

The Fundamental Divergence Between the Private and the Social Motive to Use the Legal System

Author(s): Steven Shavell

Source: *The Journal of Legal Studies*, June 1997, Vol. 26, No. S2 (June 1997), pp. 575-612

Published by: The University of Chicago Press for The University of Chicago Law School

Stable URL: <https://www.jstor.org/stable/10.1086/468008>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



The University of Chicago Law School and The University of Chicago Press are collaborating with JSTOR to digitize, preserve and extend access to *The Journal of Legal Studies*

JSTOR

THE FUNDAMENTAL DIVERGENCE BETWEEN THE PRIVATE AND THE SOCIAL MOTIVE TO USE THE LEGAL SYSTEM

STEVEN SHAVELL*

ABSTRACT

The legal system is an expensive social institution, raising the question of whether the amount of litigation is socially appropriate. The thesis developed here is that it is not—because of fundamental differences between private and social incentives to use the legal system. These differences permeate litigation, affecting decisions about the bringing of suits, settlement versus trial, and trial expenditures. The private-social divergence is attributable to two externalities: when a party makes a litigation decision, he does not take into account the legal costs that he induces others to incur (a negative externality), nor does he recognize associated effects on deterrence and certain other social benefits (a positive externality). Consequently, the privately determined level of litigation can either be socially excessive or inadequate and may call for corrective social policies. A variety of policies are discussed, including taxation versus subsidy of suit, fee-shifting, and promotion versus discouragement of settlement.

I. INTRODUCTION

THE legal system is a very costly social institution, absorbing substantial resources whether measured by the magnitude of legal expenditures, the number of lawyers, or the sheer volume of litigation.¹ That so much effort and expense are devoted to the legal system and that its use has been increasing over the years² have contributed to a widespread belief that the

* Professor of Law and Economics, Harvard Law School. I wish to thank Louis Kaplow, Daniel Klerman, A. Mitchell Polinsky, Edward Snyder, and Kathryn Spier for comments and Judson Berkey for research assistance.

¹ In 1992, for example, expenditures on legal services were \$88.7 billion, or 1.47 percent of gross domestic product; see *Statistical Abstract of the United States 1995*, at 452 and 779. The number of lawyers in the United States in 1994 was 861,000 and constituted about .66 percent of the labor force; see *Statistical Abstract of the United States 1995*, at 411. The volume of cases filed in 1994 in state courts was approximately 86.4 million, and in federal courts, 1.6 million, or about one case for every three people in the population; see Brian J Ostrom and Neil B. Kauder, *Examining the Work of State Courts, 1994*, at 20 (1996).

² Between 1960 and 1992, legal expenditures as a percentage of gross domestic product grew from .523 percent to 1.47 percent, or almost tripled; see *Statistical Abstract of the United States 1971*, at 306 and 722, and *Statistical Abstract of the United States 1995*, at

[*Journal of Legal Studies*, vol. XXVI (June 1997)]

© 1997 by The University of Chicago. All rights reserved. 0047-2530/97/2602-0010\$01.50

amount of litigation is socially excessive.³ This opinion is reflected in what may be identified as a general movement to curb litigation, exemplified by measures aimed at discouraging frivolous lawsuits, by limitations on damage awards, and by the fostering of alternative dispute resolution.⁴ At the same time, one observes employment of a number of policies designed to encourage litigation, notably, legal aid programs, pro-plaintiff shifting of legal fees, and the award of multiplied damages in certain categories of case.⁵

452 and 779. Similarly, during this period, the number of lawyers increased from 286,000 to 788,000, more than doubling; see *Historical Statistics of the United States, Colonial Times to 1970*, at 416; and also *Statistical Abstract of the United States 1993*, at 405. In comparison, between 1960 and 1992 the U.S. population rose by only 41 percent; see *Statistical Abstract of the United States 1995*, at 8. Additionally, in the 10-year period between 1984 and 1994, the civil caseload increased by over 20 percent, and the criminal caseload grew by over 30 percent, whereas the population rose by only about 10 percent; see Ostrom and Kauder, at 7 (cited in note 1).

³ The view that the degree of legal activity is excessive is well illustrated both in the popular press—see, for example, Walter Berns, *Sue the Warden, Sue the Chef, Sue the Gardener*. . . , *Wall St J* (April 24, 1995), at A12; Warren E. Burger, *Too Many Lawyers, Too Many Suits*, *NY Times Book Rev* (May 12, 1991), at 12; and Julie Johnson and Ratu Kamalani, *Do We Have Too Many Lawyers?* *Time Magazine* (August 26, 1991), at 54—and in academic writings and books—see Derek C. Bok, *A Flawed System of Law Practice and Training*, 33 *J Legal Educ* 570 (1983); Richard A. Epstein, *Simple Rules for a Complex World* (1995); Peter W. Huber, *Liability: The Legal Revolution and Its Consequences* (1988); Kevin M. Murphy, Andrei Shleifer, and Robert W. Vishny, *The Allocation of Talent: Implications for Growth*, 106 *Q J Econ* 503 (1991); and Walter K. Olson, *The Litigation Explosion: What Happened When America Unleashed the Lawsuit* (1991).

⁴ Frivolous suits are discouraged in a number of ways, notably by Rule 11 of the Federal Rules of Civil Procedure, which authorizes courts to impose sanctions when parties lack a good faith basis for litigation, and by the ability in many states of defendants to sue in tort to recover fees for malicious prosecution. On procedural sanctions and tort liability as methods of controlling frivolous litigation, see, generally, John W. Wade, *On Frivolous Litigation: A Study of Tort Liability and Procedural Sanctions*, 14 *Hofstra L Rev* 433 (1986).

Limitations on damage awards have been imposed in a variety of situations, and, one presumes, have the effect of retarding litigation and its intensity. For example, many states have placed caps on noneconomic damages and punitive damages, restricted joint and several liability, or abrogated the collateral source rule (so that a plaintiff's damages are reduced by insurance benefits received); see, for instance, Note, *1986 Tort Reform Legislation: A Systematic Evaluation of Caps on Damages and Limitations on Joint and Several Liability*, 73 *Cornell L Rev* 628 (1988). Also of interest at the national level are recently proposed "common-sense" tort reform acts, which include in their elements ceilings on punitive damage awards and constraints on joint and several liability; for a description, see Note, "Common Sense" *Legislation: The Birth of Neoclassical Tort Reform*, 109 *Harv L Rev* 1765 (1996).

Alternative dispute resolution has been widely promoted as a general and desirable substitute for litigation; see, generally, Stephen B. Goldberg, Frank E. A. Sander, and Nancy H. Rogers, *Dispute Resolution* 6–11 (2d ed 1992). Furthermore, the courts themselves have actively sought to encourage settlement through use of alternative dispute resolution, as is illustrated by Pub L No 100-702, which created a court-annexed arbitration program allowing federal courts to order cases to (nonbinding) arbitration under certain circumstances.

⁵ Legal aid programs provide free or subsidized legal services to the poor. The most important such programs are affiliated with the Legal Services Corporation, which in 1994 re-

Against this background, the question naturally arises: what is the socially appropriate amount of litigation? Should we view the extent of legal activity resulting from individuals' decisions about litigation as approximately socially correct? Or are there systematic problems with the privately generated use of the legal system that produce a socially undesirable amount of litigation?

The thesis of this article is that the level of litigation is not generally socially correct because there exist what may fairly be called fundamental differences between private and social incentives to use the legal system. These differences permeate the litigation process, from the choice of a harmed party whether to bring suit, to the plaintiff's and the defendant's negotiation over settlement versus trial, to their various decisions about legal expenditures. Furthermore, the divergence between the private and the socially appropriate use of the legal system is quite distinct from any problems that may arise from lawyers' failure to act in their clients' true interests, from possible lack of competition in the market for legal services, or from undesirable features of legal rules themselves.⁶

What underlies the divergence between private and social incentives to use the legal system? One source of the divergence is the difference between the private and the social *costs* associated with use of the legal system: the legal costs that an individual party bears generally are less than the

ceived funds of \$640 million from Congress and private donors, employed over 11,000 individuals, provided services in over 1,200 neighborhood offices, and closed about 1.7 million cases; see the Legal Services Corporation *1994 Annual Report*, at 3 and 6. For articles containing general descriptions of legal aid, see Richard L. Abel, *Law without Politics: Legal Aid under Advanced Capitalism*, 32 *UCLA L Rev* 474 (1985); and Marc Feldman, *Political Lessons: Legal Services to the Poor*, 83 *Georgetown L J* 1529 (1995).

Pro-plaintiff shifting of legal fees—that is, shifting plaintiff fees to defendants when plaintiffs prevail but not shifting defendant fees to plaintiffs when they lose—is frequently observed (and appears to be the most common form of fee-shifting in the United States). See, generally, Mary Frances Derfner and Arthur D. Wolf, *Court Awarded Attorney Fees* (1995); and for description of surveys on fee-shifting, see Note, *State Attorney Fee Shifting Statutes: Are We Quietly Repealing the American Rule?* 47 *L & Contemp Probs* 321 (1984); and Susan M. Olson, *How Much Access to Justice from State "Equal Access to Justice Acts"?* 71 *Chi Kent L Rev* 547, 553 (1995).

Awards of multiplied damages are authorized by many federal statutes, such as those providing for treble damages in antitrust actions, up to treble damages for patent and trademark infringement, and treble damages in Racketeer-Influenced Corrupt Organization (RICO) actions. Many state statutes also permit multiplied damages, for example, under actions for timber trespass, tenant holdover, landlord invasion of possessory rights, and the charging of usurious interest. On multiplied damages, see Dan B. Dobbs, *Law of Remedies* 358–63 (2d ed 1993).

⁶ With regard to the latter, consider, for example, rules that are excessively complex and therefore needlessly expensive to litigate or rules that allow parties to collect awards on weak evidence or without proper account being taken of victims' contributory behavior. Even if legal rules are unlike these and are properly formulated, the problems in this article still exist.

full social legal costs, that is, the costs borne by both litigants and by the state itself. This difference clearly creates a tendency toward overuse of the legal system. Specifically, when a person brings suit, he bears only his own legal expenses; he does not take into account that his suit will cause the defendant and possibly the court to incur legal expenses as well; a bias toward excessive suit is thus engendered. Similarly, once suit has been brought, when either litigant considers making a particular expenditure on litigation, he will not count as a cost to himself the expense that the opposing side and the court may be forced to bear as a consequence; this leads to an excessive level of litigation expenditures. Relatedly, when the disputing parties contemplate settlement, they will not credit as a savings therefrom the costs the court would have borne at trial, suggesting that they will have an inadequate incentive to settle.

But there is another important source of divergence between the private and the social incentive to use the legal system: that involving the difference between the private and the social *benefits* of its use. This divergence in benefits can work either to exacerbate or to counter the tendency toward its excessive use due to the private-social cost divergence. To explain, consider one of the principal social purposes of litigation, deterrence of unwanted behavior. This social goal has little to do with a person's decision whether to bring suit. The motive of a person who brings suit is ordinarily not chiefly, if at all, to deter socially undesirable behavior in the future. Rather it is usually to obtain compensation for harm or other relief. Therefore, the plaintiff's benefit from suit does not bear a close connection to the social benefit associated with it and may bear almost no connection at all. As I will amplify, it could be that the plaintiff's benefit from suit exceeds the social deterrent benefit (suppose that damages are high but that deterrence is slight because there is little injurers can do to reduce harm). Or it could be that the plaintiff's return from suit is less than its deterrent effect (suppose that damages would be small but that deterrence would be significant because injurers can exercise cheap and effective precautions). If the private benefit from suit exceeds the social, the tendency toward socially excessive suit owing to plaintiffs' bearing only a part of social costs will be reinforced. If the private benefit falls short of the social benefit, however, there may be too little incentive to bring suit. Similarly, the private benefit from a particular legal expenditure (developing a new legal argument, hiring an expert, whatever) is generally different from the effect that the expenditure will have on deterrence. Likewise, the private advantages of trial are likely to be different from the social benefits of trial; these benefits are associated with, among other factors, the deterrence created by information gained from trial and public exposure of defendant misbehavior.

