shavell, *Strict Liability versus*

With respect to the instantaneous level of care, strict liability and negligence both lead injurers to exercise the socially optimal level of care and, in this sense, are equivalent according to the positive model.[17] Under the stylized version of strict liability presented in economic models, injurers pay for all losses caused by their activity, whether or not they have exercised precaution.[18] Under negligence, they pay only for those losses where the cost of taking care falls below the benefit in accident loss reduction and they have not exercised precaution. Injurers will not take more care under strict liability because each extra unit of precaution above the point at which they avoid a negligence determination costs them more than they would have to pay in damages.[19]

In contrast to care levels, there is a difference between strict liability and negligence when activity levels are taken into account. Strict liability provides incentives for potential injurers to consider

---

Negligence, 9 J. LEGAL STUD. 1, 3 (1980). *See also* RICHARD A. POSNER, ECONOMIC ANALYSIS OF LAW 139-41 (2d ed. 1977).

17. *See, e.g.*, A. MITCHELL POLINSKY, AN INTRODUCTION TO LAW AND ECONOMICS 40-42 (2d ed. 1989); RICHARD A. POSNER, ECONOMIC ANALYSIS OF LAW 160 (3d. ed. 1986); Shavell, *supra* note 16, at 2. The first paper to formally prove that injurers exercise optimal precaution under negligence was Brown, *supra* note 2.

The claim that strict liability and negligence both lead injurers to exercise optimal precaution is valid only where the probability of an accident is influenced by the injurer's level of care alone. If the probability of an accident is influenced by the levels of care chosen by both the injurer and the victim, then strict liability and negligence are not equivalent in the sense stated in the text. Under negligence, the injurer and the victim will exercise optimal precaution. *See id.* at 340. Under strict liability the victim will take too little care. *See* Shavell, *supra* note 16, at 7.

18. In reality, the rules determining liability in strict liability regimes are not so simple. For example, in the products liability area, the plaintiff who brings a strict liability claim must satisfy the requirements of the RESTATEMENT (SECOND) OF TORTS § 402A (1965). Under section 402A, the plaintiff must prove that the product is "in a defective condition unreasonably dangerous." *Id.* In other words, "the product must be defective in the kind of way that subjects persons or tangible property to an unreasonable risk of harm." W. PAGE KEETON ET AL., PROSSER AND KEETON ON THE LAW OF TORTS § 99, at 695 (5th ed. 1984).

Another example is found in the workers' compensation statutes. Basically, an individual may recover under workers' compensation if (1) the injury arises out of or in the course of employment, (2) the individual is classified as an employee and not an independent contractor, and (3) the injury is of a type that is covered under the applicable workers' compensation statute. *See* 1 ARTHUR LARSON, THE LAW OF WORKMEN'S COMPENSATION § 1.10 (1991). It follows that litigation can focus on any one of these elements.

19. The point at which the injurer avoids a negligence determination is the point at which the cost of taking care is equal to the expected accident loss, on the margin. Suppose that at the injurer's current level of care, the marginal cost of care is $10 and the marginal expected accident loss is $10. In this case, the injurer will not be held negligent if an accident occurs, and, therefore, the injurer will have no incentive to exercise a higher level of care. Even if the rule governing liability is switched from negligence to strict liability, the injurer will not have an incentive to take more care. If the injurer chooses a higher level of precaution, he or she will incur precaution costs greater than $10, but the expected liability will remain $10.

accident losses in determining their activity levels.[20] Negligence provides no such incentive.[21] Despite this disparity, the first proposition stated in the textbook discussions is that the instantaneous level of care is the same under both liability rules—the optimal level of care.[22] This Article focuses on the instantaneous level of care, and ignores additional considerations such as activity levels and risk spreading. The most important reason for this approach is that the theory is quite complicated at this level alone. Taking activity levels or risk aversion[23] into account would complicate the presentation without altering the general conclusion that the tort system is unlikely to be operationally efficient.[24] One reason for leaving risk aversion out of the analysis completely is that the availability of insurance considerably weakens the argument that tort rules should serve a risk-spreading function. Finally, this Article focuses on the instantaneous level of care because this is what we are all thinking about when we discuss the incentives created by tort law. The predominant question is whether tort law makes people take care. If it does not do this, then the efficiency issue is virtually settled; second-best considerations are all that can be marshalled in favor of a particular set of tort rules.

The failure to examine risk aversion and other considerations which would require specification of the preferences of actors results in a focus on "wealth-maximizing" or "efficient," rather than "utility-maximizing," solutions to the problem of controlling accidents through private litigation. It may help to pause at this stage to restate the fundamental argument for taking such an approach. Efficient solutions, which will sometimes be referred to below as socially desirable, are preferred because they maximize the set of consumption possibilities available to society. Alternative solutions waste resources. Although it is possible that under some method of measuring happiness society can be shown to be better off at an inefficient allocation,[25]

---

20. Shavell, *supra* note 16, at 3; *see also* LANDES & POSNER, *supra* note 2, at 66-68; POLINSKY, *supra* note 17, at 47.

21. Shavell, *supra* note 16, at 2; *see also* LANDES & POSNER, *supra* note 2, at 66-68; POLINSKY, *supra* note 17, at 48.

22. *See, e.g.*, POLINSKY, *supra* note 17, at 40-42.

23. A risk averse individual would pay less than $50 for a lottery ticket which offers $100 with probability 1/2 and $0 for a ticket with probability 1/2. A risk neutral individual would pay $50 for such a ticket. For further discussion, see HAL R. VARIAN, MICROECONOMIC ANALYSIS 108-09 (1978).

24. This Article argues that tort liability does not compel injurers to exercise socially optimal precaution, because it fails to internalize social costs. Taking activity levels into account only reinforces this conclusion. If tort liability fails to internalize the relevant social costs, then activity levels will be too high.

25. This proposition, which is familiar in economics literature, is central to the arguments presented in Mark Kelman, *Consumption Theory, Production Theory, and Ideology in the*

this is an unlikely result,[26] depending heavily on the method of measuring happiness. Further, the scholars who are most likely to argue against an efficiency approach are those who think that the tort system is flawed in ways that traditional efficiency analysis has failed to recognize. That, however, is the thesis of this paper, though argued from an efficiency perspective.

To facilitate the presentation in the remaining sections of the text, consider the following example, which illustrates some of the basic propositions of the traditional economic model. Suppose if an accident occurs the loss suffered by a victim is $100. In addition, suppose there are two types of potential injurers: "low-cost" accident avoiders and "high-cost" accident avoiders. Half of the potential injurers are low-cost and the other half are high-cost accident avoiders. The cost of taking care for the low-cost injurers is $20. The cost of taking care for the high-cost injurers is $60. If the potential injurer does not take care, the probability of an accident occurring is 3/4. If the potential injurer takes care, the probability of an accident occurring is 1/4.

Under strict liability, the injurer pays for victim losses that result from his activity whether or not he has taken care (or, more precisely, was negligent). Thus, the injurer compares the cost of taking care with the expected increase in liability which results from failing to take care. For a low-cost accident avoider, that requires a comparison of a $20 cost of taking care with a $50 increase in expected liability.[27] Since the former is smaller than the latter, the potential injurer has an incentive to take care. For a high-cost injurer, the care decision requires a comparison of a $60 cost of taking care with a $50 increase in expected liability, and since the latter is smaller than the former, the high-cost potential injurer will not have an incentive to take care.

The result is the same under negligence. Under a negligence rule, the low-cost injurer compares her cost of taking care, $20, with

---

Coase Theorem, 52 S. CAL. L. REV. 669 (1979); Duncan Kennedy, *Cost-Benefit Analysis of Entitlement Problems: A Critique*, 33 STAN. L. REV. 387 (1981); Duncan Kennedy & Frank Michelman, *Are Property and Contract Efficient?*, 8 HOFSTRA L. REV. 711 (1980); *see also* Cass R. Sunstein, *Legal Interference with Private Preferences*, 53 U. CHI. L. REV. 1129, 1150-1152 (1986) (discussing endowment effects and valuation of goods).

26. To the extent that the endowment effects discussed in the literature cited *supra* note 25 can be accurately described as wealth effects, they are likely to be quite small. Robert D. Willig, *Consumer's Surplus Without Apology*, 66 AM. ECON. REV. 589 (1976).

27. If the potential injurer takes care, expected liability, which is equal to the probability of an accident mulitiplied by the loss, is (1/4)($100) = $25. If the potential injurer does not take care, expected liability is (3/4)($100) = $75. Thus, the expected increase in liability that results from failing to take care is $50.

the expected increase in liability that results from failing to take care, $75.[28] Again, the cost-minimizing decision is to take care.[29] The high-cost injurer compares the $60 cost of taking care with a $0 increase in expected liability,[30] and decides not to take care.

This example assumes that litigation is a costless activity.[31] This assumption is necessary in order to claim with confidence that under strict liability, the potential injurer's expected liability rises by the amount of the increase in expected losses. If, for example, four-fifths of the victims failed to bring suit because the litigation costs exceeded

28. If the potential injurer takes care, her expected liability is $0. If the potential injurer is a low-cost accident avoider and she fails to take care, her expected liability is (3/4)($100) = $75. Thus, the increase in expected liability is $75.

29. One might argue, in light of the analysis presented in Grady, *supra* note 1, and Marcel Kahan, *Causation and Incentives to Take Care Under the Negligence Rule*, 18 J. LEGAL STUD. 427 (1989), that the mathematical representation of the negligence rule given in the text is inaccurate. Mark Grady argues that the causation requirement of the negligence rule does not force negligent defendants to pay for the losses that would have occurred even if the due care standard had been met. The increase in expected liability will be less than $75 under this view because the expected cost of accidents which would have happened even if the defendant had exercised precaution should be subtracted from the $75 estimate. Thus, the increase in expected liability which results from failing to take care is $50, the same as under strict liability. The same result observed under strict liability will be observed under this representation of the negligence test. *See* Mark F. Grady, *Punitive Damages and Subjective States of Mind: A Positive Economic Theory*, 40 ALA. L. REV. 1197, 1199 (1989).

The following is a general description of the negligence rule that incorporates the traditional approach and the approach urged by Grady as special cases. Suppose the cost of taking care is x, and the probability of an accident occurring is p when the injurer takes care and q when the injurer does not take care. Let the loss be $100. Let y be the amount of expected damages passed on to the defendant when he has failed to meet the due care standard. An actor who is potentially negligent compares x with $(py)$$100. Under the traditional approach, $y = 1$. Under the approach urged by Grady, $y = (p-q)/p$. An intermediate rule would be one which satisfies $(p-q)/p < y < 1$.

This Article uses the more traditional mathematical description of negligence in the text because adopting the more restrictive interpretation of the negligence test urged by Grady does not alter the results.

