

ISSN 1045-6333

HARVARD

JOHN M. OLIN CENTER FOR LAW, ECONOMICS, AND BUSINESS

ECONOMIC ANALYSIS OF LITIGATION
AND THE LEGAL PROCESS

Steven Shavell

Discussion Paper No. 404

02/2003

Harvard Law School
Cambridge, MA 02138

This paper can be downloaded without charge from:

The Harvard John M. Olin Discussion Paper Series:
http://www.law.harvard.edu/programs/olin_center/

**ECONOMIC ANALYSIS OF LITIGATION
AND THE LEGAL PROCESS**

Steven Shavell*

ABSTRACT

This paper contains the chapters on litigation and the legal process from a general, forthcoming book, *Foundations of Economic Analysis of Law* (Harvard University Press, 2003). In chapter 17, I consider the basic theory of litigation. Here I describe the three phases of litigation: its initiation through suit, the determination of whether the parties will settle their case or proceed to trial, and, if trial results, the trial expenditures. I also analyze the social desirability of their decisions, a major theme being that the private incentives to litigate may diverge from what is socially desirable. In chapter 18, I extend the basic theory of litigation, examining among other issues the bringing of negative value suits, shifting of legal fees to losers at trial, lawyer-client fee arrangements, and the influence of insurers on litigation. Then, in chapter 19, I discuss several general aspects of the legal process not considered in the basic theory and its extensions, including private systems of adjudication, the value of accuracy in adjudication, the appeals process, and the function of legal advice.

*Samuel R. Rosenthal Professor of Law and Economics. Research support from the John M. Olin Center for Law, Economics, and Business is greatly appreciated.

Table of Contents

Economic Analysis of Litigation and the Legal Process

Chapter 17. Basic Theory of Litigation

1. Bringing of Suit
2. Fundamental Divergence Between the Private and the Socially Desirable Level of Suit
3. Settlement versus Trial
4. Divergence between the Private and the Socially Desirable Level of Settlement
5. Trial and Litigation Expenditure

Chapter 18. Extensions of the Basic Theory of Litigation

1. Negative Value Suits
2. Sharing of Information Prior to Trial
3. Forced Disclosure of Information Prior to Trial: Discovery
4. Shifting of Legal Fees to the Loser at Trial
5. Difficulty of Statistical Inference from Trial Outcomes
6. Elements of Trial Outcomes Apart from the Judgment
7. Role of Lawyers in Litigation
8. Role of Insurers in Litigation

Chapter 19. General Topics on the Legal Process

1. Public versus Private Legal Systems
2. Accuracy of the Legal Process
3. Appeals and the Legal System
4. Legal Advice and the Role of Lawyers in the Legal System

Summary Table of Contents of

Foundations of Economic Analysis of Law

(forthcoming 2003, Harvard University Press)

Chapter 1. Introduction

Part One. Accident Law

Chapter 2. Liability and Deterrence: Basic Theory

Chapter 3. Liability of Firms

Chapter 4. Extensions of the Basic Theory

Chapter 5. Liability, Risk-bearing, and Insurance

Chapter 6. Liability and Administrative Costs

Part Two. Property Law

Chapter 7. Definition, Justification, and Emergence of Property Rights

Chapter 8. Division of Property Rights

Chapter 9. Acquisition and Transfer of Property

Chapter 10. Conflict and Cooperation in the Use of Property: The Problem of Externalities

Chapter 11. Public Property

Chapter 12. Property Rights in Information

Part Three. Contract Law

Chapter 13. Overview of Contracts

Chapter 14. Contract Formation

Chapter 15. Production Contracts

Chapter 16. Other Types of Contract

Part Four. Litigation and the Legal Process

Chapter 17. Basic Theory of Litigation

Chapter 18. Extensions of the Basic Theory of Litigation

Chapter 19. General Topics on the Legal Process

Part Five. Public Law Enforcement and Criminal Law

Chapter 20. Deterrence with Monetary Sanctions

Chapter 21. Deterrence with Nonmonetary Sanctions

Chapter 22. Extensions of the Theory of Deterrence

Chapter 23. Other Functions of Sanctions: Incapacitation, Rehabilitation, Retribution

Chapter 24. Criminal Law

Part Six. General Structure of the Law

Chapter 25. The General Structure of the Law and Its Optimality

Part Seven. Welfare Economics, Morality, and the Law

Chapter 26. Welfare Economics and Morality

Chapter 27. Implications for the Analysis of Law

Chapter 28. Income Distributional Equity and the Law

Chapter 29. Commonly-Raised Questions about Economic Analysis of Law

Economic Analysis of Litigation and the Legal Process

(part of Foundations of Economic Analysis of Law)

Steven Shavell

©2003. Steven Shavell. All Rights Reserved.