Although the divergence in the last paragraph concerns the social benefit of deterrence, a private-social divergence plainly may exist with respect to other social benefits of use of the legal system, too, such as compensation of victims of harm or the elaboration of the law through its interpretation and the setting of precedent. The reason is that private parties are not usually concerned, or are not exclusively concerned, with the social purposes of litigation, whatever may constitute these purposes; private parties are primarily concerned with their selfish benefits from litigation.

The general consequence of the possibility of a divergence between the private and the social incentive to use the legal system is that social intervention may be justified. Where private incentives to use the legal system are excessive, restrictions on its use may be desirable. Such restrictions should be expansively interpreted to include all manner of policies discouraging suit (including its outright prohibition in some areas), policies promoting settlement, and policies intended to reduce litigation expenditures (such as procedural rules limiting discovery, the time allowed for the filing of motions, or the number and qualifications of experts). Conversely, where private incentives to use the legal system are inadequate, encouragement of litigation may be beneficial, and policies to effect this should likewise be broadly viewed. As I will comment, however, our ability to control intelligently the amount of litigation will often be circumscribed by our imperfect knowledge of its effects.

Another conclusion that will be developed is that defendants should pay for more than just the harm that they cause in adverse events: they should incur a bill equal to harm plus total litigation costs. The reason is that, when an injurer causes harm and is sued, the true cost that society is forced to bear equals the harm plus total litigation costs. Accordingly, for injurers' incentives to reduce risk to be appropriately strong, they need to pay for more than only the direct harm that they cause.

The plan of this article is to study first the divergence between private and social incentives to bring suit; this is the subject of Sections II and III. Then the private-social divergence with respect to the decision about settlement versus trial will be investigated in Section IV, and the divergence with respect to the choice about litigation expenditures will be examined in Section V. Section VI concludes.⁷

⁷ Prior work on the private versus the social incentive to use the legal system has concerned mainly the decision whether to bring suit; see, originally, Steven Shavell, *The Social versus the Private Incentive to Bring Suit in a Costly Legal System*, 11 J Legal Stud 333 (1982); and further discussion in Louis Kaplow, *Private versus Social Costs in Bringing Suit*, 15 J Legal Stud 371 (1986); Peter S. Menell, *A Note on Private versus Social Incentives to Sue in a Costly Legal System*, 12 J Legal Stud 41 (1983); and Mark Geistfeld and Susan Rose-Ackerman, *The Divergence between the Social and Private Incentives to Sue: A Com-*

In the course of the analysis, I will be considering, among other questions, the following: Are there particular legal domains in which litigation activity is likely to be excessive? Where litigation is excessive, are there simple policies, such as legal cost-shifting schemes or imposition of fees for use of the courts, that would act as a corrective? Is it necessarily socially wasteful for parties to spend more on litigation than the amount at stake? Does the private-social divergence to use the legal system provide a rationale for legal aid programs? Is there warrant for the state to adopt a general policy of encouraging settlement (as it seems to have done)? How do the implications of economic analysis relate to notions of fair access to the legal system and to the belief that settlement may compromise litigants' basic right to trial?

Before proceeding, I should mention two limitations in the scope of this article. First, for the most part, I restrict attention to civil disputes, not criminal ones, because the decision whether to initiate legal action against a criminal defendant is vested mainly with the public prosecutor rather than with private parties.⁸ Second, I consider the use of the legal system and of legal services only in relation to legal disputes that have already arisen. I

ment on Shavell, Menell, and Kaplow, 16 J Legal Stud 483 (1987). There has been little published economic analysis of the private versus the social motive to settle (as opposed to the motive to sue); the literature on settlement (see note 48 below) has been predominantly descriptive in orientation. Moreover, the private versus the social incentive to spend on litigation has not been generally investigated, although aspects of it have been examined in Louis Kaplow and Steven Shavell, *Accuracy in the Assessment of Damages*, 39 J Law and Econ 191 (1996).

Perhaps because the literature on the private versus the social incentive to use the legal system analyzes formal models and is somewhat narrow in focus, its message appears to be largely unappreciated by legal academia and policy makers, and it even seems frequently to be overlooked within the law and economics community. For example, in the chapter devoted to litigation of Richard Posner's law and economics text, the entire issue of the private versus the social incentive to litigate is left essentially unconsidered; see Richard A. Posner, *Economic Analysis of Law*, ch 21 (4th ed 1992). Further, none of the recently published books emphasizing problems concerning the level of litigation—see the works by Epstein, Huber, and Olson (cited in note 3)—reflects real awareness of the private-social divergence under discussion. Also failing to recognize the basic elements of the private-social divergence is an often-cited article on the number of lawyers: Robert C. Clark, *Why So Many Lawyers? Are They Good or Bad?* 61 Fordham L Rev 275 (1992).

This article adds to the existing literature on the private-social divergence by considering the three types of litigation decision (about suit, settlement versus trial, and legal expenditures) in a unified way, by its attention to the private versus the social motive to settle, and by its consideration of corrective policy. Also, its informal and general approach should make it accessible to a wide audience. A formal analysis of many of the points in the article is contained in Steven Shavell, *The Level of Litigation: Private versus Social Optimality* (Discussion Paper No 184, Harvard Law School, John M. Olin Center for Law, Economics, and Business, June 1996).

⁸ Nevertheless, there are some lessons that can be learned about the criminal context from the analysis of this article, and I comment on them at the ends of Secs. III and IV.

do not consider the use of legal services prior to the occurrence of harm and disputes, notably, for the drafting of contracts and for advice about how law will apply to contemplated behavior. Use of legal services for such forward-looking purposes involves issues that are different from those that I will be analyzing.⁹

II. THE VOLUME OF SUIT

A. *Basic Theory: The Fundamental Divergence*

To understand the divergence between private and social incentives to bring suit, let us consider initially a stylized situation with the following features. Injurers may cause harm, the probability of which they (but not victims) may be able to reduce by taking precautions. If harm occurs, a victim can bring suit but will incur an expense to do so; further, the injurer will bear a defense cost in the event of suit, and the state will bear certain costs as well. The victim who sues will be supposed definitely to prevail¹⁰ and to collect a judgment equal to his losses (I am not yet taking the possibility of settlement into explicit account).¹¹

We will assume that the social objective is simply minimization of the sum of social costs: the harm from injury to victims, plus the costs of precautions, plus the costs associated with use of the legal system—these comprising victims', injurers', and the state's legal costs. (I will discuss later how the analysis would change if we were to consider other social objectives, including compensation of victims.)

Why the Volume of Suit May Be Either Excessive or Inadequate. The private decision to bring a suit rests on a simple comparison. Suppose that a victim has suffered harm of \$1,000. Because he will collect this amount if he brings suit, he will proceed if his legal costs would be less than his award of \$1,000; for example, he would sue if his costs of suit were \$300. The victim's decision about suit is based on a comparison of his own costs of suit with his benefit from suit—in the form of the judgment equal to the harm he has sustained.

⁹ On the private versus the social incentive to use legal services before litigation, see Louis Kaplow and Steven Shavell, *Private versus Socially Optimal Provision of Ex Ante Legal Advice*, 8 J L, Econ, & Org 306 (1992).

¹⁰ This would be the situation if strict liability governed. If the negligence rule applies, the thrust of the conclusions to be reached is still valid, but see the discussion of this rule in Sec. III.

¹¹ The assumption that settlement does not occur will allow us to understand the basic issues pertaining to the volume of suit in an uncluttered way, but the reader can, without real harm, interpret the victim's receipts as a settlement amount. In any event, I will take settlement into explicit account in Secs. III and IV.

From the social standpoint, however, the costs of a suit are, as mentioned above, higher than the private costs because the injurer will defend himself and the state will bear costs as well. Thus, if the victim's costs are \$300, the injurer's costs are \$200, and the state's costs are \$100, then the total social litigation costs associated with a suit are not \$300 but $\$300 + \$200 + \$100$, or \$600.

The social benefits of suit inhere in the deterrent effect that suit has on the exercise of precautions by injurers and thereby on the frequency of harm.¹² (More precisely, the social benefits of suit equal the reduction in expected harm net of the costs of precautions taken.) If injurers know that they will be sued and will have to pay \$1,000 plus their defense costs of \$200 when they cause harm, they may be led to exercise precautions. Whether they will do so will depend on the cost of the precautions and their effectiveness in reducing the risk of harm. If precautions are cheap and effective, the precautions injurers choose to exercise may result in a significant decline in the incidence of harm. If, however, there is little that injurers can do by taking precautions, at least at reasonable cost, they will take few precautions, and the frequency of harm will not change much as a consequence of the prospect of suit. It is thus evident that the social benefits of suit are determined in important ways by the cost and effectiveness of precautions in reducing risk—in other words, by factors quite distinct from the magnitude of harm, here \$1,000. Accordingly, it should not be surprising to the reader that the social benefits of suit could be either less than or greater than the private benefits, and thus that suit might be brought when it should not be, or that suit might not be brought when it should be.

To verify that both possibilities may obtain, and to gain insight into the private-social divergence, it will be helpful to consider several examples in which we can calculate how victims and injurers behave and the associated levels of total social costs.

Let us first consider a situation in which suit is excessive: suit occurs, but it would be better that it does not. Suppose that legal costs are as I stipulated above, that harm is certain to occur and that there is nothing that can be done to prevent it. Suit will plainly be brought because a victim's legal costs are \$300, which is less than the \$1,000 he will collect. Hence, social costs (per victim and injurer pair) will be \$1,000 plus \$600 in total litigation costs, resulting in a sum of \$1,600. In contrast, were suit not brought, total social costs would be only \$1,000, as no litigation costs would be incurred. The reason that the amount of suit is excessive in this example is obviously

¹² I am assuming for simplicity that the level of activity of injurers is fixed. Otherwise, liability-induced reductions in their activity levels might constitute a benefit of suit along with liability-induced increases in their levels of precaution.

that there is *no* deterrent effect of suit, no social benefit whatever associated with suit.

What is the conclusion, however, if there is a positive deterrent associated with suit and suit induces an economically “efficient” precaution? To show that the level of suit still may be socially excessive when this is so, modify the example by assuming that the likelihood of harm will be reduced from a certainty to 75 percent by a precautionary expenditure of \$150; thus, the precaution brings about a 25 percent reduction in the probability of harm. Note that the precautionary expenditure is socially worthwhile, or economically efficient, because its \$150 cost is less than the expected reduction in harm that it accomplishes, namely, 25 percent \times \$1,000 = \$250. In this case, the possibility that suit will be brought will have an effect on injurers. They will be led to spend the \$150 on the precaution: whenever an accident occurs, they will bear expenses of \$1,200, which is their judgment plus litigation defense costs, and the expected reduction in these costs will be 25 percent \times \$1,200 = \$300, which exceeds \$150. Total social costs will be \$150 + 75 percent \times \$1,600 = \$1,350, for \$150 is spent on precautions, harm of \$1,000 occurs with probability 75 percent, and, when that happens, since suit is brought and total litigation costs are \$600, the social costs incurred are \$1,600. This means, though, that suit is not socially worthwhile, for if there is no suit, social costs are only \$1,000, which is less than \$1,350. The reason that suit is not socially worthwhile in this version of the example is that the net deterrent benefit of suit is not sufficient to offset the total litigation costs of suit.