30. The increase in expected liability is $0 for the high-cost injurer because that injurer will not be found negligent under the Hand formula.

31. This has been the standard approach in the theoretical literature. The model developed in John Brown's paper, Brown, *supra* note 2, which assumes that litigation is costless, has been used in much of the theoretical work examining deterrence under the tort system. In one of the earlier papers in this area, Peter Diamond noted the restrictiveness of the assumption of a costless or "frictionless" litigation system. Peter Diamond, *Accident Law and Resource Allocation*, 5 BELL J. ECON. 366 (1974).

The standard approach of assuming a frictionless legal system creates an inconsistency in the literature. The assumption of frictionless litigation is equivalent to assuming that the information needed to determine liability can be generated without cost, which is essentially a "zero transaction cost" assumption. However, a basic assumption of the economic theory of tort law is that positive transaction costs prevent potential injurers and potential victims from bargaining ex ante over the level of precaution to be exercised by the injurer. The standard approach therefore assumes positive transaction costs before the accident, and zero transaction costs after the accident.

the anticipated damage award, then expected liability would increase by $10,[32] which would not be enough to induce the potential injurer to take care.

Thus, under the assumption of costless litigation, the following propositions emerge. First, tort law causes potentially negligent actors to take care when and only when taking care is socially desirable. Second, tort plaintiffs bring suit when and only when suit is socially desirable. This is so because every victim brings suit when the cost of litigating is zero, and because every victim brings suit, victim losses are internalized to injurers. The third proposition is that compensatory damages is the optimal level of damages. This Article challenges all these propositions.

III. COSTLY LITIGATION AND THE INEFFICIENCY OF TORT LAW

A. *The Incentive to Bring Suit*

When litigation is costly, it is no longer obvious that plaintiffs bring suit only when suit is socially desirable.[33] In other words, private and social incentives to bring suit diverge. There are a number of ways to demonstrate this proposition. The straightforward approach is to examine the net benefit from private enforcement, assuming that litigation is costly.

For simplicity, this Part considers a regime of strict liability and assumes that suit is certain to be brought (i.e., the probability that suit will be brough is equal to one) after an accident occurs. This Part continues to assume that there are high-cost and low-cost accident avoiders; however, the following example assumes that the cost of taking care is $80 for high-cost injurers. These assumptions do not affect the validity of the claims made in this Part.

Suppose the cost for a plaintiff to bring a claim is $50 and the cost for a potential injurer to defend himself against a claim is also $50. The expected net benefit from enforcement is equal to the net

---

32. In this case, if the potential injurer takes care, expected liability is (1/4)(1/5)$100 = $5. If the potential injurer does not take care, expected liability is (3/4)(1/5)$100 = $15. Thus, the increase in expected liability which results from failing to take care is $10.

33. This proposition was first demonstrated in Shavell, *supra* note 5. It was disputed in Peter S. Menell, Note, *A Note on Private versus Social Incentives to Sue in a Costly Legal System*, 12 J. LEGAL STUD. 41 (1983). Louis Kaplow, *Private Versus Social Costs in Bringing Suit*, 15 J. LEGAL STUD. 371 (1986) presents an effort to resolve the dispute. The dispute is resolved, fortunately with sufficient rigor and clarity to eliminate incentives for further writing on the dispute itself, in Susan Rose-Ackerman & Mark Geistfeld, *The Divergence Between Social and Private Incentives to Sue: A Comment on Shavell, Menell, and Kaplow*, 16 J. LEGAL STUD. 483 (1987). Shavell's result is confirmed in a more general model of litigation in Hylton, *supra* note 6.

benefit from care (i.e., the care that is taken because enforcement has a deterrent effect) less the expected social cost of suit. These amounts can be calculated using the example discussed earlier. The net benefit from care is the reduction in expected victim losses less the expected cost of taking care. Since only low-cost injurers will take care in this example,[34] the expected cost of taking care is

$$(1/2)\$20 + (1/2)\$0 = \$10.$$

The reduction in expected victim losses is \$25.[35] Thus, the net benefit from private enforcement under strict liability is

$$\$25 - \$10 = \$15.$$

The expected cost of suit is the sum of the total cost of suing multiplied by the probability of an accident occurring.[36] Since only low-cost injurers take care in this example, the probability of an accident occurring is

$$(1/2)(1/4) + (1/2)(3/4) = 1/2.$$

The expected social cost of suit is therefore $(1/2)(\$100) = \$50$. The net benefit from private enforcement under strict liability is

$$\$15 - \$50 = -\$35.$$

A negative net social benefit implies that an activity is draining resources from society. Society would be wealthier without the activity. Thus, if the net benefit from private enforcement is negative, as in this example, society could make itself better off by barring suit. Better to lose \$10 in accident costs than to spend \$100 in trying to prevent their occurrence.[37]

---

34. Only low-cost accident avoiders will take care in this example because they will compare the cost of taking care, \$20, to the increase in expected liability, which is $(3/4\text{-}1/4)(\$100+\$50) = \$75$ (recall that this example assumes that all victims bring suit). Since the increase in expected liability is greater than the cost of taking care, the low-cost accident avoider will have an incentive to take care. The high-cost injurer will compare the cost of taking care, \$80, to the \$75 increase in expected liability, and thus will not have an incentive to take care.

35. Expected loss is $(3/4)\$100 = \$75$ if the potential injurer does not take care, and $(1/4)\$100 = \$25$ if the potential injurer takes care. Thus, the reduction in expected losses that results from taking care is \$50. Since only half the potential injurers (the low-cost accident avoiders) take care, the reduction in expected losses is \$25.

36. This formula ignores the administrative costs of running the court system because these costs normally would be incurred anyway, i.e. they are sunk costs.

37. Although this result may seem harsh, it is based on the assumption that actors either are risk neutral or fully insured. In reality, of course, not everyone will be fully insured, allowing for a possible insurance benefit provided by shifting the loss from an accident. Nevertheless, this does not alter the conclusion of the analysis. As long as the insurance benefit is less than \$20, private enforcement is a net drain on society's resources.

This raises the question whether, and under what circumstances, will potential plaintiffs bring suit when the net benefit from private enforcement is negative. Transaction costs determine whether potential plaintiffs will have incentives to bring suit when suit is socially desirable. In a high transaction cost world (i.e., in a world where it is difficult for potential injurers and plaintiffs to identify each other and bargain over waiving the right to bring suit before an accident occurs),[38] the private incentive to bring suit is determined by the victim's loss and litigation cost. If the anticipated damage award exceeds the cost of litigating, plaintiffs will have incentive to bring suit. In this example, since the loss of $100 exceeds the plaintiff's cost of bringing suit, $50, suit will always be brought, assuming transaction costs prevent parties from bargaining ex ante over waivers.

If transaction costs are low, potential plaintiffs and defendants may bargain ex ante over waiving the right to bring suit.[39] The plaintiff's value of the right to bring suit equals the value of the loss he would suffer if the defendant were not threatened with suit, and therefore did not take care, less the expected cost of bringing suit. The potential plaintiff must receive a payment at least as great as this amount in order to profit ex ante from waiving the right to bring suit. The potential defendant will profit from purchasing a waiver from the potential plaintiff whenever the price of the waiver is less than the potential defendant's expected liability, plus the cost to the potential defendant of taking care. Thus, whenever the defendant's expected costs exceed the value to the potential plaintiff of the right to bring suit, there is an incentive for the parties to trade cash for a waiver of the right to sue.

The question whether there is incentive to trade cash in exchange for a waiver of the right to bring suit arises only when suit is inefficient. If this is so, then inefficient suits occur only in areas where transaction costs prevent parties from bargaining ex ante over waivers. This is almost certainly a very large set of cases. However, if the problem of inefficient suits is limited to areas where transaction costs

38. This definition of transaction costs is narrower than the notion of transaction costs implicit in the Coase theorem. R.H. Coase, *The Problem of Social Cost*, 3 J.L. & ECON. 1 (1960). For example, this Article recognizes informational asymmetry as a separate issue. Under a broader definition of transaction costs, informational asymmetry would simply be a type of transaction cost, or a feature of the market that exists because of transaction costs.

39. In general, parties may contract with each other so that one party is not liable for his negligence toward the other party. *See, e.g.*, Mayfair Fabrics v. Henley, 226 A.2d 602, 605 (N.J. 1967). However, there are exceptions to this rule, generally based on public policy. The premier case is probably Tunkl v. Regents of the University of California, 383 P.2d 441 (Cal. 1963), which held that exculpatory clauses are void where the transaction involves the public interest.

prevent parties from bargaining ex ante, then the claim that private and social incentives to bring suit diverge is considerably narrower than the proceeding example suggests.

If potential plaintiffs and potential defendants are equally informed, inefficient suits will not occur in areas where, because transaction costs are low, parties bargain ex ante over waivers; but if the parties are unequally informed, inefficient suits will be observed even though parties bargain ex ante over waivers. The reader who finds this claim intuitively plausible may wish to skip over the following series of examples.

Example 1: Suppose potential plaintiffs and potential defendants are fully informed. This implies that the potential plaintiff knows whether the potential defendant is a low-cost accident avoider. In this case the value of the right to bring suit against a low-cost accident avoider is[40]

$$\$75 - (1/4)\$50 = \$62.5.$$

The expected liability of a low-cost accident avoider under a regime in which the potential plaintiff is free to bring suit is[41]

$$\$20 + (1/4)(\$50 + \$100) = \$57.5.$$

Because the potential plaintiff would set a price for a waiver that is no less than \$62.5, and the potential defendant would be willing to pay no more than \$57.5, clearly no exchange will take place between potential plaintiffs and low-cost accident avoiders. Nevertheless, this is an efficient result because suit against low-cost accident avoiders is socially desirable. To see this, note that the net benefit of suit if only low-cost avoiders exist is \$50 - \$20 - (1/4)\$100 = \$5, and since the net benefit is positive, such suits should occur.

---

40. The value of the right to bring suit is determined as follows. In a regime in which victims cannot bring suit, injurers would not take care; therefore, the probability of an accident would be 3/4, and victims' expected loss is (3/4)(\$100) = \$75. In a regime in which victims can sue, the probability of an accident would be 1/4, because low-cost injurers would take care given the threat of suit. The expected net benefit to the victim from having the right to sue would be (1/4)(\$100) - (1/4)(\$100) - (1/4)(\$50) = - (1/4)(\$50) = - \$12.5. In other words, in a regime in which victims have the right to bring suit, they suffer an expected loss of \$12.5. The reason they suffer a net loss is that suit only allows them to recover the injury loss, not the private cost of suing. Comparing the losses across regimes, the net gain to the potential victim from switching from a regime in which suit cannot be brought to one in which it can is \$75 - \$12.5 = \$62.5. This example proves the earlier assertion that the value of the right to bring suit is equal to the value of the loss the victim would suffer if the injurer did not take care, less the expected cost of bringing suit.