Chapter 17

BASIC THEORY OF LITIGATION

In this chapter, I take up the basic theory of litigation. Three stages will be considered. In the first, the party who has suffered a loss decides whether to bring a suit. A party who brings a suit is, of course, the *plaintiff*, and the party who is sued is the *defendant*. In the second stage, the plaintiff and the defendant decide whether or not they will settle the case -- in the event of a settlement, the plaintiff agrees to drop the suit, usually in exchange for a payment from the defendant. If a settlement is not reached, the third stage, that of trial, occurs.

In each stage, I discuss how the parties behave and then how this compares to what is socially desirable. A theme that I emphasize is that the private incentives of parties in litigation may diverge significantly from socially appropriate incentives, given that litigation is a costly activity.

1. Bringing of Suit

1.1 Definition of suit. By a suit, I will mean the taking of a costly initial step that is a prerequisite to further legal proceedings and trial. One may interpret suit as a formal legal action, such as the filing of a complaint, or as an action short of that, notably, the hiring of a lawyer. Bringing a suit involves costs; the plaintiff will expend time and energy and possibly incur a bill for legal services and filing fees.

1.2 Private incentive to sue. The plaintiff will sue when his cost of suit is less than his expected benefits from suit.¹ The plaintiff's expected benefits from suit involve possible settlement payments or gains from trial. However, for the purposes of this and the next section on suit, it will be sufficient to assume simply that there is an expected benefit from suit, without inquiring into its determination.

From the simple description of suit, note that suit is more likely the lower the cost of suit, the greater the likelihood of winning at trial, and the greater the plaintiff's award conditional on winning. Suit is also more likely the less averse to risk the plaintiff is and the more averse the defendant is.

1.3 Volume of suit and its costs. About 2.3 million non-criminal cases were filed in federal courts in 2000, and approximately 20.1 million non-criminal cases were filed in state

¹In this chapter it will be assumed that parties are risk-neutral unless otherwise noted.

courts.² These figures are probably underestimates of the truth, because many individuals who threaten suit settle without ever having filed suit in a formal sense and many disputes are arbitrated.

The total costs of litigation, including those of suit, are great, as is indicated by the facts that legal services absorb approximately 1 percent of the labor force and 1.3 percent of gross domestic production.³ The costs of suit, that is, of initiating legal action, are difficult to estimate, but a simple calculation is suggestive. If the hourly rate of a lawyer is, say, \$250 and suit would require only 20 hours of the lawyer's time, the cost would be \$5,000, excluding the consideration of the plaintiff's time, which could well be significant.⁴ Thus, even if individuals are certain to prevail in trials, they will not bring suit unless their losses surpass a fairly significant threshold.⁵

2. Fundamental Divergence between the Private and the Socially Desirable Level of Suit

2.1 In general. I now take up the question of how the number of suits that parties are motivated to bring compares to the socially optimal level of suit. The main point that will be made is that the private incentive to bring suit is fundamentally misaligned with the socially optimal incentive to do so, and the deviation between them could be in either direction. The reasons for this conclusion may be understood as follows.

On one hand, there is a divergence between the social and the private costs of suit that can lead to a socially excessive level suit. Specifically, when a plaintiff contemplates bringing suit, he bears only his own costs; he does not take into account the defendant's costs or the state's costs that his suit will engender. Hence, the plaintiff might be led to bring suit when the total costs associated with suit would make that undesirable.

On the other hand, there is a difference between the social and private benefits of suit that may lead to a socially inadequate level of suit or may reinforce the cost-related tendency toward excessive suit. In particular, the plaintiff would not usually be expected to treat as a benefit to himself the social benefits flowing from suit, notably, its deterrent effect on the behavior of injurers (and more generally, other effects as well).⁶ What the plaintiff does consider as the benefit from suit is the gain he personally would obtain from prevailing. This private gain is not the same as the social benefit from suit -- the private gain is a transfer from the defendant and, as will be seen, could be either larger or smaller than the social benefit from suit.

2.2 The divergence in a simple model. To clarify these points, let us consider for concreteness the model studied in Part 2 of the book, on accidents, in which injurers can exercise

²See Ostrom, Kauder, and LaFountain 2001a, 13. The number of cases stated in text for state courts excludes not only criminal cases, but also juvenile cases and traffic cases.

³See Statistical Abstract of the United States 2001, Tables 596, 641, pp. 384, 418.