If we alter the example again such that deterrence is more effective, we can readily produce a case in which suit is socially worthwhile, so that it is both brought and ought to be. For instance, let the reduction in the probability of harm due to the precautionary expenditure be twice as much as before, from 100 percent to 50 percent. Then it may be verified that injurers will spend the \$150 on the precaution, in which case social costs will be \$150 + 50 percent \times \$1,600 = \$950. Since \$950 is less than \$1,000, suit is socially worthwhile.¹³

Let us now illustrate the possibility that the level of suit might be socially inadequate rather than excessive: suit might not be brought when it would be socially advantageous for it to be brought. Suppose that the magnitude of harm is not \$1,000 but only \$100. Then suit will not be brought because a victim would not find it worthwhile spending \$300 to gain only \$100. Social costs will therefore be \$100 because the harm of \$100 will occur

¹³ In this example, suit became worthwhile because the precaution reduced risk to a greater degree; more generally, suit might also become worthwhile because of a lower cost of the precaution.

with certainty. Suit might be socially worthwhile, however, if its prospect would lead to an inexpensive and substantial reduction in the risk of accidents. Assume that by spending a mere dollar on a precaution, injurers would lower the risk of harm from certainty to only 10 percent. In this situation, were suit to be brought, injurers would be induced to take the 1 dollar precaution, and social costs would thus be only $\$1 + 10 \text{ percent} \times \$700 = \$71$ (note that when harm of \$100 occurs, litigation costs are \$600, so total costs are \$700). Since \$71 is less than \$100, it would be socially desirable for suit to be brought. The reason that suit is not in fact brought, however, is that a victim sees as his reward from suit only the \$100 judgment; he does not take into account that his willingness to bring suit would produce a socially valuable incentive for injurers to take precautions.

The reader should now better appreciate that in principle there is no necessary connection between the bringing of suits and the social value of suits.

Externalities and the Interpretation of an Incorrect Volume of Suit. One way of explaining the conclusion that the volume of suit may be socially inappropriate is to observe that the market for legal services for suit is affected by externalities. This market involves both negative externalities—because suit imposes costs on injurer-defendants and on the state—and positive externalities—because the prospect of suit creates deterrence of unwanted behavior. The existence of these externalities means that no matter how competitive and well-functioning the market for legal services may be, we cannot have confidence that the volume of suits tends toward the socially desirable level. The market for legal services for suit is quite unlike the market for, say, plumbing services. When a person hires a plumber to replace a leaking pipe in his house, there is usually no external cost imposed or external benefit conferred on others, so that we can say that the amount of plumbing services determined by market forces is approximately socially rational.

A Misconception about Suit—That Suit Is Not Justified If Litigation Costs Exceed the Dollar Amount at Stake but Is Justified Otherwise. Our analysis of the volume of suit also sheds light on a view that is often encountered, namely, that suit is socially irrational if litigation expenditures exceed the amount at stake but rational if litigation expenditures are less than the amount at stake. Although this view is appealing, involving as it does a simple comparison of litigation costs with the potential dollar benefit from litigation, it is fallacious.

To amplify, it is perfectly possible for more to be spent on suit than the amount in question and yet for suit to be socially desirable. Indeed, it may be desirable for the state to encourage litigation even when total litigation costs exceed the amount at issue. This was precisely the situation in the

example where harm was \$100. In that example, it was advantageous for the state to stimulate litigation even though total litigation cost in each suit was \$600, fully six times the amount at stake.¹⁴

How can one explain the point that it may be socially desirable for more to be spent on suit than the amount at stake? The answer is that, on one hand, spending on suit may create substantial deterrence and that, on the other, the costs of suit are borne only when harm occurs, so that the expected costs for society may be much smaller than the realized costs in the event of a suit. In the example where harm is \$100 and litigation costs are \$600, recall that the prospect of suit induces injurers to spend 1 dollar on a precaution that reduces the probability of an accident from certainty to only 10 percent. This deterrence benefit is substantial; it is worth 90 percent \times \$100 = \$90; net of the precaution cost of \$1, it is worth \$89. And what is the expected social cost of litigation? It is only 10 percent \times \$600, or \$60. The fact that *when* suit occurs it costs \$600 does not tell us *how often* it occurs. The view that it is socially irrational to spend more than the amount at stake on litigation is wrong because it overlooks both the social benefit of suit and the fact that the costs of suit must be discounted by the probability with which suit is brought and costs are actually incurred.

At the same time, the notion that it is socially rational for litigation expenditures to be made as long as their total is less than the amount at stake is mistaken. That is, it may be socially undesirable for litigation expenditures to be incurred on suit even though they are less than the amount at stake. This was the situation in the first several examples I presented, where the sum of litigation costs was also \$600 but the harm was \$1,000 and suit was brought even though it was undesirable. Suit was undesirable because it did not result in any, or much, reduction in the likelihood of harm. Again, the error in the view that litigation expenditures should be compared to the amount at stake in deciding on their social rationality is that the amount at stake cannot be equated with, and indeed is only very loosely related to, the social benefit from litigation.

Why the Level of Precautions Tends to Be Inadequate. An issue that we have not yet discussed concerns the precise level of precautions that injurers

¹⁴ It is also easy to supply an example in which private parties have an incentive to bring suit, total litigation costs exceed the amount at stake, and suit is socially desirable. Consider a modification of our example in which the harm caused is \$1,000: let the litigation costs of the victim, the injurer, and the state each be \$400, so that total litigation costs are \$1,200, and assume that the exercise of the \$150 precaution lowers the likelihood of harm to 20 percent. In this case, victims will bring suit, because \$400 is less than the \$1,000 that they will win, and, as a consequence, injurers will be induced to exercise the precaution. Thus, social costs will be $\$150 + 20 \text{ percent} \times \$2,200 = \$590$, which is less than \$1,000, social costs if suit were not brought. Hence, the bringing of suits is eminently worthwhile even though total litigation costs are \$1,200 and exceed the harm of \$1,000.

are led to take by the prospect of suit. As a general matter, injurers' suit-related incentive to take precautions will be too low. This can easily be understood from the example in which harm was \$1,000. In that situation, each time that an injurer is sued, he causes not only direct harm of \$1,000, but he also induces expenditures all around of \$600 on litigation; thus, he engenders \$1,600 in total social costs. But the injurer bears costs of only \$1,200—damages of \$1,000 plus his own litigation costs of \$200. Because, then, injurers bear less than the true social costs of accidents, they might not be induced to exercise a properly high level of precautions. To illustrate, reconsider the version of the example where the precaution cost was \$150 and it reduced the accident probability to 50 percent. Suppose that there is another precaution that could further reduce the probability to 45 percent and that it costs an additional \$75. An injurer would not take this extra precaution because its expected benefit to him would be only 5 percent \times \$1,200 = \$60. However, it would be socially worthwhile for him to take the added precaution, for 5 percent \times \$1,600 = \$80, which exceeds the \$75 cost. The point, then, is just a special case of the general proposition that for incentives to take precautions to be appropriate, the injurer should pay for the full measure of social costs that he causes, and here that means that he should pay for the immediate harm plus total litigation costs.¹⁵

B. *Basic Theory: Corrective Social Policy*

Given that the volume of suit and the level of precautions are generally socially inappropriate, the question arises of how to rectify them.

The Volume of Suit. It should be straightforward in principle for the state to remedy an imbalance between the privately determined and the socially best level of litigation. If there is excessive litigation, the state can discourage it by imposing a properly chosen fee for bringing suit or by some other device to make suit more expensive; the state could also refuse to allow unwanted categories of suit to be brought. If there is inadequate litigation, the state can subsidize or otherwise encourage suit.

However, the state requires a great deal of information to be able to assess the socially correct volume of suit. As we have seen, to determine whether suit is socially desirable, the state must ascertain not only the costs of litigation for both sides but also the deterrent effect of the prospect of suit. This means that the state needs to deduce the nature and cost of the

¹⁵ It should be observed that the point I am making here applies whether or not the suit is brought when that is socially optimal. Even if suit is brought when it should not be, it is still the case that causing harm leads society to incur litigation costs. Thus, the level of precautions should still reflect the harm plus total litigation costs.

opportunities for preventing harm. I will comment in Section III below on the practical ability of the state to obtain such information.

Policies Wrongly Thought Generally to Improve the Volume of Suit: Making Plaintiffs Pay the State's Litigation Costs, and Loser Pays Fee-Shifting. It follows from what has just been said that two policies that are popularly suggested as cures for an improper volume of suit cannot be taken to be so in any general sense. The first policy is making those who sue pay for the state's litigation costs, on the ground that it is economically rational for a party to have to purchase the services that he uses.¹⁶ It is true, of course, that society would usually want a person to pay for the cost of provision of any service that he enjoys. If a farmer uses the services of a government veterinarian, we would want him to pay for the veterinarian's time, for then the farmer will use veterinary services when and only when he places a value on them exceeding their cost; in other words, he will use the services when and only when that creates social surplus. But if a person uses government services in the course of a lawsuit—by using the judge's time, the jury's time, a room in the courthouse, and so forth—the situation is different from that of the farmer. As I have emphasized throughout, there are external effects associated with suit. Making the victim pay for the state's litigation costs would only partially address the cost externality of suit because it omits the injurer's legal expenses, and such an approach would not reflect the positive externality due to the deterrent effect of suit. More to the point, we know that the policy of making plaintiffs pay for the state's costs cannot be desirable in all cases at least because it might discourage suit when suit would be socially beneficial.

The second policy that many find appealing is two-way fee-shifting, under which the losing party pays the fees of the winning party. This policy is sometimes advanced on the theory that it is only right that someone who has been shown wrong in court should pay the other side's legal expenses.¹⁷ The question for us to consider, however, is whether loser pays fee-shifting will have a beneficial effect on the volume of suit.¹⁸ In the examples and the analysis that I presented above, fee-shifting would simply serve to encourage litigation, for victims—who we have assumed have good cases—would always be able to fob off their litigation costs on injurers. Thus, fee-

¹⁶ For discussion of this economic argument, see, for example, Martin D. Beier, *Economics Awry: Using Access Fees for Caseload Diversion*, 138 U Pa L Rev 1175 (1990); Rex E. Lee, *The American Courts as Public Goods: Who Should Pay the Costs of Litigation?* 34 Cath U L Rev 267 (1985); and Richard A. Posner, *The Federal Courts* 136 (1985).

¹⁷ For a discussion of this fairness rationale, see, for example, Thomas D. Rowe, *The Legal Theory of Attorney Fee Shifting: Critical Overview*, 1982 Duke L J 651, 653–57.

¹⁸ Not of concern to us here is that fee-shifting also has effects on the propensity to settle and on the incentive to spend on litigation.

shifting would tend to worsen any problem of excessive suit. Moreover, in models where those who bring suit do not win with certainty, it can also be demonstrated that there is no general basis for a policy of loser pays fee-shifting. This is not to deny that fee-shifting can have socially desirable effects in particular circumstances. For instance, in the example where it was desirable for the state to subsidize suit, fee-shifting would equally induce suit; and in situations where some individuals would bring unmeritorious suits to extract settlements, fee-shifting might be desirable because it would discourage suits that are unlikely to succeed. But we are here examining the issue whether there is a systematic advantage of fee-shifting in regard to controlling the volume of litigation, and there is not.

The two policies just considered thus cannot serve as general correctives for problems with the volume of suit. Furthermore, a little reflection reveals that there does not exist any simple policy tool, any “magic bullet” for achieving that purpose, because appropriate social policy depends inherently on assessment of the deterrent effect of suit, and this is intrinsically complicated.

The Level of Precautions. To induce injurers to exercise the proper level of precautions, a level that reflects the full measure of social costs that are incurred when they cause harm and are sued, the amount that they pay must equal the direct harm that they cause plus the sum of all litigation costs. A straightforward way to ensure this result is for the court to impose a tax on injurers equal to the sum of the victim’s litigation costs and the state’s litigation costs (in our previous example, the tax would equal $\$300 + \$100 = \$400$). Because injurers also bear their own litigation costs, use of the tax would imply that injurers’ total payments equal the harm plus all litigation costs.