41. The left hand side of the numerical expression is the sum of the cost of taking care, \$20, and the expected liability facing the injurer. The expected liability is the expectation of the sum of the defendant's cost of litigating and the damage award, which is equal to (1/4)(\$50 + \$100).

The conclusion is different where bargaining takes place between potential plaintiffs and high-cost accident avoiders. The value of the right to bring suit against a high-cost accident avoider is

$$\$75 - (3/4)\$50 = \$37.5.$$

The expected liability of such a potential injurer is[42]

$$(3/4)(\$50 + \$100) = \$112.5.$$

Exchange will occur in this case, with waivers trading at some price above $37.5 and below $112.5. This is the efficient result because suit is inefficient against high-cost accident avoiders. The net benefit of suit against such a group is $50 - $80 - (3/4)($100) = -$105.

This example demonstrates that when parties are fully informed and transaction costs are low enough for them to bargain ex ante, there is incentive to trade waivers of the right to sue only when suit would be inefficient.

Example 2: Consider the case in which neither potential plaintiffs nor potential defendants know ex ante whether the potential defedant is a low-cost accident avoider. The value of the right to bring suit to the potential plaintiff is

$$\$75 - [(1/2)(1/4)+(1/2)(3/4)]\$50 = \$50.$$

The expected cost to a potential defendant under a regime allowing suit is

$$(1/2)\$20 + (1/4)(\$50 + \$100) + (1/2)(3/4)(\$50 + \$100) = (1/2)\$57.5 + (1/2)\$112.5 = (1/2)\$170 = \$85.$$

Since the potential plaintiff demands at least $50, and the potential defendant is willing to pay any price less than $85 for a waiver, exchange will occur in this case. Further, these exchanges will be efficient ex ante.

Example 3: Suppose the potential plaintiff does not know whether the potential defendant is a low-cost avoider, but the potential defendant does know ex ante what type of avoider he is. The value of the right to bring suit to the potential plaintiff is

$$\$75 - [(1/2)(1/4)+(1/2)(3/4)]\$50 = \$50.$$

The expected cost to a low-cost avoider is

$$\$20 + (1/4)(\$50 + \$100) = \$57.5.$$

42. Because the high-cost accident avoider does not have an incentive to take care, he or she incurs only the expected liability from an accident. That amount equals the probability of an accident involving such an injurer, 3/4, mutliplied by the sum of the defendant's litigation cost and the damage award.

Exchange will occur, but this happens only because the potential plaintiff undervalues the right to sue a low-cost accident avoider. If the potential victim correctly valued the right to sue a low-cost accident avoider, he would demand a payment of at least $62.5 before waiving the right to bring suit. As a result of this undervaluation, potential plaintiffs too frequently will trade away their right to bring suit, and too few suits will be brought against low-cost avoiders.

For a high-cost accident avoider, expected liability is

$$(3/4)(\$50 + \$100) = \$112.5.$$

Because this amount exceeds $62.5, trading in waivers will occur. However, some inefficient exchanges will occur because potential plaintiffs overvalue the right to sue high-cost avoiders; too many suits will be brought against high-cost accident avoiders. Thus, private and social incentives to sue diverge despite the trading in entitlements to bring suit.

This disparity covers a potentially broad set of circumstances. It certainly covers situations in which transaction costs prevent potential plaintiffs and potential defendants from bargaining ex ante. Informational asymmetries also prevent the realization of a fully efficient waiver market when transaction costs are so low that such parties can bargain ex ante. Thus, even where bargaining ex ante occurs, the social and private incentives to bring suit may differ.

This conclusion does not rest on the assumption that the cost of taking care differs for every potential injurer. The claim that private and social incentives to sue diverge remains true as long as transaction costs prevent bargaining ex ante over waivers. Nevertheless, the assumption of heterogeneity with respect to the cost of taking care is required by the informational asymmetry argument.

This discussion demonstrates that there are two sources of inefficiency lurking beneath the claim that suit may be brought when it is not socially desirable. The first source is the cost of bringing suit itself. Suit may be inefficient because the cost of bringing suit against any defendant exceeds the net benefits from the deterrence provided by the threat of suit. This source requires that litigation costs be sufficiently high to cancel the deterrence benefits of private enforcement. The second source of inefficiency is heterogeneity in the cost of taking care. This source renders suit inefficient because potential plaintiffs are unable to identify those potential defendants who will find it inefficient to bring suit. This latter source of inefficiency exists as long as litigation costs are positive, no matter how small they are.

The proposition of divergent private and social incentives to

bring suit probably should not be interpreted as an externality problem.[43] The proposition is not attributable to the simple fact that the plaintiff does not bear the defendant's litigation costs.[44] Rather, it arises from transaction costs and informational asymmetry, which prevent efficient trading of entitlements to bring suit, and therefore lead to inefficient suits.

B. *Costly Litigation and the Level of Care*

Perhaps the most important proposition of the positive economic model of tort law is that potential injurers will exercise the optimal amount of precaution under either strict liability or negligence. This proposition depends on the implicit assumption that litigation is a costless activity. This Part challenges the traditional model, and demonstrates that potential injurers will exercise less than optimal precaution under both strict liability and negligence.

This Part assumes that courts decide cases without error. Additionally, it assumes that potential plaintiffs and potential defendants have information on the distribution of accident losses and precaution costs, and use this information to make rational forecasts of the probability of suit and of a negligence verdict.

1. STRICT LIABILITY AND THE PROBLEM OF EXTERNALIZATION

As a group, potential injurers exercise less than the socially optimal level of precaution under strict liability.[45] The reason for this is simple. Generally, an actor will be compelled to exercise the socially optimal amount of precaution when all of the external costs are "internalized."[46] The external costs of failing to take care in a costly litigation system are the losses imposed on victims and the total litiga-

43. Divergent incentives to bring suit are an externality problem in the sense that the potential plaintiff's incentive condition is not the same as the "social incentive condition," i.e. the condition which describes when suit enhances social welfare. However, the problem is not simply a matter of the plaintiff imposing costs on others. If this were the case, a solution could be found by imposing an appropriately set tax on the plaintiff when he or she brings suit. But as Shavell recognizes, there is no simple (implementable) taxation solution to the plaintiff's incentive problem. Shavell, *supra* note 5, at 337. For this reason, it seems inappropriate to refer to this as an externality problem. Further, timing is an important feature of the problem. When the plaintiff decides whether to bring suit it is always after an accident has occurred. The social incentive condition describes the net benefit from suit from an ex ante perspective.

44. This is proven by working through the argument in this part under the assumption that the plaintiff must pay the defendant's cost of litigating. Even under this assumption, it is apparent that the private and social incentives to bring suit diverge.

45. Hylton, *supra* note 6, at 164.

46. This is a basic result of the theory of externalities, usually traced to A.C. PIGOU, THE ECONOMICS OF WELFARE (4th ed. 1946). For an introductory discussion, see VARIAN, *supra* note 22, at 203-07.

tion costs incurred by victims who bring suit.[47] These costs must be internalized in order to achieve optimal deterrence under strict liability.

Optimal deterrence cannot be achieved under strict liability because the system cannot fully internalize the social costs of accidents. In particular, two types of social costs will not be internalized under strict liability: (1) the litigation costs born by victims who bring suit, and (2) the losses suffered by victims who choose not to bring suit because the anticipated damage award exceeds their litigation costs.[48]

The following example demonstrates the underdeterrence problem. Suppose the cost of taking care is $20 for a low-cost accident avoider, and $60 for a high-cost accident avoider. The cost of litigating is, again, assumed to be $50 for each litigant. Assume that the probability of an accident occurring is again 1/4 if the potential injurer takes care and 3/4 otherwise. Finally, assume that half of the victims suffer a loss of $160 from an accident, and the other half suffer a loss of $40.[49] The average loss from an accident remains $100.

Under these assumptions, it is socially optimal for both high- and low-cost avoiders to exercise precaution because the marginal cost of taking care is less than the marginal cost of not taking care for each type of actor. On the margin, the social cost of failing to take care is the sum of the increase in expected losses imposed on victims and the increase in expected total litigation costs. The increase in expected loss is $50.[50] The increase in expected total litigation costs is $25.[51]

---

47. The litigation costs borne by injurers are not external costs. Such costs are part of the social cost of failing to take care, but they are not externalized because they are borne by the injurer.

48. One might argue that this view of the incentive to bring suit is too simplistic. In reality, part of the value of bringing suit is the portfolio of options that are made available once the suit is filed. Thus, the expected value of bringing suit is the anticipated damage award plus an additional "option value." Suit may be brought even in cases where the cost of litigating far exceeds the anticipated damage award. *See* Bradford Cornell, *The Incentive to Sue: An Option-Pricing Approach*, 19 J. LEGAL STUD. 173 (1990).

Although Bradford Cornell's option pricing approach provides a more accurate description of the incentive to bring suit, it has no impact on the general argument of this Article. The option pricing approach would require recognizing that an additional term should be added to the value of an award in order to reflect real incentives, but even with this additional option value term, there presumably would be some claims that would remain unprofitable.

49. The group that suffers a loss of $40 will never have an incentive to bring suit because of the $50 cost of bringing suit.

50. Recall that the expected loss is (3/4)$100 = $75 if the potential injurer fails to take care, and (1/4)$100 = $25 if the potential injurer takes care.

51. If the potential injurer takes care, the total expected litigation cost is (1/4)(1/2)$100 = $12.5. In this example, once an accident has occurred the probability that suit will be

Therefore, the social cost of failing to take care is $75. Since this amount exceeds the cost of taking care for both types of actors, efficiency requires that they take care.

Even though it is socially preferable that potential injurers take care under strict liability, high-cost accident avoiders will not have incentives to take care. The private cost of failing to take care equals the increase in expected liability, which includes the defendant's litigation costs. In this example, the increase in expected liability is $52.5.[52] Since the cost of taking care is $60 for high-cost accident avoiders, they will not take care.

A more difficult question is whether potential injurers have incentives to meet the standard of care that is required by the Hand formula for negligence.[53] Under the Hand formula, a potential injurer must take care to avoid liability when the cost of taking care falls below the expected loss imposed on a victim by failing to take care. The question then becomes whether, under strict liability, all potential injurers for whom the cost of taking care falls below the expected loss suffered by an accident victim will have incentives to take care. Although the Hand formula's failure to consider litigation costs results in its inability to define the socially optimal level of care, examination of the Hand formula remains important because of its significant presence in tort law.