⁴The average hourly rate of a partner in a law firm as of January 1, 2001, was \$246; see The 2001 Survey of Law Firm Economics, p. 11-39.

⁵This statement applies as well if the lawyer is bearing his own costs, under a contingency fee (see section 7 of chapter 18).

⁶One way of expressing this point about deterrence is to observe that by bringing suit, plaintiffs contribute to the general impression that potential injurers have that they will be sued if they cause harm. Were the law only on the books, but never to result actually in suit, potential injurers would have nothing to fear.

care to lower the risk of accidents, and the social welfare goal is minimization of total social costs, comprised of the costs of precautions, those of accidents that occur, and also the costs of litigation. Thus, the socially optimal amount of suit is that which minimizes total social costs. We want to show that the amount of suit that private parties find it in their interest to bring could be above or below the socially optimal amount. Let us assume that liability is strict (I will comment on the negligence rule below), and consider two examples.

Example 1. To illustrate the possibility of socially excessive suit, suppose that the losses a victim would suffer in an accident are \$10,000, that a victim's cost of bringing suit is \$3,000 and an injurer's cost of defending is \$2,000, that the probability of accidents is 10 percent, and that there is no precaution that injurers can take to lower the accident risk.

Victims will then bring suits whenever accidents occur, for suing will cost a victim only \$3,000 and yield him \$10,000. From the social perspective, this outcome is undesirable. Suit creates no beneficial deterrent, as injurers by assumption cannot do anything to lower risk. Yet suit does generate legal costs: Expected legal costs are $10\% H (\$3,000 + \$2,000) = \$500$. The bringing of suits is not socially desirable in this example because there are no incentives toward safety created by the suits. This fact is of no moment to victims, nor are other parties' litigation costs. Victims bring suits for their private gain of \$10,000.//

Although in this example there was no deterrent benefit whatever from the bringing of suits, it should be obvious that the point of the example would hold if the deterrent benefit were positive and not too large.

Let us now illustrate the opposite possibility, that suits might not be brought even though it would be best that they are brought.

Example 2. Suppose here that the losses victims suffer in accidents are \$1,000 and that an expenditure of \$10 by injurers will reduce the probability of accidents from 10 percent to 1 percent. The costs of suit and of defending against suits are as in the previous example.

In this case victims will not bring suits, as doing so would cost a victim \$3,000 but yield him only \$1,000. Because injurers will not be sued, they will have no reason to take care to reduce risk, and total social costs will therefore be $10\% H \$1,000 = \100 .

It would be socially desirable for victims to bring suit, however. If they were to do so, injurers would be led to spend \$10 to lower risk from 10 percent to 1 percent, and total social costs would thus be $\$10 + 1\% H (\$1,000 + \$3,000 + \$2,000) = \$70$. The bringing of suits is socially worthwhile here (\$70 being less than \$100), because of the significant reduction in accident losses that would result. (And observe that this is true even though the total legal costs of \$5,000 exceed the victim's losses of \$1,000.) But victims do not take the deterrence-related benefits of suit into account. Each victim looks only to his own gain from suit, which is negative.//

As emphasized, a victim does not bring suit in this example because his private gain, the harm he has sustained, is not sufficient to outweigh his legal costs, even though the general deterrent that would be engendered by the bringing of suits would so reduce accident losses that the bringing of suits would be socially worthwhile.⁷

⁷A formal explanation of the points made in this section is as follows. Suppose that liability is strict. Victims will sue if and only if $c_p < h$, where c_p is the victim's (plaintiff's) cost of suit and h is the harm suffered and what he will collect if he sues. Let q be the probability of harm h if suit is not brought, let q' be the probability of harm if suit is brought, and let x be the precaution expenditures that injurers will be induced to make if there is suit. (Thus, q' will be less than q if x is positive.) If suit is not brought, total social costs will simply be qh . If suit is

2.3 The divergence continued; its importance and interpretation. I now mention a number of issues that bear on the foregoing and its interpretation.

(a) *Negligence rule and the divergence between the private and the socially optimal level of suit.* Under the negligence rule, there is some reason to believe that the possible problem of excessive suit is less serious than under strict liability. The main reason is that, under the negligence rule, a harmful outcome is less likely to produce a suit than under strict liability because negligence must be shown in order to prevail. Indeed, in a hypothetical, perfectly working negligence system, a harmful outcome that is not the result of negligence would never result in suit, and this turns out to imply that there will never be a problem of excessive suit, and indeed that it would be desirable for the state to subsidize suits.⁸ Of course, the negligence system does not operate perfectly in practice; victims sometimes bring suit against non-negligent injurers and injurers sometimes act negligently. Therefore, problems of socially excessive suit may well exist under the negligence rule, but it is plausible that they would be worse under strict liability.