Of course, another way to achieve the same result is for injurers to reimburse victims for their legal expenses and to pay a tax equal only to the state’s litigation-related expenses (or to pay these as well to victims). Note, however, that such a policy would increase victims’ incentives to bring suit and thus would usually affect the corrective policy needed to control the volume of suit.

C. The Practical Relevance of the Divergence between Private and Social Incentives and the Possibilities for Corrective Social Policy

Let us now consider the extent of the divergence between private and social incentives to bring suit and the policies that might realistically be employed to correct problems with the volume of suit.

Divergence in Private and Social Costs and Benefits. The difference in private and social costs of suit is often large, at least in percentage terms.

The private cost divergence is that victims do not take into account injurers' and the state's litigation costs, as the reader knows. Thus, it is not unreasonable to expect that victims may fail to take into account around half of total litigation costs.¹⁹

The difference between the private and the social benefits of suit can also be substantial. First, many harms are large in magnitude and give the victim significant incentives to sue, yet deterrence effects may be relatively small for a variety of reasons. To illustrate, let us consider two important areas of litigation: automobile accident litigation and product liability litigation.

With regard to automobile accidents, we know that harms are sufficient to generate a tremendous volume of suit: it is estimated that they constitute about 50 percent of all tort litigation.²⁰ However, intuition suggests that liability-related deterrence of these accidents may be modest. Individuals have good reasons not to cause automobile accidents apart from wanting to avoid liability: they may be injured themselves, and they face fines for traffic violations and also serious criminal penalties for grossly irresponsible behavior (drunkenness, excessive speed). Given the existence of these incentives toward automobile accident avoidance, and given that the deterrent due to liability is dulled by ownership of liability insurance, one wonders how much the threat of tort liability adds to deterrence. The data on the deterrent effect of tort liability on automobile accident rates—obtained mainly by comparing accident rates where tort liability governs to that where it does not and no-fault statutes apply—are contradictory. Some studies show little or no effect of tort liability; other studies claim that elimination of tort liability would increase accident rates, perhaps by 15 percent or more.²¹ If we assume, for example, a 7.5 percent deterrent effect, elimination of tort liability for automobile accidents would have meant an increase in total accident costs of about \$3.75 billion in 1985.²² What would such a change to a no-fault regime for automobile accidents have saved? It is estimated that transaction costs account for about half of injurer payments under the tort sys-

¹⁹ For example, in 1985, victims' fraction of total litigation expenses in automobile litigation was about 54 percent (so they failed to take into account 46 percent of litigation expenses), and their fraction of total litigation expenses in nonautomobile tort litigation was about 42 percent; see Deborah R. Hensler, Mary E. Vaiana, James S. Kakalik, and Mark A. Peterson, *Trends in Tort Litigation* (Rand Corp 1987).

²⁰ See James S. Kakalik and Nicholas M. Pace, *Costs and Compensation Paid in Tort Litigation* (Rand Corp 1986).

²¹ See Don Dewees, David Duff, and Michael Trebilcock, *Exploring the Domain of Accident Law* 22–26 (1996).

²² In 1985, the total costs of automobile accidents in the United States were estimated in Leonard Evans, *Traffic Safety and the Driver* 7 (1991), to be about \$50 billion, and 7.5 percent \times \$50 billion is \$3.75 billion.

tem.²³ Significant reductions in these transaction costs would be expected to result under a no-fault system, and it is not hard to imagine that the savings in 1985 would have exceeded \$3.75 billion.²⁴ While this comparison is the result of rough guesswork, it does suggest that our litigation-related expenditures on automobile accidents might not be repaid by the social benefits of deterrence.²⁵

In the area of product liability, it also appears plausible that the incentives created by the prospect of liability are not substantial in relation to litigation costs. Whether or not they will be held liable, firms do not want their products to harm their customers because, if this occurs, firms will tend to lose business and/or have to lower their prices as a consequence. It is true, admittedly, that this inducement toward safety relies on the assumption that consumers will learn about the true risks of different product defects, whereas consumer information is almost never perfect. But it may well be that consumer information is often tolerably good and that, on net, the marginal deterrence engendered by the threat of product liability is not great. Although study of the effect of product liability on accidents is surprisingly sparse, it is not inconsistent with this hypothesis of low deterrence.²⁶ The litigation costs of product liability are, however, high. Therefore, as with automobile accidents, it appears conceivable that the private

²³ See Dewees, Duff, and Trebilcock, at 36 (cited in note 21) and Steven Shavell, *Economic Analysis of Accident Law* 263 (1987).

²⁴ For example, if victims were compensated for \$10 billion of their \$50 billion of losses in 1985 from the tort system, the implied magnitude of transaction costs would be \$10 billion. If half of these costs would be averted under a no-fault system, the savings would be \$5 billion, exceeding \$3.75 billion. In fact, the data on the reduction in transaction costs that would follow from adoption of automobile no-fault regimes are not clear. In Dewees, Duff, and Trebilcock, at 36–38 (cited in note 21), figures for the savings from automobile no fault ranging from nothing to 80 percent of expenses are noted. More generally, we know that the transaction cost of first-party insurance may be quite low: for some government programs, it is only several percent; see, for example, Shavell, at 263 (cited in note 23). To estimate the savings from adoption of no-fault, we would have to combine estimates of the reduction in transaction costs with data on the fraction of individuals who are presently compensated under the tort system. We would also have to estimate the value of the time that individuals would save by avoiding litigation.

²⁵ Note also that, even if it turns out that suit is socially worthwhile for the entire group of drivers, it may well still be the case that suit is not socially worthwhile for some easily identified subpopulation of drivers, such as drivers who are over the age of twenty-five and who have a good driving record.

²⁶ For example, in one of few empirical studies on product liability and deterrence, George Priest (*Products Liability Law and the Accident Rate*, in Robert Litan and Clifford Winston, eds, *Liability Perspectives and Policy* 184 (1988)), examined aggregate statistics on accident and fatality rates; he found no evidence that the amount of litigation activity influenced injury or death rates. Also, Dewees, Duff, and Trebilcock, at 205 (cited in note 21), conclude their survey of empirical literature on product liability by saying that there is little evidence that strict product liability has brought socially desirable safety gains.

incentive to sue has resulted in an excessive volume of suit, and study of this issue seems warranted.²⁷

The opposite possibility, that the volume of suit is socially inadequate, is also of practical significance. Recall that I illustrated this outcome with an example in which an individual's losses were low, \$100, so that suit would not be brought, but in which the frequency of harmful events could be fairly cheaply reduced. This example seems to be of relevance. One can readily imagine situations in which firms know that the harms that they cause will not be of sufficient importance to be worth a typical victim's while to pursue, even though the incidence of the harms can be decreased substantially by modest expenditures. (Consider low-level pollution damage, such as more frequent peeling of paint in a neighborhood near a factory, and which the factory could eliminate by installing inexpensive smoke scrubbers.) One can also envision situations in which, even though the magnitude of harm might be high, the expected value of suit is still low because of difficulty in proving causation. (Suppose the pollution from the factory can produce cancer, but its etiology is hard to demonstrate.) If, once causation were established, many other suits could easily be brought, then it might be socially valuable for suit to be filed in the case at hand even though that would not be advantageous to the plaintiff. Another example is, perhaps, certain civil rights litigation where the combination of low monetary damages and difficulty in proving liability may discourage suit even though the effect of prevailing might be socially significant and justify it.

Corrective Policy: Volume of Suit. The actual use of policies to correct the volume of suit in broad categories of case appears to be quite plausible. If serious study of the automobile and product liability evidence points in the directions that I suggested, I could imagine imposition of fees on plaintiffs for bringing suit, or even a ban on suit in these areas. Similarly, if we concluded that we ought to encourage suit for certain types of low-magnitude harms for the reasons outlined above, one could also envision the adoption of subsidies or loser-pays fee-shifting to stimulate legal action. Such policy responses are fairly easily envisaged because they would not be difficult to implement and because we have observed their use before. On one hand, worker's compensation legislation²⁸ and automobile no-fault statutes have removed important categories of accident from the domain of

²⁷ Note that the argument of this paragraph does not apply with regard to product liability provisions that the parties write into their contracts, but only to court-mandated product liability. Note also that the argument obviously does not apply to harms to parties who are not injuring firms' customers (such as people who live near a factory and who are harmed by an explosion), for in the absence of liability, firms will have no financial incentive to avoid causing them loss.

²⁸ See, for example, Dewees, Duff, and Trebilcock, at 387–94 (cited in note 21).

tort law,²⁹ New Zealand has barred suit for all cases of personal injury,³⁰ certain jurisdictions have shifted legal fees to losing plaintiffs to discourage litigation, and so forth. On the other hand, as mentioned, legal aid programs have been employed to subsidize litigation (see below), statutes have authorized the award of multiplied damages to spur certain types of litigation, and legal fee-shifting has been permitted in some areas for the same reason.³¹

Whether, however, judges could reasonably decide to discourage or encourage suit in individual cases is another matter. To do this, judges would need to obtain information about, or gain a fairly good intuitive understanding of, the deterrence effect of bringing a specific type of case. One suspects that judges would not usually be able to evaluate deterrence on such a refined level. Unless they could, the most likely type of policy that could justifiably be employed to correct the volume of suit would be legislative actions applying to whole categories of case.

Comment on Legal Aid. Legal aid programs generally require that the victim be sufficiently poor to receive assistance in bringing a case.³² But the fact that a person is poor does not, in itself, imply that he has an inadequate private incentive to bring suit. For example, a poor victim of an automobile accident might have a strong incentive to bring suit because of the amount of damages he could collect. In this situation, the contingency fee system furnishes lawyers with a financial reason to take the victim's case even though he might not be able to afford to pay legal fees in advance. Moreover, because such a suit might have a low social benefit, as I suggested, were the help of legal aid needed to bring the suit, it is hardly clear that the suit would be socially advantageous. It is true, though, that some types of case that legal aid programs may be likely to bring could find some justification in the private-social divergence in incentives to bring suit. I have mentioned that the private incentive to bring suit may be inadequate where harm is low or where suit might have a substantial symbolic effect. Legal aid programs are often said to favor cases of this type and to that extent may be acting socially beneficially to encourage suit. Still, legal aid pro-

²⁹ See, for example, Robert H. Joost, *Automobile Insurance and No-Fault Law* (2d ed. 1992); and Stephen J. Carroll, James S. Kakalik, Nicholas M. Pace, and John L. Adams, *No-Fault Approaches to Compensating People Injured in Automobile Accidents* (1991).

³⁰ See, for example, Geoffrey Palmer, *New Zealand's Accident Compensation Scheme: Twenty Years On*, 44 U Toronto L J 223 (1994).

³¹ This is hardly to say that these policies have been adopted for the reasons I have discussed. For example, it seems that workers' compensation legislation was enacted primarily because workers were not able to obtain compensation under the tort system, not because the costs of the tort system outweighed its deterrence benefits.

³² I am considering here only the role of legal aid for plaintiffs.

grams appear problematic as social policy because the wealth status of a victim would seem to be only incidentally related to the presence of inadequate private incentives to bring suit.³³

Corrective Policy: Level of Precautions. Correction of the incentives of injurers to take precautions (as opposed to correction of victims' incentives to bring suit) is a goal that can fairly readily be met because the informational burden of the courts would not be great. Courts would not need to ascertain the incentive effect of suit; they would need only to estimate the legal costs of plaintiffs and of the state and add these to determine the tax the defendant must pay. While this task would not always be easy, it does not seem especially difficult and is something courts already occasionally undertake.³⁴ It should also be remarked that the policy under discussion does not give plaintiffs a perverse incentive to spend on legal services: because the policy is a tax paid by the defendant to the state, meaning that the plaintiff still bears his own legal expenses, the plaintiff will naturally want to limit them.