Whether, under strict liability, potential injurers will meet the standard of care required by the Hand formula depends on whether the expected loss suffered by victims who choose not to bring suit exceeds the cost to the potential injurer of defending a claim.[54] In other words, the answer depends on whether the defendant's cost of litigating effectively internalizes the losses suffered by victims who fail to bring suit under strict liability. Recall that in the example discussed above, the increase in expected loss that results from failing to take care is $50. The private cost (i.e., the increase in expected liability) of failing to take care is $52.5. Thus, in this example, the defend-

---

brought is 1/2. If the potential injurer fails to take care, the total expected litigation cost is (3/4)(1/2)$100 = $37.5. The difference, $25, represents the increase in the total cost of litigation that occurs because the potential injurer has not exercised precaution.

52. The increase in expected liability that results from failing to take care is (3/4 - 1/4)(1/2)($160 + $50) = $52.5. This reflects that only half of the potential victims have an incentive to sue. Although this group actually has greater losses, the *expectation* of their losses, which is the important consideration to the potential injurer, is less than the $100 average loss. This result does not depend on the numbers chosen. For a general proof, see Hylton, *supra* note 6.

53. *See supra* note 10.

54. *See* Hylton, *supra* note 6, at 167 n.13; *see also* KEITH N. HYLTON, STRICT LIABILITY VERSUS NEGLIGENCE: A PIGOVIAN REAPPRAISAL (American Bar Foundation Working Paper No. 9001, 1990).

ant's level of liability "overinternalizes" the expected losses suffered by victims.

2. NEGLIGENCE AND THE PRIVATE ENFORCEMENT PARADOX

Positive theory recognizes two basic propositions stemming from the negligence rule. First, potential injurers will exercise the level of care required to avoid liability. Second, if the negligence standard is determined by the Hand formula, the level of care exercised by potential injurers is socially optimal.[55] The existence of positive litigation costs invalidates both propositions.

Addressing the second issue first, recall that the Hand formula requires a comparison of the cost of taking care with the expected losses imposed on an accident victim. A proper efficiency test would require the potential injurer to compare the cost of taking care with the sum of the expected losses imposed on a victim and the litigation costs imposed on society.[56] As long as the due care standard fails to take litigation costs into account, it cannot be said that potential injurers exercise socially optimal precaution even in an equilibrium in which they all obey the due care standard.

The first proposition is similarly unsound. Potential injurers, as a group, will not exercise the level of care that is required to avoid liability under the due care standard. Equilibrium in a negligence regime where litigation is costly requires the existence of actors who refuse to obey the due care standard.[57]

When litigation is costly, plaintiffs will bring suit only if the probability of winning a lawsuit is positive. But if every potential injurer obeys the due care standard, the probability of winning a negligence suit would be zero. If potential plaintiffs (more precisely, plaintiff's attorneys) are aware of this, they would never have an incentive to bring suit.[58] If no potential plaintiff has an incentive to bring suit, potential injurers will not take care.

---

55. *See* Brown, *supra* note 2, at 343. The Brown article also argues, however, that this proposition does not hold if the court does not have enough information to determine the socially optimal level of care. *Id.* at 343-44. *But see* Roberty Cooter et al., *Liability Rules, Limited Information, and the Role of Precedent*, 10 BELL J. ECON. 366 (1979) (arguing that optimal liability rules result from courts' discovery of optimal care levels over time).

56. Hylton, *supra* note 6, at 163.

57. Janusz A. Ordover, *Costly Litigation in the Model of Single Activity Accidents*, 7 J. LEGAL STUD. 243, 244 (1978); *see also* Marilyn J. Simon, *Imperfect Information, Costly Litigation, and Product Quality*, 12 BELL J. ECON. 171, 178 (1981) (reaching a similar result). A simpler proof of Ordover's result is provided in Hylton, *supra* note 6, at 168.

58. Recall that it has been assumed attorneys make rational forecasts of the probability of a negligence verdict. Plaintiffs would refrain from bringing suit because the rational forecast of the probability of a negligence verdict would be zero. Because the expected award would be

The two player game matrix popularized by the Prisoner's Dilemma game provides an alternative interpretation of this impossibility argument.[59] Suppose the cost of taking care is $X. Suppose the expected liability from an accident, given that the injurer acted negligently, is $100, representing the probability that suit will be brought, multiplied by the sum of the cost of litigating and the damage award. Suppose the expected liability from an accident when the injurer has taken care is $50, representing the probability that suit will be brought multiplied by the injurer's cost of litigating. The following diagram describes the incentives of a typical injurer.

FIGURE 1: INJURERS' INCENTIVES TO TAKE CARE

		ONE RATIONAL INJURER	
		Takes Care	Does Not Take Care
ALL OTHER INJURERS	Take Care	$X	$0
	Do Not Take Care	$(X + 50)	$100

Each cell in Figure 1 shows the payoff to a typical injurer, given the actions of all other potential injurers. The cell in the upper left corner describes the situation in which all potential injurers take care. Because no victim has an incentive to bring suit in this situation, the rational injurer incurs only the cost of taking care, $X. The cell in the

zero, and the cost of litigating positive, each negligence claim would have a negative expected value.

This argument assumes that the accident itself does not reveal whether the injurer acted negligently. If events surrounding the accident revealed whether the defendant acted negligently, the victim would not need to forecast the probability of a negligence verdict.

59. *See, e.g.*, R. DUNCAN LUCE & HOWARD RAIFFA, GAMES AND DECISIONS 94-102 (1957). The analogy to the Prisoner's Dilemma is imperfect because the equilibrium under negligence is not a dominant strategy equilibrium. Yet the Prisoner's Dilemma remains instructive because the intuitive arguments associated with it make it easier to present the argument in this section.

upper right corner describes the case in which the rational injurer fails to take care while all other injurers continue to take care. In this case, an injurer avoids the cost of taking care and incurs no increase in expected liability. The injurer incurs no increase in expected liability because when all other injurers are taking care, as is assumed in this case, no victim will expect to win a negligence suit and one injurer's failure to take care will not affect this.[60] In this sense, one injurer can free ride on the caretaking efforts of others. The lower left cell describes the payoff to the rational injurer if he takes care when all other injurers are not taking care. The rational injurer incurs the cost of taking care, $X, plus the expected cost of defending himself against a negligence claim, $50. The expected litigation cost to the injurer who takes care when everyone else behaves negligently is likely to be significant, because if an accident occurs, the victim is likely to assume that it was the result of negligence. The lower right corner cell describes the case in which no injurers take care, and all face the full expected liability of $100.

The most important conclusion that can be drawn from Figure 1 is that the outcome in which all injurers take care (the upper left cell) is not a self-enforcing or Nash equilibrium.[61] Neither is it a dominant strategy equilibrium. If everyone takes care, one injurer can always benefit by not taking care.[62]

The conclusion of this discussion should be clear: under a negligence system in which litigation is costly and potential plaintiffs make rational forecasts of the likely outcome of a trial, some group of potential injurers must be acting negligently if the system is to reach an equilibrium in which potential plaintiffs have incentives to bring suit.

---

60. In other words, it is assumed that the actions of one potential injurer will not affect the expectations of victims. Those expectations arise from the behavior of the group of injurers. Therefore, victims will not expect to win a negligence suit unless a significant share of potential injurers behave negligently.

61. A self-enforcing or Nash equilibrium is one in which neither player has an incentive to deviate from its chosen strategy, given the strategy choice of the other. For a simple introduction to basic concepts of game theory, see MICHAEL D. INTRILIGATOR, MATHEMATICAL OPTIMIZATION AND ECONOMIC THEORY 134 (1971); ERIC RASMUSEN, GAMES AND INFORMATION: AN INTRODUCTION TO GAME THEORY 21-41 (1989).

62. A few minor conclusions should be mentioned. For values of x greater than 50, the strategy "don't take care" becomes dominant for the rational injurer (and the game becomes similar in form to the Prisoner's Dilemma). Thus, if x is the same for all injurers, and is greater than 50, there is an equilibrium in which no one takes care. Of course, if the cost of taking care is extremely high (and given the numbers in Figure 1, 50 is extremely high), it is unlikely that the injurer would ever be found negligent anyway. For values of x less than 50, we can only conclude that the equilibrium will be one in which some choose to take care and some choose not to take care. The answer will depend on the chosen strategies and the specific payoffs assumed in Figure 1.

Because this conclusion has the flavor of a paradox, it will, no doubt, be questioned. Indeed, this Article eventually argues that the private enforcement paradox fails to hold under a different, and more reasonable, set of assumptions.[63] Nevertheless, this conclusion arises from well established assumptions of the positive economic model of the tort system. The only novel assumption is the uncontroversial one that litigation is costly.

This is by no means the first paradox derived from an economic analysis. Two paradoxes, whose proofs are broadly similar to the one stated here, come to mind. The first is attributable to Schumpeter, who noted that innovation, being costly, would never occur if the zero-profit equilibrium implied by the model of perfect competition held at all times.[64] The second and more recent is the efficiency paradox of Grossman and Stiglitz, which states that securities markets cannot be informationally efficient, for if they were, there would be no incentive to engage in costly research, and without such research the market could not remain informationally efficient.[65] In this Article, private enforcement is the analogue of costly research in the Grossman-Stiglitz argument. Under negligence, private enforcement cannot work as well as the traditional model has posited; for if it did, the outcome would not be compatible with individual incentives to litigate.

C. *The Ambiguous Desirability of Compensatory Damages*

From the foregoing it should be clear that if litigation is costly, compensatory damages will not always be the socially optimal level of damages.[66] Consider a strict liability regime. Since injurers exercise less than the socially optimal level of care, social welfare could be improved to the extent that a higher damage award compels potential injurers to exercise more precaution. But this is not a complete answer to the question of optimal damages because an increase in the damage award will cause more victims to bring suit, thereby increasing the social cost of litigation. It is therefore not clear that increasing damages above the compensatory level will enhance social welfare.

---

63. The more reasonable set of assumptions includes the assumption that courts occasionally make mistakes. *See infra* Part IV.A.

64. JOSEPH A. SCHUMPETER, THE THEORY OF ECONOMIC DEVELOPMENT 128-56 (1934).

65. Sanford J. Grossman & Joseph E. Stiglitz, *On the Impossibility of Informationally Efficient Markets*, 70 AM. ECON. REV. 393, 404-05 (1980).

66. The premise that setting damages at a level that simply compensates the victim is not necessarily optimal was first demonstrated in Polinsky & Rubinfeld, *supra* note 6, at 152. The discussion in this part is based on the Polinsky & Rubinfeld paper and Hylton, *supra* note 6, app. B. (extending the Polinsky & Rubinfeld analysis by showing that the costly litigation model yields an upper limit on optimal damages).