(b) *Generality of divergence between socially optimal and privately induced amount of suit.* It should be clear from the discussion in section 2.1 that the point that the private and the social incentives to bring suit may diverge is robust. On one hand, it will always be the case that the private cost of use of the system will be less than the social cost. And, on the other hand, the private benefits from suit will be what the plaintiff will win from suit, usually money, whereas the social benefits from suit will ordinarily be different: They will always include deterrence benefits and may also include compensation of victims (if insurance is unavailable) and the setting of precedent. Potential plaintiffs will tend not to take all of these social benefits into account.

However, the private and social incentives to bring suit may roughly coincide in contract disputes. In such cases, the interested parties (with the exception of the state) are present at the bargaining table and actively design their contracts before disputes arise. Hence, we would expect contracting parties to set the terms of their agreement to minimize their expected losses *plus* their joint litigation costs. For example, to avoid litigation costs, contracting parties might stipulate that certain problems will not be deemed breaches of contract, or they might require that disputes will be inexpensively arbitrated. Nevertheless, many contracts will not be sufficiently detailed to include such remedies for problems of inappropriate suit, so that an

brought, social costs will be $q/h + x + q(c_p + c_D + c_S)$, where c_D is the injurer's (defendant's) litigation cost and c_S is the state's cost. Hence, the bringing of suit will lower total costs, be socially worthwhile, if and only if $q(c_p + c_D + c_S) < (q/h - x)$. In other words, suit is socially worthwhile if the expected litigation costs are less than the net deterrence benefits of suit. It is clear that this social condition for when suits should be brought, and the private condition for when they will be brought, namely, $c_p < h$ are very different. Whether victims will sue does not depend on the costs c_D and c_S . Moreover, the private benefit of suit is h , the amount of harm (conditional on harm occurring), because this is what the victim will receive as a damages award; in contrast, the social benefit depends on the harm weighted by the reduction in the accident probability, q/h net of the cost of precautions x . Therefore it is evident (as the examples illustrate) that victims might sue when suit is not socially optimal, and that victims might not sue even when suit would be socially optimal.

⁸If the state's policy is to subsidize suit such that victims would always be willing to bring suit for negligently caused harm, injurers would be led by the threat of suit to take due care. Hence, no suits would actually result, just because injurers would take due care. Thus, no litigation costs would in fact be incurred by society.

attenuated private-social divergence to sue may still exist.

(c) *Practical importance of the divergence.* The difference in the private and the social costs of suit is often large, at least in percentage terms. As emphasized, the private cost divergence is that victims do not take into account injurers' and the state's litigation costs. Thus, it is not unreasonable to expect that victims may fail to take into account around half of total litigation costs.⁹

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 automobile accident disputes comprise at least half of all tort litigation.¹⁰ 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 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.¹¹

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 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, as with automobile accidents, surprisingly sparse, it is not inconsistent with the hypothesis of low deterrence.¹² The litigation costs of product liability, however, are high. Therefore, it again appears conceivable that the private incentive to

⁹For example, in 2000, victims' fraction of total litigation expenditures by victims and defendants in tort cases was approximately 52 percent; see Tillinghast-Towers Perrin 2002. Also, in 1986, victims' fraction of total litigation expenses in automobile litigation was about 54 percent, and their fraction of total litigation expenses in non-automobile tort litigation was about 42 percent; see Hensler et al. 1987.

¹⁰See Kakalik and Pace 1986, x, and Ostrom, Kauder, and LaFountain 2001b, 28.

¹¹Deweese, Duff, and Trebilcock 1996, 15-26, and Sloan 1998 survey empirical studies of the effects of liability on automobile accidents; as they discuss, the evidence is mixed and in some respects contradictory.

¹²For example, in one of the few empirical studies on product liability and deterrence, Priest 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, Deweese, Duff, and Trebilcock 1996, 205, 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.

sue has resulted in an excessive volume of suit.¹³

The opposite possibility, that the volume of suit is socially inadequate, also seems to be of practical significance. Recall Example 2 in which an individual's losses were relatively low, 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 real relevance. One can readily imagine situations in which firms know that the harms that they cause will not be great enough to be worthwhile for a typical victim to pursue, even though the incidence of the harms could be decreased substantially by modest expenditures. (Consider low-level pollution damage, such as more frequent peeling of paint in a neighborhood near a factory, that 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.