Not only is it feasible to impose a tax on defendants to correct their incentives to take precautions, it also seems important to do that because of the magnitude of plaintiffs' and the state's litigation-related costs. Indeed, a rough estimate is that they could average about a third of harm. Recall that studies suggest that total litigation expenses are in the neighborhood of what plaintiffs receive from defendants. If this is so and we suppose that plaintiffs' and the state's expenses are about half of total litigation expenses, then they must be about one-third of harm.³⁵ Society should rather clearly want to include in the financial penalty of injurers a component of social harm that is as important as one-third of what we presently say to be

³³ I have not mentioned two justifications for legal aid that are often advanced. One is that it is a form of income redistribution and therefore is a good thing. While income redistribution is socially valuable in the opinion of many, and is included in the conventional social welfare calculus of economists, they generally argue that the best way to redistribute income is through the income tax and transfer systems, not through subsidy of the purchase of particular goods. (One way of explaining why is to observe that the poor will be better off if given a cash grant equal to the amount spent on legal services in their behalf: the poor could then exercise the option of spending the cash grant on other goods that they might prefer, like food and shelter.) Thus, the desire to redistribute income cannot serve as a sound rationale for legal aid.

The other common justification for legal aid is that all people have a basic right to access to the courts, and to guarantee the poor this right they need legal aid. To this, my response is essentially what I say below in Sec. III about that right.

³⁴ For example, courts need to determine legal costs when they engage in fee-shifting.

³⁵ To illustrate, suppose that a plaintiff receives \$10,000 and that total litigation costs are \$10,000. If the plaintiff's litigation costs are half of the total, his litigation costs are \$5,000. Since what the plaintiff receives is harm minus his litigation costs, harm must have been \$15,000—for \$15,000 minus \$5,000 yields \$10,000, what he received. Hence, \$5,000 represents one-third of harm.

the harm; otherwise we substantially dilute the proper incentive of injurers to take care. To put the point a little differently, could we imagine agreeing to exclude from damages a direct component of harm that amounted to one-third of its magnitude?

Socially Perverse Interest of the Bar, and Possibly of Legal Academia, against Curtailment of Suit. A final point bearing on the possibility of the use of policies to correct the volume of litigation concerns the interest of the bar and possibly of legal academia. The bar has a strong economic interest against social policies that would curtail demand for legal services. Thus, where there is an excessive private incentive to bring suit and the called-for social policy is to discourage suit, we can expect the bar to resist.³⁶ The bar can not only exercise its considerable political power, it can also make two appealing types of argument: that individuals have a basic right to access to the courts, and that discouraging suit would constitute interference with the natural forces of supply and demand in the market for legal services. I would also expect many legal academics to argue against social policies discouraging suit because they tend to view access to the courts as a fundamental right and to regard compensation as a social purpose of suit rather than of the insurance system (on which, see below).

III. EXTENSIONS OF THE ANALYSIS

I consider here a number of comments on and elaborations of the analysis of the volume of suit that will help to round out our understanding of it.

Goals of Suit apart from Deterrence. How would goals of suit apart from deterrence affect our analysis? One such goal of course is *compensation* of victims. Perhaps the most important point to make about compensation is that it can be much more efficiently distributed by the insurance system (including the social insurance system) than by the legal system.³⁷ Accordingly, had I included compensation as a social goal, we would not have been led to conclude that the socially appropriate volume of suit would be much different from what I identified it to be.

Nevertheless, suppose that we ignore the presence of insurance and view the legal system as the sole means of alleviating risk. That is, suppose that we treat the allocation of risk between victims and injurers as a social goal of the legal system. How would the private and the social incentives to use the legal system differ with respect to this social goal? The answer is that

³⁶ Of course, where the problem is an inadequate level of suit, I would not expect there to be any problem garnering support from the bar for corrective policy.

³⁷ The average cost of furnishing each dollar of compensation to victims through the tort system appears to be approximately 1 dollar, whereas it is typically less than 20 cents under the insurance system; see, for example, Shavell, at 263 (cited in note 23).

victims' private incentive to use the legal system to alleviate their risk would exceed the social incentive.

To understand this claim, it is clarifying to consider a situation where injurers have no ability to affect accident risks, so that deterrence is not at issue. Suppose, for instance, that there is a 10 percent risk of a victim suffering a \$1,000 loss and that there is no way injurers can alter this risk. What will be the private incentive to sue, and will it be socially appropriate? We know that a victim will sue as long as his own litigation costs are less than \$1,000, and regardless of the magnitude of the injurer's and of the state's litigation costs. Thus, the private behavior of victims leads in effect to a system whereby large sums, possibly exceeding \$1,000, might be spent to deliver compensation of \$1,000 to victims (note, too, that this is true even if victims are not very averse to risk). Socially, however, it would not be rational for what is effectively an insurance system to be established to compensate the \$1,000 losses unless the total litigation costs are sufficiently small.³⁸

Furthermore, it is narrow to view compensation of victims as the only socially relevant aspect of the allocation of financial risk: injurers also may be risk averse, meaning that the risk that suit imposes on injurers must be entered into the social welfare calculus. But victims do not recognize this social cost when they bring suit. In sum, then, victims have an excessive incentive to use the legal system with regard to the social goal of the proper allocation of risk due to failure to take adequately into account both litigation costs and the possible risk aversion of injurers.

Another potential goal of suit is to foster the *amplification of law* through its interpretation and the setting of precedents. In considering this goal, two factors should be borne in mind. First, of course, because the great majority

³⁸ To illustrate, suppose that injurers are risk-neutral manufacturers of a product who pass on their liability costs to victim-consumers in the form of higher prices. Assume that, if the product fails, it will cause harm of \$1,000 to a consumer, and that he will incur \$300 in litigation costs to collect \$1,000 in damages and the manufacturer will absorb \$200 in litigation costs. Therefore, the product price will rise by $10 \text{ percent} \times \$1,200 = \$120$, so that victims in effect will be paying a premium of \$120 for an insurance policy that gives them \$1,000 with probability 10 percent and involves their bearing a collection cost of \$300 in the event of an accident—meaning that their coverage is really \$700. They would not want such an insurance policy unless they were fairly risk averse. Thus, the best social decision might be to bar suit in this case. To continue, suppose that victims had to spend \$999 in litigation expenses to collect \$1,000. They would still spend this amount even though they would be made worse off by their ability to sue, no matter how risk averse they are: for their “premium” of \$120 in the product price would now purchase for them only \$1 in coverage. In this case, victims would definitely be better off were suit barred. The general point under discussion here, that the private incentive to sue is excessive in relation to the social goal of insuring risk-averse victims, is closely related to the conclusion of Louis Kaplow, *Optimal Insurance Contracts When Establishing the Amount of Losses Is Costly*, 19 Geneva Papers on Risk and Insurance Theory 139 (1994).

of cases are resolved short of trial—apparently over 95 percent³⁹—they cannot result in development of the law. Second, of the cases that do go to trial, many do not result in changes in law, as is noted below in Section IVB. Although these factors limit the general importance of the amplification of law as a social reason for suit to be brought, it may still be of substantial significance in some areas of adjudication, especially where the law is in flux.

An additional goal of suit concerns the utility that parties obtain from *participation in legal proceedings*, from speaking their piece, explaining themselves.⁴⁰ The importance of this consideration is also reduced by the extent of settlement, for settlement means that a full trial is not held. In any event, whatever is its importance, the utility that individuals derive from participation in legal proceedings is not an apparent source of divergence between private and social incentives to use the legal system because, one presumes, potential litigants will take all benefits that they might obtain from litigation into account in deciding whether to bring suit. In particular, therefore, the value of participation in legal proceedings does not provide a reason for the state to encourage suit.

Two final, and oft-asserted, goals of suit are to *avert individuals from engaging in self-help* and to *avoid the development of undesirable social attitudes*. Allowing parties to sue makes it unlikely that they will resort to socially destructive forms of self-help against injurers; and because suit means that victims' grievances will not fester unsatisfied, victims of harm will not develop attitudes of disrespect for the law and government generally, which can have real and negative repercussions for society.⁴¹ To the extent that these benefits of suit are important, more suit than otherwise becomes socially valuable.

But how much self-help would there really be, and would significant antisocial attitudes in fact be likely to emerge, in the absence of the opportunity to sue? A victim's desire for self-help seems to be mitigated by two factors. First, because individuals are often insured against loss (and are sometimes provided with social insurance), they already receive compensa-

³⁹ Recent data on state courts show that in fiscal year 1992, over 96 percent of civil cases were settled or otherwise disposed without a trial; see Ostrom and Kauder, at 28 (cited in note 1). Similarly, statistics on civil cases in U.S. District Courts during the 1995 fiscal year show that almost 97 percent of cases were resolved without trial; see *Judicial Business of the United States Courts, 1995 Report of the Director*, at 162.

⁴⁰ On the value of participation in legal proceedings and related benefits of proceedings, see, for example, Frank I. Michelman, *The Supreme Court and Litigation Access Fees: The Right to Protect One's Rights—Part I*, 1973 Duke L J 1153, 1173–77.

⁴¹ Examples of such negative repercussions are that citizens might not give information to the police about crimes that they witness and that they might decide against serving in socially valuable volunteer organizations.

tion for losses and thus would not need to engage in self-help to obtain recompense in the absence of the ability to sue. Second, because the state is able to prosecute injurers criminally, the types of actions that are likely to provoke a strong desire for punishment frequently do result in punishment; the inability to bring an action for civil damages would not alter this fact. Additionally, where victims' ability to sue has been eliminated, for instance, in the area of workplace accidents and in jurisdictions with no-fault automobile accident statutes, we have not experienced notable problems of self-help. These examples suggest that inability to sue does not necessarily spawn antisocial attitudes toward the legal system.

Settlement and the Costs of Suit. Let us now turn to a factor that was not considered in the previous analysis and ask how it would affect our conclusions. Settlement was, as mentioned at the outset of Section II, not explicitly introduced, but it is certainly relevant because, as mentioned, it is what usually occurs when suit is brought, and settlement means that the costs of trial are avoided. Thus, we might best interpret the litigation costs of the parties to be not trial costs but a properly weighted average of possible settlement costs and of possible trial costs. (Under this interpretation, it is important to note that the settlement costs of the state will still be positive because before reaching settlement, parties generate costs for the state by filing their claims, making motions about discovery, and the like.)

How does recognition of settlement alter our conclusions? Because the chance of settlement lowers effective litigation costs, the socially optimal volume of suit will be higher than would be true were all suits followed by trial. Also, because the possibility of settlement lowers the effective cost of bringing suit, it tends to increase the private incentive to bring suit. Therefore, it is not clear in general whether settlement increases or reduces the divergence between private and social incentives to sue, even though it raises the socially optimal volume of suit.

The possibility of settlement also affects the tax that should be imposed on injurers in order that their incentives to take precautions be correct: if there is a settlement, an injurer's tax should equal the sum of the victim's and the state's settlement costs (rather than trial costs), for then the injurer will be paying the true social cost of his having caused harm.

However, a policy under which an injurer who settles must pay a tax to the state raises a question. Would not the parties have an incentive to keep their settlement secret so that the injurer can avoid the tax? The answer is in the affirmative, for the injurer could share his savings from not paying the tax with the victim by giving him a higher settlement amount. To prevent secret settlements from undermining the tax, the state could adopt the simple policy of refusing to enforce settlements that were not accompanied by payment of the tax. Under this policy, no injurer would rationally pay

money in a secret settlement, for then the victim could still turn around and sue the injurer, since his settlement agreement would not prevent the victim from doing so.

Negligence Rule. It was generally assumed in the prior analysis that if a victim sued for harm, he would always prevail, in other words, that strict liability for harm governed. Under the negligence rule, according to which a victim will prevail only if the injurer was negligent, there are differences in the analysis that bear comment. To appreciate them, consider first a perfectly functioning negligence rule, in which courts never err in determining negligence. In this case, it is in principle socially desirable for victims *always* to bring suit for negligently caused harm, however odd this claim may initially strike the reader. That is, no matter how high the litigation costs or how small the harm, the state should subsidize suits for negligence which victims otherwise would not bring.