Although the optimal level of damages under strict liability generally is unclear, the condition that determines optimal damages can be stated with clarity. The marginal benefit of an increase in damages arises from enhanced deterrence. Enhanced deterrence reduces the losses imposed on victims and, consequently, the amount of litigation required to compensate victims. The marginal cost of an increase in damages is the cost resulting from the increased litigation that occurs because more victims have incentives to bring suit. Optimality occurs when, on the margin, the social benefit of enhanced deterrence equals the social cost of additional litigation.[67] A priori, it is impossible to say where this equality occurs, and there is no reason to believe that it occurs at the point where damages just compensate for loss. In spite of this ambiguity, purely theoretical considerations can provide some additional guidance.

First, increasing damages only makes sense if the marginal benefit from enhanced deterrence is positive. If it is zero, the only consideration in the social accounting matrix is the loss from additional litigation. Thus, it never makes sense to increase damages beyond the point at which all of the deterrence benefits have been exhausted, and at which the marginal deterrence benefit that results from an increase in damages is zero.

Second, once the social costs of failing to take care are effectively internalized, the additional deterrence achieved by increasing damage awards creates a negative marginal benefit (i.e., a marginal cost). Increasing damages beyond the point at which relevant social costs are fully internalized compels potential injurers to exercise precaution beyond the socially optimal level.

Third, the social costs of failing to take care can be fully internalized by adopting a simple damages rule: make the defendant pay the victim's cost of litigating in addition to compensating the victim. With damages set equal to the compensatory level plus the victim's cost of litigating, all victims will sue.[68] Therefore, all of the social

67. This is a special measure of the social cost of litigation. Generally, increasing damages has two effects on the total cost of litigation: it reduces litigation costs by reducing the number of accidents, and increases litigation costs by increasing incentives to sue. The reduction in litigation costs which results from a reduction in the number of accidents, i.e. from enhanced deterrence, is measured as part of the social benefit from deterrence. Thus, what is referred to here as the "marginal social cost of litigation" reflects only the increase in litigation which results from an enhancement of plaintiff incentives to bring suit.

68. This assumes, as has been assumed throughout, that victims will bring suit only when the anticipated damage award exceeds the litigation costs. The more important assumption underyling this approach is that the damage award and the cost of suit capture the victim's benefits and costs of litigating. This implicitly excludes areas in which the damage award and cost of bringing suit do not reflect the benefits and costs to the victim of bringing suit. For

costs of failing to take care will be internalized to the injurer.

From the three preceding points, it follows that optimal damages will never exceed the sum of the victim's loss and the victim's litigation costs. This provides an upper limit on optimal damages. This result has implications for any claim that punitive damages are optimal.[69]

One economic argument for punitive damages is that because litigation is costly, not all victims bring suit. Punitive damages, therefore, are necessary to internalize the losses of victims who choose not to bring suit because the costliness of litigation makes it unprofitable.[70] To the extent that punitive damages are based on an attempt to correct the inefficiencies caused by the costliness of litigation, this theory suggests that such damages should not exceed the sum of the victim's litigation costs and the amount required to compensate for the victim's loss. Alternatively, if people fail to bring suit for reasons other than the cost of litigation,[71] the optimal damage measure would not exceed the sum of the victim's loss, divided by the probability that suit will be brought, plus the victim's litigation costs. If the victim's loss is $100, and the cost to the victim of litigating is $50, then in a regime in which only the profitability of litigation matters to potential

---

example, in defamation suits, often the victim will bring suit in order to enhance his reputation, even though the cost of suit exceeds the expected damage award. Given the high rate at which defamation plaintiffs lose, *see* Marc A. Franklin, *Suing Media for Libel: A Litigation Study*, 1981 AM. B. FOUND. RES. J. 797, 797 (five percent of plaintiffs in sample emerged from appellate process with judgments compared with more than 60 percent of defendants), bringing such a suit in order to bolster one's reputation probably explains a significant number of such claims. Although this Article implicitly excludes cases in which the damage award and cost of bringing suit do not reflect all of the benefits and costs to the victim of bringing suit, the theoretical approach remains valid. With fairly minor alterations, the model can be applied to the excluded cases. The more troubling question is whether the cases which implicitly have been excluded are in some sense more representative of tort litigation.

69. *See* David Friedman, *An Economic Explanation of Punitive Damages*, 40 ALA. L. REV. 1125, 1127-28 (1989) (arguing that litigation costs provide a justification for punitive damages). The Friedman paper indepedently derives results similar to those presented in Hylton, *supra* note 6, app. B, and in Polinsky & Rubinfeld, *supra* note 6, but does not show that there is an implied ceiling on damages. The point that the theory implies an upper bound on optimal damages was first demonstrated in Hylton, *supra* note 6, app. B.

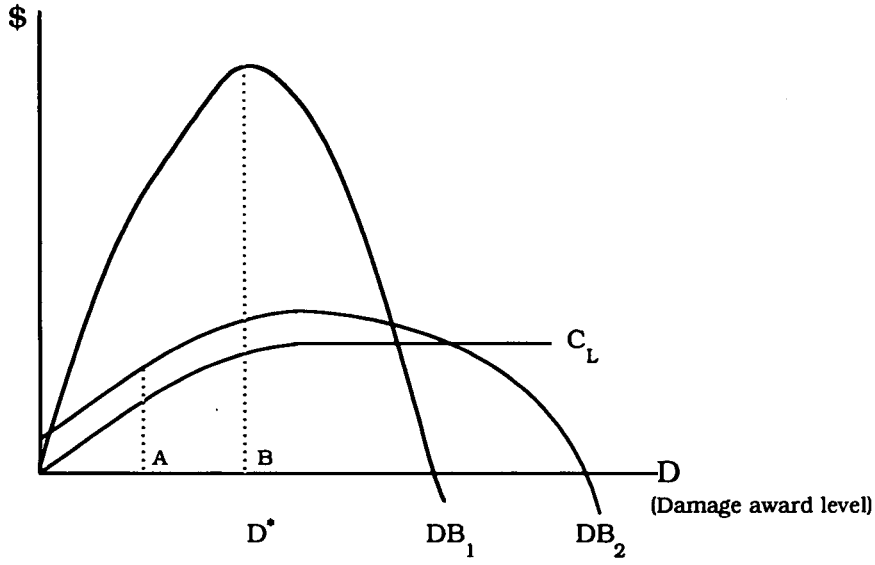
70. An alternative argument is that punitive damages allow the plaintiff to recover the costs of bringing suit. A number of jurisdictions allow the jury to consider the plaintiff's litigation costs as one element of an award of punitive damages. *See* 1 LINDA L. SCHLUETTER & KENNETH R. REDDEN, PUNITIVE DAMAGES § 2.2(B)(1) (2d ed. 1989).

71. For example, suppose a substantial number of tort victims do not sue because of an irrational distaste for courtrooms and lawyers. Consider, as another example, the discussion of defamation suits, *supra* note 66. A politician may be reluctant to bring a defamation suit against a newspaper because of fear that it may affect his popularity with other newspapers. Generally, if there are reputational concerns that affect the decision to bring suit, or if victims have an irrational distaste for litigation, there will be cases in which a claim is not brought even though it seems to be profitable.

plaintiffs, the optimal damage measure will not exceed $150. On the other hand, if the probability that suit is brought is, say, 1/2 even when suit is profitable (and at all levels of profitability), then the optimal damage measure is $250. As this argument suggests, it is possible, and not terribly expensive, to determine empirically whether punitive damages approximate the upper bound on optimal damages.

In any event, it does not follow from the existence of an upper limit on optimal damages that the damage award should always be set at the limit. The upper limit on damages provides an optimal damages formula only if, on the margin, the social benefit from enhanced deterrence exceeds the social cost of additional litigation at all damage levels below the the upper limit. If the marginal social benefit of deterrence does not exceed the marginal social cost of additional litigation at all damage levels below the upper limit, some damage level strictly below the upper limit will be optimal. Figure 2 illustrates this result.

FIGURE 2: THE SOCIAL BENEFITS AND COSTS OF DAMAGES AWARDS



D* = Victim's Loss + Victim's Litigation Costs
Note: At point A, the slopes of DB_2 and C_L are equal.

As Figure 2 demonstrates, the social benefit from deterrence is increasing for damage levels below the upper limit, and decreasing for damage levels above the upper limit. This reflects deterrence beyond the point at which actors internalize the social costs of failing to take

care, which forces actors to exercise more than the socially optimal level of precaution. The curve C_L represents the social cost from additional litigation generated by raising damages, which makes suit more attractive to a larger number of potential plaintiffs. This measure of cost does not increase after the level of damages has reached the upper limit (point B in Figure 2), because once the level of damages reaches the upper limit, increasing it will not cause more victims to sue. Suit is free for plaintiffs once damages are set at the upper limit. Therefore, every victim brings suit.

Two solutions for the optimal level of damages are shown in Figure 2. The first solution is the upper limit, which corresponds to the curve measuring deterrence benefits labeled DB_1. If DB_1 measures deterrence benefits, then the optimal damage level is equal to compensatory damages plus the victim's litigation costs. The other solution for the optimal damages level (point A in Figure 2) is strictly less than than the upper limit, and is the solution consistent with the curve labeled DB_2. Whether the former solution (point B) or the latter solution (point A) will describe the optimal solution depends on the rate at which deterrence benefits respond to changes in damage levels. Obviously, the curve DB_1 depicts a situation in which deterrence benefits are more responsive to changes in the level of damages than does the curve DB_2. Generally, deterrence benefits will be more responsive to changes in damage levels (1) the more productive[72] the injurer's care is and (2) the more likely it is that the injurer pays attention to damage levels in deciding whether to take care.[73] Thus, the curve labeled DB_1 reflects a situation in which the potential injurer's care is more productive and/or the potential injurer is more likely to take damages into account in choosing his level of care.

IV. LEGAL UNCERTAINTY, COSTLY LITIGATION, AND THE OPERATIONAL INEFFICIENCY OF TORT LAW

To this point it has been assumed, in accordance with the positive economic model of the tort system, that courts decide cases without error. This section goes beyond the standard model to consider

72. The injurer's care is productive if it has a great effect on the probability of an accident occurring or on the amount of injury from an accident. Care is unproductive if exercising a greater level of precaution does little to alter the probability of an accident occurring.

73. Some degree of attentiveness to changes in expected liability is an obvious requirement of any deterrence-based theory of liability. Calabresi's preference for placing liability on the "cheapest cost avoider" obviously requires an examination of attentiveness to changes in expected liability. *See* CALABRESI, *supra* note 3, at 135-40. For an argument that such attentiveness should be the predominant concern, see Howard A. Latin, *Problem-Solving Behavior and Theories of Tort Liability*, 73 CAL. L. REV. 677 (1985).

the implications of judicial uncertainty.[74] Two types of uncertainty are discussed here: (1) uncertainty as to whether strict liability or the negligence rule should govern in a particular fact setting, and (2) misapplication of the negligence test.[75] Misapplication of the test can be divided into Type I error, in which the court erroneously protects a defendant who should be held liable, and Type II error, in which the court erroneously holds a defendant liable who should not be.[76]

A. *Negligence, Costly Litigation, and Legal Error*

The private enforcement paradox no longer holds under negligence when Type II error is possible, or when there is uncertainty as to whether negligence or strict liability governs.[77] Victims will bring suit even when all potential injurers obey the due care standard as long as there is a positive probability that damages will be awarded.