The rationale for the policy of subsidy is that, given our assumptions, there will never in fact *be* any suits for negligence. Therefore, no litigation costs will actually be borne, yet injurers will be led to act appropriately. Specifically, if injurers know that they will definitely be sued for negligently causing harm, they will be induced to act nonnegligently. And since our supposition is that courts will never make a mistake in assessing negligence, there would never be any findings of negligence, so that victims would never bring suit.⁴² Thus, it is socially desirable for victims always to be ready to sue for negligence, in order to induce nonnegligence on the part of injurers, in which case suits are not brought.⁴³

When, however, we take into account various complications in the working of the negligence rule—notably, imperfect ability of courts to determine optimal standards of behavior and to ascertain actual conduct, inability of injurers to control their momentary behavior, imperfect ability of firms to control the behavior of employees, and the judgment proof problem—we see that there will be findings of negligence and therefore suit. (Furthermore, even when injurers would not be found negligent, there may still be suit because victims might believe that injurers would be found negligent.) Because, then, use of the negligence rule may result in findings of negligence and suit, the situation becomes like that analyzed previously: there may be too much or too little suit, and the state should take appropriate corrective measures, either to discourage or to encourage its volume. Nevertheless, there is a sense in which the conclusions from the case of a perfectly functioning negligence rule apply. Namely, since nonnegligent be-

⁴² This assumes as well that victims can observe that injurers were nonnegligent.

⁴³ Under strict liability, by contrast, suits will be brought when harm occurs regardless of the degree of care exercised by injurers, so litigation costs will be incurred.

havior often means that suit will not be brought (or that a claim will quickly be dropped), the negligence rule does involve less actual suit than strict liability, and the problem of socially excessive suit may thus be of less significance under the negligence rule than under strict liability.

Lawyers as Imperfect Agents for Clients. Another factor that was not taken into account in the foregoing analysis is that lawyers may not act in the way clients would prefer. Such “agency” problems are different from the social-private divergence that is the subject of this article; even in the absence of any agency problems, there will very much be a divergence between social and private incentives to sue. Still, because agency problems can affect the private incentive to sue, they may influence the social-private divergence. For example, consider a lawyer who works on an hourly fee basis. He has an incentive to exaggerate to the victim the probability of winning at trial so that he will be hired (later he can drop or settle for a low amount). Therefore, more cases will result in suit on account of this agency problem. If the lawyer works on a contingency fee basis, however, he will tend to take fewer cases than our simple theory indicates.⁴⁴ These agency effects would need to be considered in a more detailed analysis than that undertaken here to determine the private-social divergence to bring suit.

However, it should be noted that agency problems alone are not a reason for a divergence between the socially desirable and the privately determined use of lawyers. In the absence of the externalities associated with use of the legal system under discussion in this article, the agency problem would not lead to any divergence between the private and socially desirable use of lawyers.⁴⁵

Scope of Civil Law for Which There Is a Divergence between Private and Social Incentives to Sue. The divergence between private and social motives to bring suit arises, as we know, because parties who decide to bring suit do not take into proper account the costs they impose or the benefits in deterrence that suit may bring about. Thus, the divergence applies quite broadly to civil litigation. The only possible exception concerns contracts because parties to a contract will be motivated to set the terms of their agreements to minimize their expected losses *plus* their joint litigation costs

⁴⁴ For example, the lawyer who will obtain a third of the award as a contingency fee but who bears the entirety of litigation costs himself will only take a case if its expected value is at least three times litigation costs, rather than whenever its expected value exceeds litigation costs.

⁴⁵ Agency problems make lawyers effectively more expensive and less efficient than they would otherwise be, but this will be taken into account in purchase decisions by clients, presuming that they are aware of the existence of the agency problem. In particular, then, there is no reason for the state to reduce the demand for lawyers because of agency problems alone.

(but not the litigation costs of the state). For example, a supplier and a buyer of machines might be concerned that the law would allow the buyer to sue for many modest problems and that this could result in their wasting resources on litigation and dispute settlement. To avoid these costs, they might state in their agreement that their modest problems will not be deemed breaches of their contract, or they might stipulate that their disputes will be inexpensively arbitrated. Nevertheless, many contracts will not be sufficiently detailed to include such remedies for problems of inappropriate suit, so that a private-social divergence to sue will still exist.

Public Prosecution of Criminal Cases. The private-social divergence in incentives to litigate also bears on criminal law, in that it helps explain why society cannot easily rely on private parties to bring criminal cases and in fact resorts to public prosecution. The private benefit from bringing a criminal case consists mainly in the satisfaction of seeing a criminal punished for his misdeed. The benefit does not involve monetary relief, for that is largely the province of tort law, and in any event many if not most criminals have only meager assets so could not pay much to their victims.

The *social* benefits of criminal prosecution are, however, frequently much higher, due to deterrence effects and the incapacitation of criminals. Hence, the social benefits of prosecution may substantially outweigh the private. Furthermore, were it left solely to private parties to decide whether to prosecute, criminals would be likely to engage in threats and violence to prevent their prosecution because the payoff would be great (even with public prosecution this can be a difficulty). In effect, the costs of private prosecution might become extremely high.

One suspects, therefore, that, were society to depend on private prosecution of criminal cases, the number of cases brought would be grossly inadequate. Public prosecution may thus be viewed as an answer to this latent problem.⁴⁶

⁴⁶ However, the history of prosecution may at first seem to contradict this opinion: English and American legal systems relied primarily on private prosecution of criminal cases until at least the nineteenth century; see Douglas Hay, *Controlling the English Prosecutor*, 21 *Osgoode Hall L J* 165 (1983); John Langbein, *Understanding the Short History of Plea Bargaining*, 13 *L & Socy* 262, 266–67 (1979). But there were also important features of these legal systems that encouraged private prosecution—and which may be regarded as confirming the general thesis emphasized in the text. First, to give victims incentives to prosecute, the state sometimes offered reimbursement for certain prosecutorial expenses (id at 171). Second, the state at times paid rewards to successful prosecutors; see John Langbein, *Shaping the Eighteenth Century Criminal Trial: A View from the Ryder Sources*, 50 *U Chi L Rev* 1, 105–114 (1983). Third, at least for a period, victims were able to obtain money payments from defendants in exchange for agreeing to drop their prosecutions (on which, see note 58 below). Fourth, collective private efforts were undertaken to encourage prosecution: people in a given locality frequently formed associations for the prosecution of felons, in which members pledged themselves to prosecute if victimized and their expenses were paid out of a fund financed by annual fees (Hay, at 171). Many of these points are stressed in David

IV. SETTLEMENT VERSUS TRIAL

Given that a victim has decided to bring suit, the injurer and he will decide either to settle or to proceed to trial. The great majority of cases settle, but the absolute number of trials constitutes a substantial residual.⁴⁷ The question that I want to address is, What is the private incentive to settle, and how does it contrast with the social incentive? In principle, then, we want to compare the circumstances in which victims and injurers will decide to settle with those in which it would be socially desirable for them to settle.⁴⁸

A. *Basic Theory*

We will consider here four important factors that bear on the private and the social desirability of settlement.

Costs of Trial. Because parties can save the costs of trial by settling, these costs represent a private motive for settlement. For example, if the victim's costs of proceeding to trial rather than settling are \$400 and the injurer's trial costs would be \$200, they would jointly save \$600 by settling.

Friedman, *Making Sense of English Law Enforcement in the Eighteenth Century*, 2 U Chi L School Roundtable 475 (1995).

⁴⁷ For example, if we assume that the proportion of civil cases that go to trial is in the 2–4 percent range (see note 39 above) and we apply these figures to the approximately 75 million noncriminal cases filed in 1994 (see Ostrom and Kauder, at 20 (cited in note 1)), we conclude that these cases generate between 1.5 million and 3 million trials.

⁴⁸ The economic literature on settlement versus trial is almost exclusively concerned with description of parties' bargaining behavior and with whether they settle; see, for example, William Landes, *An Economic Analysis of the Courts*, 14 J Law and Econ 61 (1971); John Gould, *The Economics of Legal Conflicts*, 2 J Legal Stud 279 (1973); Steven Shavell, *Suit, Settlement, and Trial: A Theoretical Analysis under Alternative Methods for the Allocation of Legal Costs*, 11 J Legal Stud 55 (1982); Lucian Arye Bebchuk, *Litigation and Settlement under Imperfect Information*, 15 Rand J Econ 404 (1984); the survey by Robert D. Cooter and Daniel L. Rubinfeld, *Economic Analysis of Legal Disputes and Their Resolution*, 27 J Econ Literature 1067 (1989); and Posner, at 554–560 (cited in note 7).

The only economic articles addressing the social desirability of parties' decision about settlement appear to be A. Mitchell Polinsky and Daniel L. Rubinfeld, *The Deterrent Effects of Settlements and Trials*, 8 Intl Rev L and Econ 109 (1988); and my recent paper, Shavell, cited at note 7; but Kathryn E. Spier, *Settlement Bargaining and the Design of Damage Awards*, 10 J L, Econ, & Org 84 (1994), comments on the issue in her n6 at 85.

There are, however, a number of articles in the traditional legal literature on the question of the private versus the social motive to settle, the best known of which is probably Owen Fiss, *Against Settlement*, 93 Yale L J 1073 (1984). See also Jules Coleman and Charles Silver, *Justice in Settlements*, 4 Soc Philosophy & Policy 102 (1985); Frank H. Easterbrook, *Justice and Contract in Consent Judgments*, 1987 U Chi Legal Forum 19; Marc Galanter and Mia Cahill, *Most Cases Settle: Judicial Promotion and Regulation of Settlement*, 46 Stan L Rev 1339 (1994); David Luban, *Settlements and the Erosion of the Public Realm*, 83 Ga L J 2619 (1995); and Judith Resnik, *Managerial Judges*, 96 Harv L Rev 376 (1982).

In general, the sum of the parties' legal costs constitutes a joint reason for settlement.

From the social perspective, however, the savings from settlement are larger than the sum of the parties' cost savings because society also avoids incurring the court's expenses. If in the example the court's trial expenses would be \$500, the full savings from settlement would be \$1,100 rather than only \$600. Thus, although the factor of trial costs creates a private motive to settle, it is smaller than the associated social reason to settle.

Risks of Trial. Parties may also want to settle if they are risk-averse, for by settling the parties avoid the risk of losing. If, for example, the odds of winning are 50 percent and the amount at stake is \$2,000, a risk-averse victim would prefer receiving \$1,000 for sure than taking the gamble of trial, and so would a risk-averse injurer prefer paying \$1,000 in settlement to taking the gamble of trial. Consequently, avoidance of risk is a reason why parties settle, independent of the outright cost savings that they can thereby achieve.

The avoidance of risk by risk-averse parties also is a social benefit of settlement, for social welfare depends on the well-being of individuals and, in particular, of litigants. The factor of trial risks thus creates both a private and a social reason for settlement.

Differences in Information and in Beliefs about Trial Outcomes. Now consider differences in information and in opinions about trial outcomes. These can generate impasses in private bargaining over settlement and produce trial. For example, suppose that a victim's losses are in fact high, \$10,000, but that the injurer cannot be convinced of this and thinks the losses are in the usual range of \$5,000. Consequently, the injurer might offer an amount much lower than the amount that would satisfy the victim, and there will be trial. Or suppose that a victim believes that his chances of prevailing are very good and his opinion about the law is different from the injurer's. Then again the injurer might not be willing to offer the victim enough to induce him to agree to settlement, so trial will result.

Although differences in information or in opinions may produce trial, this is not necessarily associated with a social reason for trial to occur. Trial may be largely the by-product of a disagreement about the outcome of the trial process and not produce better incentives. However, as will be noted later, it is possible that trial would produce better incentives, so an unambiguous conclusion about the relationship between differences in information and the social versus the private desirability of trial cannot be drawn.

Deterrence. The occurrence of trial may affect deterrence, but this is not a factor about which the parties themselves will ordinarily care, as I stressed in the analysis of the private incentive to bring suit. Socially, however, deterrence is of relevance. Hence, to the extent that trial may enhance

deterrence, there will be an insufficient private motive to go to trial and a socially excessive motive to settle.

The importance of the foregoing point, however, may be less than is sometimes suspected. Consider situations where injurers are able to obtain very favorable settlement agreements from victims (due, say, to victims' risk aversion) and hence where it appears that settlement results in systematic dilution in deterrence.⁴⁹ Here, it might be thought that trial would be socially beneficial because it would raise deterrence. But this is a mistaken view, given the possibility of imposing taxes on injurers (or of raising the measure of damages) to offset the dilution in deterrence due to settlement. It would be socially irrational for the state to countenance additional costly trials when the socially cheap alternative of taxes can be employed to raise deterrence.

A second reason to be skeptical about the social value of trials is that even if new information is developed at trial, this may not improve deterrence. For example, trial may result in a determination of negligence liability where the injurer is uncertain what the court will decide. But if the injurer is unclear what is considered negligence, how would his foreknowledge that the issue would be clarified through trial affect his incentives to take precautions? Such foreknowledge would not improve his incentives in an obvious way; in order that his behavior be appropriately channeled, he must know in advance what defines negligence, rather than know in advance that *later* the determination will be made.⁵⁰ Likewise, suppose that trial results in generation of more accurate information about the magnitude of the victim's losses than was known to the injurer during settlement negotiations. The injurer's foreknowledge that the harm he might do will be fairly precisely ascertained at trial can hardly improve his incentives to take appropriate precautions (high precautions if the harm would be great, low precautions if the harm would be small) if he does not know in advance what the quantum of harm will be.

These remarks are not, of course, meant to suggest that the information that is revealed at trial can never improve incentives.⁵¹ If an actor knows the definition of negligence and that negligent behavior will probably come to light at trial, then his foreknowledge that trial will occur will improve his motivation to act nonnegligently. Rather, the above remarks are intended to

⁴⁹ One of the main points of Polinsky and Rubinfeld, cited in note 48, is that settlement may dilute deterrence.

⁵⁰ If, however, the determination about negligence establishes a precedent, then it would have social value because it would beneficially affect behavior in the future.

⁵¹ On the value of accuracy in adjudication, see Louis Kaplow, *The Value of Accuracy in Adjudication: An Economic Analysis*, 23 J Legal Stud 307 (1994); Kaplow and Shavell, cited in note 7; and, closely related, Spier, cited in note 48.

explain why production of new information at trial does not necessarily imply that incentives are thereby improved.

Conclusion as to Private versus Social Incentives to Settle. Litigants have two basic motives to settle: to save themselves the costs of trial and to avoid risk. But their own cost savings are smaller than the social cost savings because they do not bear the public costs of trial, indicating a tendency toward too much trial. Litigants also may have reasons to go to trial—concerning differences in information or beliefs about trial outcomes—but these reasons to go to trial do not have a necessary connection to social deterrence-related gains from trial. This factor also, therefore, may lead to a tendency toward socially excessive trial. However, litigants do not usually take deterrence produced by trial into account, even though it is socially relevant. This factor works in the opposite direction and suggests that there might be too little trial. Yet the added deterrence likely to be produced by trial is often not apparent, so that a conjecture is that a problem of socially excessive trial is more likely than a problem of socially excessive settlement.⁵²

Policies to Promote Settlement or to Encourage Trial. The state can do various things to foster settlement if that is perceived to be the problem. It can impose penalties on parties for failing to settle, and it can attempt to increase the information parties have about each other to promote settlement. If, however, the state desires to encourage trial, it can do the opposite, impose penalties on settling parties and not seek to improve parties' information about each other.

B. Discussion

Let me now note a number of issues going beyond those just discussed and make several observations about settlement versus trial.

The Future Relationship between Opposing Parties. It is often said that a function of settlement is to prevent permanent rupture of the relationship that may exist between parties in a dispute, for settlement means that the parties themselves have made peace with each other, that they have not needed an outside authority to bring finality to their problem. This virtue of settlement one supposes can be of significance, especially in small communities or in other contexts where the litigants are likely to have repeated

⁵² That is, I think that analysis of various models will show this in some systematic sense. For example, in my paper (Shavell, cited in note 7), in which uncertainty concerns the amount of harm sustained by plaintiffs, the only possible problem is one of excessive trial. However, Kathryn E. Spier, *A Note on the Divergence between the Private and the Social Motive to Settle under a Negligence Rule*, in this issue, demonstrates that settlement may be socially excessive under the negligence rule.

contact with one another, because, if the parties are able to continue a workable relationship, they can cooperate in social or business affairs and increase their welfare. This consideration would not be important, though, where the litigants are strangers to one another (for instance, where two people who live miles apart are involved in an automobile accident). In any event, the value of settlement to the future relationship between litigants is something that, in the main, one would think the parties themselves would see as a benefit. Therefore, it does not seem to constitute a strong argument for the state to foster settlement.

Settlement as a Means of Securing Privacy and Maintaining Secrecy. Another factor not addressed above is that, when parties settle, various facts that would have emerged at trial do not come to the notice of the broader public. The privacy and secrecy that is achieved by settlement in these cases often constitutes an advantage to the settling parties and helps to explain why they settle. For example, a defendant firm whose product was defective may not want this information to come to public light and may be willing to pay an extra amount for that reason to achieve settlement (the victim may not much care whether the information is revealed). Or a victim might be embarrassed by the facts of a dispute (suppose the suit is for sexual harassment) or not want to acquire a reputation as a troublemaker, and so might wish to settle for that reason. In some cases, the parties' desire for privacy may be socially beneficial, but many times it seems that society would benefit from the information that would be revealed through trial. This would be the situation with regard to the firm that wants to keep its product defect secret: if the public learns about the defect, perhaps people can take precautions to reduce harm and, further, the firm will suffer adverse consequences, leading to improved deterrence. In circumstances like this, then, the private motive to settle may be excessive. (However, perhaps allowing settlement but mandating disclosure of certain facts about the dispute and the settlement terms would serve the social purpose.)

Unequal Wealth of Litigants. It is sometimes suggested that the unequal wealth of litigants creates an argument for encouragement of trial.⁵³ The argument begins with the observation that poorer litigants are systematically disadvantaged in settlement negotiations: they have a greater need for cash, may be more averse to the risk of trial, and will tend not to be as well represented by counsel as wealthier litigants. In the trial process, however, the aspiration is to treat individuals equally, regardless of their wealth. Hence, the argument goes, the less well-off would be better served by trial. But this conclusion is a non sequitur. If the trial process would result in better out-

⁵³ See Fiss, at 1076–78 (cited in note 48).

comes for poorer litigants than their settlements, one presumes that these litigants would not have agreed to their settlements in the first place; forcing poorer litigants to go to trial against their will must generally make them worse off.⁵⁴ Only an argument that poorer litigants (and their counsel) are systematically fooled into making disadvantageous settlements would justify a recommendation for trial that is based on the wealth status of litigants.

Development of Law. An additional issue is that valuable interpretations of the law may be made and new precedents established through trial. But, as I noted earlier, many trials do not make new law. Trials are often the result of disagreements about facts not about law. And where trial is the product of disagreement about law, published opinions that create new law are hardly automatic. Furthermore, a natural conjecture is that the number of cases the courts need to hear in order to modify the law is small in relation to the total volume of cases that do go to trial.⁵⁵ Therefore, it would seem that usually, though of course not always, the issue of the elaboration of the law will not be of real importance to the social versus the private incentive to go to trial.

Validation of Social Norms. A closely related issue to the development of law is that social norms may be validated through trial in the sense that the norms may be publicly called on to justify legal outcomes. Thus, the importance of the norms may be reinforced by the trial process. Because this factor is not one that individuals would tend to count as a private benefit of trial, it suggests a tendency toward insufficient trial. However, in assessing the relevance of the factor, one must consider not only the limited value of trials (as opposed to parents, schools, religious institutions) in teaching morality, but also that, as with the development of law, the number of trials necessary to meet this goal is probably not great.⁵⁶

The Paradoxical Economic Purpose of Trial. To this point, although a number of reasons why trial may be socially beneficial have been mentioned, the thrust of the analysis and discussion has been that the social

⁵⁴ This conclusion is perfectly consistent with the two premises of the argument under criticism. Notably, poor plaintiffs may well agree to settlements that substantially discount the expected court award they would receive, and court awards may well reflect a trial process that would treat poor plaintiffs on more or less the same plane as wealthy defendants—but this hardly alters the point that the poor plaintiffs would not agree to the settlement terms (however disadvantageous they might appear to a legal commentator) unless they preferred them to the gamble and expense of trial.

⁵⁵ Moreover, there are arguments for not viewing trial as necessary for the development of the law: courts could issue advisory opinions interpreting the law in the absence of particular disputes; in principle, there is no reason why courts require the occasion of a legal dispute to interpret the law.

⁵⁶ It would seem that the potential for teaching morality lessons from trials can be well achieved through the publicizing of a small and select number of trials.

value of trial is, on reflection, weaker than might otherwise appear. What then can be said to be the main social justification for the institution of trials? It seems that *the most important justification for society's having established the legal apparatus for the holding of trials is, paradoxically, not actually to have trials occur. Rather, it is to provide victims with the threat necessary to induce settlements.* As has been explained, settlement payments, together with appropriate taxes imposed on injurers (and, perhaps, ancillary policies like forcing disclosure of settlement terms), can achieve much desirable channeling of behavior, often making trial itself of marginal social value, of lesser value than its social cost.

Actual Legal Policy Bearing on Settlement versus Trial. The legal policies that we observe bearing on settlement versus trial predominantly foster the former. This is accomplished by allowing parties to engage in discovery, sometimes requiring them to participate in nonbinding arbitration prior to trial, holding settlement conferences, and so forth. These devices work principally by increasing the information that parties have about each other and thus tend to reduce the possibility of bargaining impasse. In addition, the law promotes settlement negotiations by prohibiting settlement offers from being introduced as evidence at trial and by the basic policy of making settlement agreements binding on the parties (otherwise the meaningfulness and value of settlements would be compromised). Further, courts often use their powers of suasion to encourage settlement.

The justification that one usually sees offered for the general promotion of settlement is that it clears dockets and saves public and private expense. This justification comports with economic analysis in the obvious sense that the parties do not consider the court's time and other public costs associated with trial as a saving from settlement. The policy of fostering settlement is also justified by what is not usually mentioned by commentators, that parties may want to go to trial for reasons that often are not socially relevant, notably, because of disagreement about the likely trial outcome. The possibility that trial ought to be held despite the parties' wishes to settle receives relatively little attention. One wonders, for example, about the wisdom of promoting settlement, let alone allowing it, in situations where deterrence is likely to be compromised by the fact that the identity and/or important aspects of the defendants' conduct do not become public knowledge.