Three types of equilibria can occur in a negligence regime in which litigation is costly and there are the types of uncertainty identified above. One is the undercompliance equilibrium described ear-

74. One reason for judicial uncertainty is the juries' inability to collect and correctly evaluate all of the relevant facts. Although the law and economics literature generally treats the jury as a black box that somehow produces correct decisions, the psychological decision theory literature provides a number of reasons to expect fact finders to make mistakes. For example, preconceptions often interfere with observation, influencing the conclusions drawn from a given set of facts. Also, people generally do not pay attention to the quality of information. They use biased samples without attempting to correct for bias. *See generally* RICHARD NISBETT & LEE ROSS, HUMAN INFERENCE: STRATEGIES AND SHORTCOMINGS OF SOCIAL JUDGMENT (1980); Amos Tversky & Daniel Kahneman, *Judgment under Uncertainty: Heuristics and Biases*, 185 SCIENCE 1124 (1974); Paul Slovic et al., *Behavioral Decison Theory*, 28 ANN. REV. PSYCHOL. 1 (1977). In addition, there is a body of literature examining the reliability of jury decisions which has pointed to specific types of biases common among jury decisions. *See* AUDREY CHIN & MARK A. PETERSON, DEEP POCKETS, EMPTY POCKETS: WHO WINS IN COOK COUNTY JURY TRIALS (1985); Dale W. Broeder, *The University of Chicago Jury Project*, 38 NEB. L. REV. 744 (1959); Jonathan D. Casper et al., *Cognitions, Attitudes and Decision-Making in Search and Seizure Cases*, 18 J. APPLIED SOC. PYSCHOL. 93 (1988).

75. Such uncertainty may be generated by inconsistent jury decisions or by the development of complex legal rules. For an argument stressing the latter, see Anthony D'Amato, *Legal Uncertainty*, 71 CAL. L. REV. 1 (1983).

76. Usually the definition of Type I or Type II error depends on the null hypothesis. This article follows the standard definition of these terms. *See, e.g.*, A. Mitchell Polinsky & Steven Shavell, *Legal Error, Litigation, and the Incentive to Obey the Law*, 5 J.L. ECON. & ORG. 99 (1989).

Earlier articles examining the influence of legal error on deterrence include Richard A. Posner, *An Economic Approach to Legal Procedure and Judicial Administration*, 2 J. LEGAL STUD. 399 (1973); Isaac Ehrlich & Richard A. Posner, *An Economic Analysis of Legal Rulemaking*, 3 J. LEGAL STUD. 257 (1974); Donald Wittman, *Two Views of Procedure*, 3 J. LEGAL STUD. 249 (1974). Closely related is Ronald A. Heiner, *Imperfect Decisions and the Law: On the Evolution of Legal Precedent and Rules;* 15 J. LEGAL STUD. 227 (1986) (arguing that stare decisis is an optimal adaptation to unavoidable Type I and Type II errors).

77. The discussion presented in this section is based on Hylton, *supra* note 7.

lier.[78] The second is perfect compliance, where only actors who would be found negligent if courts operated without error have incentives to take care.[79] The third is an overcompliance equilibrium, in which even actors who would never be found negligent if courts operated without error have incentives to take care.[80]

A brief game-theoretic explanation shows why perfect and overcompliance equilibria are possible in negligence regimes. In a regime in which neither of the types of uncertainty discussed in this section exists, undercompliance results because potential tortfeasors know that if suit is costly and the expected award is zero, victims will not bring suit in an equilibrium in which there is perfect or overcompliance. The possibility of Type II error, or uncertainty about whether negligence or strict liability applies, renders this logical inference invalid. Both types of uncertainty essentially permit plaintiffs to adopt a mixed strategy in which suit is brought under either a negligence or a strict liability standard. Under these conditions, it is no longer rational for a potential tortfeasor to deviate from the care standard in any equilibrium in which there is perfect or overcompliance.

As this discussion suggests, the possibility of Type II error and uncertainty about whether strict liability or negligence applies combine to produce a mixed strict liability/negligence regime in which a fraction of the disputes in a negligence regime are decided under a strict liability standard. This fraction is the effective probability that strict liability applies. In perfect and overcompliance equilibria, the probability of a negligence verdict is equal to the effective probability that strict liability applies. This must be true because there are no negligent actors in perfect and overcompliance equilibria.[81]

The implications of this analysis for the optimality of precaution

---

78. *See supra* Part III.B.2.

79. Under the Hand formula, actors who "could be found negligent" are those for whom the cost of taking care falls below the expected loss imposed on victims. If a member of this group fails to take care, he will be held liable. *See supra* note 10.

80. Actors who "would never be found negligent" are those for whom the cost of taking care exceeds the loss imposed on victims. Richard Craswell & John E. Calfee, *Deterrence and Uncertain Legal Standards*, 2 J.L. ECON. & ORG. 279, 280 (1986) (noting that overcompliance is likely to be common where uncertainty with repect to the legal standard is "relatively small").

81. To see this, let s be the probability that strict liability applies. Let v_1 equal the probability of Type I error under negligence. Let v_2 equal the probability of Type II error under negligence. Let w equal the probability that the defendant is negligent. Then the probability of a verdict in favor of the plaintiff is $s + (1-s)[w(1-v_1) + (1-w)v_2]$. In a perfect or overcompliance equilibrium, $w = 0$ (because by definition there are no negligent tortfeasors). Thus, in a perfect or overcompliance equilibrium the probability of a verdict in favor of the plaintiff is $s + (1-s)v_2$, which can be interpreted as the probability that strict liability effectively governs.

are somewhat complicated, because they depend on the type of equilibrium that results under negligence. Nevertheless, the general conclusion of the earlier negligence discussion remains: potential injurers, as a group, exercise too little precaution, irrespective of the type of equilibrium.[82] Consider those potential injurers for whom the cost of taking care exceeds the expected loss imposed on a victim. Because their failure to take care also imposes litigation costs on society, they should take care whenever that cost is less than the sum of expected victim losses and litigation costs. Members of this group will be found negligent only when a victim brings suit, an event which will occur with probability less than 1, and even then only when either Type II error occurs or when the court decides that strict liability applies. Thus, potential injurers will have insufficient incentives to take care.[83] This is true regardless of the type of equilibrium realized under negligence, so negligence will fail to optimally deter this type of actor.

With respect to those actors for whom the cost of taking care is less than the expected loss imposed on a victim, the implications are dependent upon the type of equilibrium that is realized under negligence. They will exercise less than socially optimal precaution in an undercompliance equilibrium.[84] However, in a perfect or in an overcompliance equilibrium, they will take care, and thus, exercise optimal precaution.

While it may be considered good news that legal uncertainty need not result in undercompliance under negligence, this discussion has troubling implications. In particular, it implies that in an equilibrium in which perfect compliance occurs, negligence suits will still be brought. Indeed, suit will necessarily be brought against non-negligent defendants. Moreover, since courts make mistakes, some non-negligent defendants will be required to pay damages.

Unfortunately, it is up to the class of hapless victims of aggressive litigants to support a perfect or overcompliance equilibrium.[85]

---

82. This conclusion is valid under the assumption that the due care standard is determined by the traditional Hand formula test for negligence. If the test for negligence is modified to incorporate litigation costs, then a negligence regime optimally deters in a perfect compliance equilibrium. *See* Hylton, *supra* note 7, at 445.

83. *See* Hylton, *supra* note 7, at 443.

84. Recall that in an undercompliance equilibrium, potential injurers fail to exercise the level of precaution required by the Hand formula. However, the Hand formula understates the socially optimal level of precaution because it ignores litigation costs. Therefore in an undercompliance equilibrium, potential injurers exercise less than the socially optimal level of precaution.

85. An "aggressive litigant" is one who brings suit when he otherwise would not if the probability of an erroneous verdict in favor of the plaintiff were zero. This definition

Ideally, everyone obeys the due-care standard and no one brings suit, provided that the due-care standard requires injurers to exercise the socially optimal level of precaution. However, this ideal seems, at least in theory, to be unattainable under negligence. The logic of the negligence system gives us a disappointing choice: accept undercompliance, and the higher rate of accidents that it necessitates, or have full, or even more than full, compliance with the only activity in the courts being sham litigation.

B. *Strict Liability and the Influence of Error*

As in negligence, error operates on two levels in a strict liability regime. One is the determination whether strict liability governs the relationship between the plaintiff and defendant.[86] More precisely, in deciding whether strict liability should govern, error influences the determination whether, for example, the defendant's product falls under Section 402A of the *Restatement (Second) of Torts*,[87] or whether the defendant's actions generally can be classified as "abnormally dangerous" and therefore eligible for classification under Sections 519 and 520 of the *Restatement*, which embody the doctrine of *Rylands v. Fletcher*.[88] Error also may affect the actual determination

---

encompasses several types of litigants who would not ordinarily be considered aggressive. For example, consider a victim who suffers a loss of $500. Suppose the cost of litigating is $40 and the likelihood of error in favor of the plaintiff is 1/10. Then, in an equilibrium in which all injurers were complying with the negligence standard, the net expected value of such a victim's claim is (1/10)($500) - $40 = $10. A victim in this situation would have an incentive to bring suit. This definition of "aggressive litigant" includes not only individuals who knowingly bring frivolous claims; it includes some individuals who bring claims that appear to be "legitimate."

Although one might argue that at least some of these aggressive litigants will be deterred by the threat of Rule 11 sanctions, deterrence of aggressive litigation, however defined, is not understood to be the aim of Rule 11. *See, e.g.*, Westmoreland v. CBS, Inc., 770 F.2d 1168, 1180 (D.C. Cir. 1985) (creative litigation is encouraged and aggressive litigation should not be penalized). In addition, the American Rule of awarding attorney's fees upon a showing of bad faith may not deter aggressive litigation. *See, e.g.*, Shimman v. International Union of Operating Eng'r Local 18, 744 F.2d 1226, 1232 (6th Cir. 1984) (attorney's fees awarded for bad faith but not for aggressive litigation).

86. The influence of Type II error on the incentives of injurers to take care is not the same under strict liability and under negligence. Under strict liability, Type I error occurs when an actor who is not in a relationship in which strict liability governs is held to be in such a relationship. In a "pure strict liability" regime, i.e., a regime in which injurers can sue only under the doctrine of strict liability, this should not affect the incentives of the relevant group of injurers—injurers who are in a relationship in which strict liability governs. Type II error, under pure strict liability, is analogous to a lump-sum tax imposed on everyone. On the other hand, Type I error dilutes incentives to take care in a pure strict liability regime.

In a "mixed" negligence/strict liability regime, injurers who cannot bring a strict liability claim may bring a negligence claim. Type II error will affect incentives of injurers under such a regime.

87. RESTATEMENT (SECOND) OF TORTS § 402A (1967).

88. 3 L.R.-E. & I. App. 330, 338 (H.L. 1868). For recent cases discussing Sections 519

of whether the defendant acted negligently.

The description of the influence of error and uncertainty operating at two stages is artificial because the potential injurer who is aware that she may be held strictly liable essentially will operate under a mixed negligence/strict liability tort system.[89] The deterrence properties of such a regime are the same as those of a negligence regime. However, the mixed negligence strict liability system may have merits over a negligence regime in which error is possible and a system of pure strict liability.[90] The mixed negligence/strict liability regime that arises because of error combines the strengths and weaknesses of both regimes. The resulting regime can avoid an undercompliance equilibrium.[91] Further, mixed negligence/strict liability can outperform pure strict liability on deterrence grounds.[92] Because perfect or overcompliance is possible in such a negligence system, it is possible, at least in theory, that a mixed negligence/strict liability regime could be superior on deterrence grounds to a regime of pure strict liability.

V. POTENTIAL SOLUTIONS TO THE PROBLEM OF SUBOPTIMAL DETERRENCE

It hardly needs saying that the tort system fails to compensate all victims. Indeed, tort law's failure to compensate victims has been the central argument of advocates for no-fault systems.[93] Up to this point, this Article has argued that from a theoretical perspective the tort system is far from ideal on deterrence grounds as well. The fol-

---

and 520 and the doctrine of *Rylands v. Fletcher*, see Williams v. Amoco Product. Co., 734 P.2d 1113 (Kan. 1987); New Jersey Dep't of Env'tl Protection v. Ventron, 468 A.2d 150 (N.J. 1983); Branch v. Western Petroleum, Inc., 657 P.2d 267 (Utah 1982).

89. Because of uncertainty as to whether the doctrine of strict liability or negligence applies to a given set of facts, one can argue that the relevant doctrine is, in effect, a mixture of several theories of liability. *See, e.g.*, 2 MARSHALL S. SHAPO, THE LAW OF PRODUCTS LIABILITY, ¶ 26.02 (2d ed. 1990).

90. "Pure strict liability" refers to a regime in which claims can be brought only under the doctrine of strict liability, i.e. negligence claims are not available.

91. The undercompliance paradox discussed *infra* Part III.B.2. can be avoided as long as there is a positive probability that the defendant will be held strictly liable.

92. Suppose that under negligence, injurers comply, in equilibrium, with the Hand formula standard. Suppose that under pure strict liability, the same group of injurers fail to exercise the level of precaution required by the Hand test. This result is a more likely when the probability of Type II error is significant, and the defendant's cost of litigating under strict liability is small. Nothing in the model of the tort system presented in this Article would prevent this result. For a more detailed presentation of this argument, see HYLTON, *supra* note 54.

93. *See* ROBERT E. KEETON & JEFFERY O'CONNELL, BASIC PROTECTION FOR THE TRAFFIC VICTIM 1, 34-69 (1965); JEFFERY O'CONNELL, ENDING INSULT TO INJURY: NO-FAULT INSURANCE FOR PRODUCTS AND SERVICES 9-24 (1975).

lowing discussion addresses whether the deterrence features of the tort system can be improved.

A. *Strict Liability*

Consider pure strict liability. This Article has previously argued that this system underdeters because it fails to internalize the losses of non-suing victims and the litigation costs of suing victims. There are two potential solutions to this problem. First, the state could tax each defendant in a strict liability suit an amount equal to the expected loss imposed on a non-suing victim, divided by the probability that suit is brought, plus the cost of bringing a claim.[94] Second, the law could require the defendant to pay the victim's cost of litigating.

Both potential solutions raise problems. The taxation solution at least raises the possibility that an expensive administrative agency would have to be established to assess and collect the tax. Further, unless tax revenue passes directly to plaintiffs who bring claims, such a system will create incentives for defendants to settle out of court by paying plaintiffs slightly more than they net from litigation, but far less than the defendant would have to pay after judgment and taxation.[95] In the worst case scenario, the problem of insufficient internalization would remain, virtually in its present form.[96] The problem of out-of-court settlements could be avoided by barring them altogether. That, however, would amount to forcing victims who bring claims to subsidize a general deterrence benefit. In addition, the mere policing of a ban on out-of-court settlements would require state monitoring, which would be expensive, probably infeasible, and quite unpopular among the supposed beneficiaries.

The problem with requiring the defendant to pay the victim's cost of litigating is that it may result in an inefficiently large amount of litigation. Recall that compensatory damages plus the cost of bringing a claim is the optimal damages formula only if the tortfeaser exercises sufficient care to reduce accident losses.[97] One cannot be sure that this describes all situations in which a claim of strict liability

94. *See* Hylton, *supra* note 6.

95. Jerry Green, *On the Optimal Structure of Liability Laws*, 7 BELL J. ECON. 553, 554 (1976) (discussing general problem of side payments that arises whenever total fines or damages exceed the losses suffered by the parties to an accident); *see also* William M. Landes & Richard A. Posner, *The Private Enforcement of Law*, 4 J. LEGAL STUD. 1, 24 (1975) (discussing incentive to bribe created by taxing private enforcers).

96. Internalization remains because defendants would settle out of court by offering plaintiffs only one penny more than they would net from bringing suit. The externalization problem identified earlier also remains.

97. In addition, the potential injurer must pay attention to liability in deciding the level of care. *See infra* Part III.C.

may be brought. Indeed, in many cases, the victim's level of care may be the most important determinant of accident losses.[98]

These considerations suggest that it is hard to solve the deterrence problem under strict liability. Each situation in which strict liability governs should be considered separately. In spite of this, there are some rather obvious "second best" suggestions. Given the problem of out-of-court settlements, it seems that the only adequate solution to the externalization problem is to require the defendant to pay the successful victim's cost of bringing a claim.[99] The only question remaining is whether the deadweight loss society bears from insufficient internalization is greater than the loss that would be imposed by excessive litigation. If so, then society would benefit by moving to a regime in which defendants pay the litigation costs of successful victims.

Consider workers' compensation as an example. The solution suggested here would require the employer to pay the cost of litigating for successful plaintiffs.[100] Although this would, no doubt, make the program more expensive to the employer, it would remove litigation costs as an obstacle to complete internalization under workers'

98. A rule requiring the defendant to pay the litigation costs of the victim might also lead to an inefficiently large amount of litigation because it would enhance incentives for victims who are not in a relationship in which strict liability governs to bring suit in order to take advantage of the possbility of Type II error. This would not be a very serious problem in a regime in which victims can only bring strict liability claims. Further, the incentive to engage in this type of rent-seeking litigation reduces greatly if defendants are only required to pay the litigation costs of successful plaintiffs.

99. One might argue that this solution is not good enough. If the defendant is required to pay the litigation costs of the successful plaintiff, litigation costs will continue to prevent some plaintiffs from bringing suit because the plaintiff's cost of litigating will not be reduced to zero. Instead, the cost will merely be discounted by the probability that the plaintiff is unsuccessful. This argument assumes that in pure strict liability regimes, likelihood of success will not be an important issue for a victim who is in a relationship with the injurer in which strict liability governs. For example, there would not be a great deal of uncertainty over whether an employer would be held strictly liable for an injury arising out of employment. If this is a valid assumption, then it is fair to conclude that a rule requiring the defendant to pay the costs of the successful plaintiff would virtually remove litigation costs as an obstacle to loss-internalization under approximately pure strict liability regimes.

100. Although most workers' compensation claims are resolved without litigation, some types of claims in certain states generate a great deal of litigation. For example, 100 percent of workers with low-back injury claims in New Jersey in 1981 had attorney representation. *See* SARA R. PEASE, PERFORMANCE INDICATORS FOR PERMANENT DISABILITY: LOWBACK INJURIES IN WISCONSIN 45 (Workers Compensation Research Institute Series Paper WC-87-4, 1987); *see also* LINDA DARLING-HAMMOND & THOMAS J. KNIESNER, THE LAW AND ECONOMICS OF WORKERS' COMPENSATION 32-33 (Rand Corp. Institute for Civil Justice No. R-2716-ICJ, 1980) ("Despite its no-fault characteristics, the compensation system as it presently operates is heavily adversarial. . . . Data from the 1975 Survey of Closed Workers' Compensation Claims suggest that approximately 20 percent of all claims are contested at some point . . .") (footnote omitted).

compensation.[101]

The employment situation, however, is not the typical one addressed in tort law because the employer and employee bargain ex ante.[102] Accordingly, there may be several objections to this approach. First, arguably tort law cannot provide a benefit in this situation because the employer and employee probably will draft a contract that is optimal with respect to all features, including the safety characteristics of the workplace.

The standard response is that although the employer and employee bargain, they do so in an environment in which information is asymmetrically distributed. Employees are likely not to be fully informed about the safety characteristics of the worksite.[103] If employees are relatively uninformed, employer liability for workplace accidents can best achieve an efficient solution for workplace accidents.[104] Although one could argue that to the extent the problem is one of informational asymmetry, the employer has incentives to correct the asymmetry in order to exploit the gains from bargaining that exist at any inefficient level of safety, this argument is valid only in

---

101. *See* ROBERT I. FIELD & RICHARD B. VICTOR, ASBESTOS CLAIMS: THE DECISION TO USE WORKERS' COMPENSATION AND TORT (Workers Compensation Research Institute Series Report WC-88-5, 1988) (discussing the effect of litigation costs on the decision to pursue a workers' compensation claim). For articles discussing the effects of workers' compensation benefits on safety, see Richard J. Butler & John D. Worrall, *Workers' Compensation: Benefit and Injury Claims Rates in the Seventies*, 65 REV. ECON. & STAT. 580 (1983); James R. Chelius, *The Influence of Workers' Compensation on Safety Incentives*, 35 INDUS. & LAB. REL. REV. 235 (1982); John W. Ruser, *Workers' Compensation Insurance, Experience-Rating, and Occupational Injuries*, 16 RAND J. ECON. 487 (1985). The externalization that results because litigation is costly is not examined in any of these articles, however.

102. This was the argument of Chief Justice Shaw in Farwell v. Boston & Worcestor R.R. Corp., 45 Mass. (4 Metc.) 49 (1842), an opinion which seems to have anticipated the Coase theorem.