The Prohibition against Private Settlement of Criminal Cases. Before continuing, let me note that our discussion of the private versus the social incentive to settle bears on an aspect of criminal law. Namely, not only is the initiation of criminal cases placed in the hands of public prosecutors, but so is, in the main, the decision to settle (agree on a plea) versus to proceed to trial; a victim of a crime does not have authority to settle with the criminal. This stands in great contrast, of course, to the law's encourage-

ment of settlement in civil cases. The legal policy of forbidding settlement between victims and injurers in the criminal context can be explained by a private-social divergence similar to that discussed earlier in relation to the incentive to initiate a case against a criminal. The victim's incentives in bargaining about the severity of a criminal sanction would often be very poorly aligned with society's. How much would a victim of a theft care about the length of the thief's jail sentence? The victim's interest is more likely to be centered on return of his goods and would not normally reflect society's interest in punishment to accomplish deterrence and incapacitation.⁵⁷ Additionally, as earlier noted, were victims allowed to make settlements, criminals would have an incentive to use threats and violence to induce victims to make bargains on generous terms.⁵⁸

V. LITIGATION EXPENDITURES

Last, let us briefly examine the level of expenditures on litigation. By such expenditures, I refer both to expenses incurred prior to settlement and to those incurred during the course of trial.⁵⁹

A. *Private versus Social Incentives to Spend on Litigation*

As a general matter, many of parties' litigation expenditures are fully offsetting in the sense that the two litigants would be in the same position had neither made their expenditures. A classic instance is when both parties devote effort to legal arguments of roughly equivalent weight but in support of opposite claims, or when both hire experts who produce equally convincing reports favoring opposite assertions. The reason that the two parties

⁵⁷ There are obvious qualifications to this statement. A victim might have an interest in incapacitation if the criminal would be likely to steal from him in the future. Also, a victim might be interested in revenge. In such cases, although the victim has a positive interest in seeing the criminal suffer a sanction, the sanction would generally not be equal to the socially appropriate sanction. It could easily be less (suppose the victim has some measure of sympathy for the victim, does not feel his act—car theft, say—deserves a long jail sentence) or more (suppose the victim is bent on punishment and wants a substantial imprisonment term imposed even though the theft was a relatively minor first-time offense).

⁵⁸ When, historically, private parties possessed the primary role in prosecution of criminal cases (see note 46 above), it is interesting that, although at first settlement (for a money payment in exchange for dropping prosecution) was allowed, the law then forbade it. Settlement of a criminal prosecution became a misdemeanor called "compounding felony" (yet the extent to which this prohibition was enforced has yet to be established). See Langbein, *Understanding the Short History of Plea Bargaining*, at 266 n8 (cited in note 46); Leon Radzinowicz, *2 A History of English Criminal Law and Its Administration from 1750*, at 313–15 (1957).

⁵⁹ I should note also that in this section I am discussing the level of litigation expenditures taking as given parties' decisions whether to sue and whether to settle or go to trial.

make offsetting expenditures may be that each fears the other will make his expenditure, so, to be safe, each does himself, or that each acts believing that he will gain an advantage over the other but turns out not to. In any case, these expenditures are socially wasteful because by assumption they do not alter trial outcomes.

Of course, many expenditures are not offsetting, or at least not fully offsetting; they do affect the position of one or the other of the parties. What can be said about these expenditures? In principle, a party's expenditure on an aspect of litigation could be either socially excessive or socially inadequate for the general reasons that I have already discussed in this article. Consider a victim's decision to hire an expert to produce a report in his behalf. The victim will not consider the cost that his expert's report imposes on the injurer: the injurer may need to hire an expert of his own to limit the effect of the victim's expert (but, our present assumption is, not fully offset it). Nor will the victim consider the cost due to the increased time the court will devote to listening to his expert and the injurer's expert should there be a trial. However, the victim will also fail to consider the effect of his expert's report on deterrence (suppose the report would unmistakably establish the injurer's liability). Victims' ignoring the costs they impose may lead them to spend socially excessively on experts, but their not considering the deterrence produced by their experts' reports could result in their not spending enough on experts.⁶⁰ Similar observations apply to vir-

⁶⁰ Explicit examples illustrating both possibilities can be constructed along the lines of those I discussed in Sec. II on the decision to sue (although the examples are, of necessity, more complicated). For instance, to illustrate that a victim might not hire an expert when it would be better that he do that, consider the following situation. Harm of \$1,000 will occur with certainty if a precaution is not taken, whereas harm will occur only with a probability of 40 percent if a precaution that costs \$400 is taken. Further, the cost of suit for a victim is \$100, he will prevail at trial with a likelihood of 60 percent if he does not hire an expert before trial to prepare a report to help establish causation, whereas he will prevail with a probability of 90 percent if he does engage the expert, at a cost of \$350. Assume, too, for simplicity that the parties will settle and the injurer will pay the victim his expected judgment because they agree about the likelihood of the trial outcome and would face equal additional trial costs.

In this situation, the victim will decide not to hire an expert: without the expert, his expected gain from trial—and thus his settlement—will be 60 percent \times \$1,000 = \$600, and with the expert his settlement will be 90 percent \times \$1,000 = \$900, so the value of the expert to him will be only \$300, less than the cost of the expert. (Note, though, that the victim will spend \$100 to sue since he will obtain \$600 in settlement from bringing suit.) Because victims do not hire experts and obtain only \$600 in settlements, the cost to injurers of causing harm will be \$600. In consequence, injurers will not take the precaution: if they take the precaution, they will reduce their expected liability from \$600 to 40 percent \times \$600 = \$240, saving themselves \$360, but the precaution costs \$400. Therefore, social costs will be the certain harm plus the costs of suit, namely, \$1,000 + \$100 = \$1,100.

Social costs would be lower, however, were victims to hire experts, for that would lead injurers to take the precaution. In particular, injurers would then settle for \$900, so that the precaution would then reduce their expected liability from \$900 to 40 percent \times \$900 =

tually any type of legal expenditure,⁶¹ and whether they were made by victims or by injurers. Thus, although nonoffsetting expenditures do by definition help the parties who make them, we cannot say in theory whether or not they are socially excessive or inadequate; which is so depends on the particulars.

It is worth mentioning, however, that there is a broad class of facts concerning the magnitude of harm that parties may spend to demonstrate but where this will be socially wasteful because it will not improve deterrence. Specifically, this is the class of facts about the magnitude of harm that the injurer did not know when he made his decision that resulted in harm.⁶² Consider, for example, the precise extent of harm the victim suffered in an automobile accident. Much may be spent establishing the magnitude of harm by presenting evidence on medical bills, the time lost from work, and the victim's wage. But these expenditures will not improve the incentives of drivers to prevent accidents if they do not know the magnitude of harm that will result in the event that they cause an accident, that is, if all that they know is the probability distribution of possible harms (depending on who is in the car they strike, the nature of the impact, and so forth). If this is the case, injurers' incentives to prevent accidents will be essentially identical where there is no real inquiry into the scale of harm, and damages are simply set equal to the average harm that victims sustain in that kind of automobile accident.⁶³

In sum, litigation expenditures are subject to the same general divergence

\$360, or by \$540, making it worth their while to spend \$400 on the precaution. In this case, social costs would be the cost of the precaution plus expected harm plus litigation costs (including the cost of the expert), or $\$400 + 40 \text{ percent} \times (\$100 + \$350 + \$1,000) = \$1,080$, which is lower than \$1,100.

⁶¹ For example, as emphasized in Bruce Hay, *Civil Discovery: Its Effects and Optimal Scope*, 23 J Legal Stud 481 (1994), discovery expenditures may be socially excessive or socially inadequate because of the difference between private incentives to spend on discovery and its social value in improving deterrence.

⁶² The point now being discussed amplifies that made above in Sec. IVA, where I said that information about damages developed at trial might not affect deterrence.

⁶³ For example, suppose that all that injurers know when deciding on a precaution is that, if an accident occurs, it will result in harm of \$1,000, \$2,000, or \$3,000, each with equal probability—injurers thus do not know the particular magnitude of the harm that would occur. If damages are set equal to actual harm (after a perhaps costly legal determination), injurers obviously will not know ex ante what their liability will be if an accident occurs; they will know only that their expected liability given an accident will be \$2,000 (that is, $1/3 \times \$1,000 + 1/3 \times \$2,000 + 1/3 \times \$3,000$). If, instead, damages are simply set equal to the average harm of \$2,000, injurers' liability given an accident will be the same. Therefore, injurers will behave identically. (In this example, I assumed that injurers are risk neutral. If they are risk averse, the point also holds; in that case, the fixed amount of damages is set appropriately below the average harm, and risk-averse injurers will act the same way as they would if damages were set equal to actual harm.) See, more generally, Kaplow and Shavell, cited at note 7.

between private and social incentives that was discussed previously, so they may be in principle either socially excessive or socially inadequate. Further, certain categories of litigation expenditures are identifiable as socially excessive: offsetting expenditures and those determining damages more accurately than was known when precautions were taken.

B. Corrective Policy

The means of controlling litigation expenditures are straightforward. On the one hand, expenditures can be discouraged through monetary disincentives like fees or taxes, and they can also be regulated through constraints on the time parties are given to prepare for trial, restrictions on discovery, limits on the length of permitted submissions and the number of testifying experts, and so forth. In fact, controls on expenditures seem to be made largely through such forms of regulation of the pretrial and trial process rather than through financial inducement. On the other hand, expenditures can be encouraged through subsidy and other monetary rewards; regulation by itself cannot lower litigation costs for a party.

Although the means of controlling litigation expenditures are evident, the circumstances in which they ought to be used seem difficult to determine because of courts' lack of knowledge about, especially, the deterrent effects of establishing liability or more accurate damages. Sometimes, it may be possible for the state to make judgments about these matters. I suggested, for example, that in automobile accidents, the value of precise estimates of harm is dubious; if so, a policy that severely limited expenditures on proof of damages for such accidents would be valuable. But the ability of courts to make particularized determinations about the social appropriateness of legal expenditures is circumscribed.

VI. CONCLUSION

I have argued in this article that there is a basic cleavage between private and social incentives to use the legal system that pervades decisions about it, from the important choice at the outset to press a claim, to the decision about settlement versus trial, to the decisions made by parties about their levels of legal expenditure. The reasons for the private-social divergence were twofold. First, parties do not take into full account the costs that their legal actions generate for each other and for the court. And second, their personal benefit from settlement or trial is not closely related to the resulting social benefits. These social benefits I assumed to reside primarily in the channeling of behavior, but however broadly one conceives them to be, they are quite distinct from the private benefits. The difference between the private and social benefits can work either to reinforce the cost-related

tendency toward excessive litigation or to counter it and could result in socially inadequate use of the legal system.

The most broad implication of the private-social differences in incentives to use the legal system is that we cannot accept the general amount of legal activity—either the volume of suit, the number of settlements, or the level of litigation expenditures—as approximately correct. We are dealing with a market that is imbued with externalities and in which outcomes cannot be taken as tending toward the socially desirable. In consequence, social intervention to influence use of the system may be warranted. Some such interventions do not require much information on the part of the state—this is true of the recommendation that taxes be imposed on defendants, so that their level of care will reflect not only the harm they cause but also the litigation costs they generate. But much intervention, and notably that needed to control the volume of suit, requires that the state estimate the deterrent effect of litigation, something that is not easy to do, especially for narrow categories of case. For that reason, I suggested that policy to reduce or encourage litigation is probably best conceived to be applicable only to fairly broad categories of case. But there is, I argued, real scope for social advantages to be secured through such policy because of the large volume of litigation in certain areas, like that of automobile accidents. I also indicated that, with respect to settlement, there is some justification for what we observe, that the courts generally seek to foster settlement over trial, yet there is hardly an unambiguous rationale for this policy. Additionally, I discussed the level of litigation expenditures, noting some particular features of it that suggest parties have a propensity to spend excessively and calling for their curtailment, but at the same time I observed that such expenditures could be inadequate.

I have also noted why certain policies and principles that meet with popular approval do not have a sound basis. I explained that fee-shifting or making plaintiffs pay for the public costs of trial will not result in any automatic way in the right amount of suit, and indeed that there does not exist any simple policy that will generally do so. I emphasized as well that the view that a decision regarding use of the legal system is socially justified if and only if its costs are less than the expected monetary return is erroneous. Additionally, I observed that the economic interest of the bar will cause it to resist social policies to reduce legal activity and that academics are likely to disfavor these policies as well in the name of fairness or compensation for victims, even though the justification for such opposition may be problematic.