103. Safety information may be costly to gather, and because such information is a public good, the incentives to gather it are likely to be weak. Even if information on workplace safety could be gathered without cost, the cost to an individual employee of analyzing data on safety may outweight the perceived benefits. The perception and inference problems noted in the psychological decision theory literature present additional considerations. *See* sources cited *supra* note 72. For a discussion of the implications of these considerations for safety regulation in a competitive labor market, see William T. Dickens, *Occupational Safety and Health and "Irrational" Behavior: A Preliminary Analysis, in* WORKERS' COMPENSATION BENEFITS: ADEQUACY, EQUITY, AND EFFICIENCY 19-40 (John D. Worrall & David Appel eds., 1985) (concluding that safety and health regulation can improve the welfare of workers who are not aware of the safety characteristics of the worksite).

104. This assumes that only the employer's precaution is at issue (the case in which the employee's precaution is at issue is taken up below). This proposition follows from arguments presented in the products liability context in Michael Spence, *Consumer Misperceptions, Product Failure and Producer Liability*, 44 REV. ECON. STUD. 561 (1977). For elaboration, see Keith N. Hylton & Maria O'Brien Hylton, *Rational Decisions and Regulation of Union Entry*, 34 VILL. L. REV. 145, 154-58 (1989).

situations where transferring information on workplace safety is costless for the employer. If it is costly to transfer such information, the employer will not always have an incentive to inform employees about relevant characteristics of workplace safety.[105]

A second objection to the proposal of having the employer pay the cost of bringing a successful claim is that the deterrence benefits of such a requirement would be minimal because it is the victim's care that counts most in this area. It is useful to return to the Coasean argument rejected in the preceding paragraph: the employer and employee bargain. In a system in which employers bear liability for losses that result from workplace accidents, the employer will benefit directly from additional precaution exercised by the worker. Assuming that the employer is better informed than the worker about the benefits of safety, the incentive to correct workers' misperceptions of worksite safety characteristics will be greater under a system in which the employer is strictly liable than under one in which workers bear the losses. Given this, an efficient agreement[106] is more likely to result in the former regime than in the latter.

Another objection is that workers' compensation was never meant to internalize accident losses in a way that would "deter" inefficiently unsafe work environments. This perspective considers workers' compensation to be designed primarily to compensate workers in an efficient manner. The proposal offered here is not inconsistent with that goal. Indeed, it seems to be the only administratively simple way of achieving the goal. Litigation costs prevent optimal deterrence and full compensation. Shifting the cost of bringing a successful workers'

105. The gain to the employer would come from wage reductions that workers would accept in exchange for information on workplace safety. However, there is nothing to guarantee that such wage reductions will equal or approximate the value of accurately evaluated safety enhancements. Thus, the situation may arise in which the value (accurately measured) of a safety improvement exceeds the cost, so that it is efficient to enhance safety, yet workers may be unwilling to accept wage reductions large enough to finance the safety improvement. The unwillingness or, more precisely, the undervaluation of safety on the part of the employees may result because the employees are unwilling to trust the employer's reports concerning the value of safety. *See* Ian Ayres & Robert Gertner, *Filling Gaps in Incomplete Contract: An Economic Theory of Default Rules*, 99 YALE L.J. 87 (1989) (discussing the importance of informational asymmetry as a barrier to Coasean bargaining).

The agency cost literature suggests an additional consideration. *See, e.g.*, Michael C. Jensen & Wiiliam H. Meckling, *Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure*, 3 J. FIN. ECON. 305 (1976). Although an efficient safety enhancement will increase the value of the firm, such an enhancement may not occur if it conflicts with the interests of managers. *See* Hylton & Hylton, *supra* note 104, at 158.

106. If the employer is strictly liable for workplace injuries, the workers will be indifferent as between any two levels of workplace safety (assuming that losses are fully compensated). The employer, on the other hand, will not be indifferent. Thus, an efficient agreement will require the workers to exercise more precaution in exchange for higher wages.

compensation claim to the employer offers an improvement in both categories.

There may be no need to seek additional deterrence benefits through workers' compensation, given the alternative regulatory and market pressures that employers face. Reputational concerns provide market pressure, and federal regulation[107] provides additional incentives to improve workplace safety. These pressures, however, are unlikely to achieve an efficient solution. An employer concerned about the effects of workplace safety on its reputation in the labor market will choose a level of safety that enhances its ability to compete for qualified workers.[108] There is little reason to believe that this will be an efficient level of safety. Federal workplace regulation, like all regulation, suffers from incentive problems which almost surely prevent it from reaching socially optimal levels of safety.[109]

A workers compensation system that internalizes the social costs of suboptimal precaution is the most effective way of promoting efficient safety. Moreover, a system that fully internalizes accident losses is likely to be superior to a mixture of regulation, reputational concerns, and an insufficiently internalizing workers' compensation system. Thus, an additional benefit of the proposal offered here is the abolishment of government workplace safety regulation.

B. *Negligence*

Part IV presented the argument that negligence underdeters

---

107. *E.g.*, Occupational Safety and Health Act, 29 U.S.C. §§ 651 (1988).

108. Suppose the amount of worksite safety can be measured over the [0,1] interval, and suppose workers have preferences for safety amounts that are distributed uniformly over this interval. Two firms competing for workers would occupy the middle of the interval, rather than choose the more efficient system of placing themselves at spots 1/4 and 3/4 away from the boundary. This is the problem of "Hotelling competition," familiar in the economics literature. Harold Hotelling, *Stability in Competition*, 39 ECON. J. 41 (1929); Hylton & Hylton, *supra* note 104, at 158 n.34 (discussing its application in the labor market context).

109. Specifically, public enforcement agencies do not have the same incentives to enforce as do private enforcers. For a discussion of incentive problems in public enforcement, see Mark A. Cohen & Paul H. Rubin, *Private Enforcement of Public Policy*, 3 YALE J. ON REG. 167 (1985). For a more general statement of the problem, see Gary S. Becker & George J. Stigler, *Law Enforcement, Malfeasance, and the Compensation of Enforcers*, 3 J. LEGAL STUD. 1 (1974). For criticism of the Occupational Safety and Health Act (OSHA), see PAUL W. MACAVOY, THE REGULATED INDUSTRIES AND THE ECONOMY 99-104 (1979); HERBERT R. NORTHRUP ET AL., THE IMPACT OF OSHA: A STUDY OF THE EFFECTS OF THE OCCUPATIONAL SAFETY AND HEALTH ACT ON THREE KEY INDUSTRIES—AEROSPACE, CHEMICALS, AND TEXTILES, 79-85, 280-87, 537-42 (Industrial Research Unit, University of Pennsylvania, The Wharton School Labor Relations and Public Policy Series No. 17, 1978); W. Kip Viscusi, *The Impact of Occupational Safety and Health Regulation*, 10 BELL J. ECON. 117 (1979) (analysis of pooled time series and cross section data for the 1972-1975 period failed to show a significant OSHA impact).

because it fails to provide adequate incentives to potential injurers whose precaution costs exceed the victim's expected loss (i.e., injurers who would not be deemed negligent under the Hand test). There are two potential solutions to the problem of inadequate deterrence under negligence. One is to modify the due care standard so that it takes litigation costs into account. The other is to multiply damages, as under the antitrust statutes.[110]

Litigation costs are incorporated into the due care standard in the following way: Instead of asking whether the defendant's cost of taking care exceeded the expected loss, the court would ask whether the defendant's cost of taking care exceeded the expected loss *plus* the expected cost of litigation. In theory, this proposal establishes a negligence regime that provides incentives for potential injurers to exercise socially optimal precaution.[111] This proposal also suggests that an appropriate damage multiplier could eliminate the problem of insufficient deterrence.[112] Multiplying damages could effectively internalize the social costs of failing to take care to those actors whose precaution costs exceed the losses imposed on others.

It is hard to tell which is the better solution to the underdeterrence problem. Juries may lack sufficient information on the administrative costs of the tort system to properly take them into account in applying the Hand test, and it may be very costly to use experts to inform them. The damage multiplier proposal may create equal, if not greater, administrative burdens. Setting an "optimally-deterring" damage multiplier would require a great deal of information about the technology of accidents and continuous correction as the technology changed.

Modifying the Hand test to take litigation costs into account would not be inappropriate merely because such costs are difficult to foresee. Many potential defendants probably fear incurring large litigation costs as much as having to pay damages. Just as the traditional negligence test requires the potential injurer to take into account the foreseeable injury to the victim, the modified test would require the

110. For example, section 4(a) of the Clayton Act provides for treble damages. 15 U.S.C. § 15(a) (1988).

111. The Article's proposal would not guarantee socially optimal precaution. Recall that an "undercompliance" equilibrium can result under negligence, and in such an equilibrium, potential injurers exercise less than the socially optimal level of precaution. *See supra* text accompanying notes 78-83. However, if the due care standard takes litigation costs into account appropriately, the only types of equilibria under negligence would be an undercompliance equilibrium and a perfect compliance equilibrium, and in the latter equilibrium potential injurers would exercise socially optimal precaution. Hylton, *supra* note 7.

112. For a proof of this claim, see Keith N. Hylton, A Model of Strict Liability, 24 (1991) (unpublished paper, on file with author).

injurer to take into account the injury and the victim's cost of pursuing compensation.

The more important issue revealed by this discussion is the tension between doctrinal and operational efficiency that surfaces when administrative costs are taken into account. The Hand test is demonstrably efficient in a doctrinal sense. Indeed, one might say that this is the central conclusion or proposition of the positive economic theory of tort law. However, doctrinal efficiency does not guarantee operational efficiency. Under certain conditions, the modified Hand test suggested here is operationally efficient.[113]

VI. CONCLUSION

Economic analysis may provide a logically consistent explanation of tort doctrine, but providing a normative justification of the features of the tort system is an altogether different question. This Article has argued that such a justification requires a rigorous examination of the influence of costly litigation. The positive economic analysis of tort law, which has dominated the law and economics literature, has not incorporated the dynamics of litigation into its theory of deterrence.

This Article has discussed three features of the tort litigation system from a theoretical perspective: the incentives to litigate, the incentives to exercise precaution, and the level of damages. With respect to each of these features, the system is not necessarily operationally efficient, and in certain instances it may be unambiguously inefficient. For example, costly litigation prevents strict liability from internalizing all of the social costs of failing to take care. Strict liability does not internalize the losses suffered by victims who choose not to bring suit because the cost of litigating exceeds the anticipated damage award, and the litigation costs of victims who bring suit. The negligence rule, as defined by the traditional Hand test, also fails to provide optimal incentives to take care in a regime in which litigation is costly. Unless the economic theory of tort law considers these features, its normative implications will be inevitably distorted.

113. *See* Hylton, *supra* note 7.