(d) *Cost of suit leads to inadequacy of precautions.* An issue that we have not adequately discussed concerns the effect of the costs of suit on the level of precautions that injurers are led to take. As a general matter, the costliness of suit means that injurers' incentives to take precautions will be too low, for two reasons. First, and obviously, injurers might not be sued due to litigation costs faced by victims, and thus injurers' expected liability will tend to be too low; this is especially likely if the level of harm they tend to cause is lower than victims' costs of suit. Second, even if injurers expect to be sued when they cause harm for which they should be liable, their level of precautions will still be too low: The damages that injurers have to pay equal the direct harm they cause for their victims, but the full social costs include also the litigation costs associated with suit -- the full costs that society incurs when harm leads to suit is not only the direct harm but also the resources absorbed in the litigation process. Thus, for injurers' incentives to be correct, injurers should bear, in addition to the direct harm caused to victims, the litigation costs borne by victims and by the state (injurers bear their own litigation costs already). If, for example, the harm is \$10,000 and litigation costs of the plaintiff are \$3,000 and those of the state \$1,000, the injurer should pay \$14,000, not \$10,000.¹⁴ This point is significant because the litigation costs of plaintiffs are substantial; to ignore these costs is to omit perhaps a third of the damages that are needed to provide injurers with proper incentives to reduce harm.¹⁵

¹³Note 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.

¹⁴It should be observed that the conclusion of this paragraph applies whether or not the volume of suit is 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, so that the injurer should still pay for the defendant's and the state's litigation costs.

¹⁵As previously mentioned with respect to tort litigation (section 1.1 of chapter 12), total litigation costs appear to be about equal to amounts received by plaintiffs. Also, litigation costs may be borne roughly equally by plaintiffs and defendants (note 9). Hence, for every dollar that defendants pay to plaintiffs, plaintiffs may spend 50 cents in litigation costs, and defendants need to pay this (and the state's costs) as well in order for their incentives to

2.4 Corrective policy. 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. Conversely, if there is an inadequate level of suit, 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. 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 suit. This means that the state needs to deduce the cost of precautions and their effectiveness.

It should be noted that, by contrast, for the state to ameliorate the problem of inadequate precautions due to the cost of suit (as just discussed in section 2.3(d)), a corrective policy that will be helpful is easily identified and should not be difficult to implement. Namely, when suit is brought, liable defendants should have to pay more than the harm done; they should also have to pay victims' litigation costs and the states' costs, for only then will they be bearing the full social costs associated with harm.¹⁶

2.5 Comments on the foregoing. Let me now make a number of remarks about the divergence between the actual and the socially desirable amount of suit and about corrective policy.

(a) *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.¹⁷ 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. This is desirable because we assume the farmer's benefit to be coincident with the social benefit, so when he decides that his benefit exceeds the veterinarian's cost, that is true for society as well. But when a person brings suit, his benefit is, as stressed, generally different from the social benefit. Hence, it is not clear that making him pay the full social costs of suit will lead to a better decision from society's perspective. (More to the point, perhaps, we know that the policy of making plaintiffs pay for the state's costs cannot be desirable in all cases, because we know that it might discourage suit when suit would be socially beneficial.)

The second policy that many find appealing is 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 just

take precautions to be appropriate.

¹⁶If it is desirable to increase victims' incentives to sue, then the augmented damages could be paid to them; if it is best to discourage suit, certainly it would not be good to allow victims to collect more, and the extra payments of defendants should thus be made to the state.

¹⁷For discussion of this argument, see, for example, Lee 1985 and Posner 1985b, 136.

or fair that someone who has been shown wrong in court should pay the other side's legal expenses.¹⁸ The question for us to consider, however, is whether loser-pays fee-shifting will have a beneficial effect on the volume of suit.¹⁹ In the examples and the analysis that I presented above, fee-shifting would simply serve to encourage litigation because victims -- whom we have assumed to have good cases -- would always be able to lay off their litigation costs on injurers. Thus, fee-shifting would tend to worsen any problem of excessive suit. Moreover, in situations 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 in which it was desirable for the state to subsidize suit, fee-shifting would equally induce suit; and in contexts in which 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 of 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.

(b) Practical use of corrective policy. Policies to correct the volume of suit in broad categories of case may be desirable and feasible. If, for example, serious study of automobile accidents and/or product accidents suggests that suit is excessive (see section 2.3(c)), one could imagine imposition of fees on plaintiffs for bringing suit, or even a ban on suit, for these accidents. Similarly, if it were concluded that we ought to encourage suit for certain types of low-magnitude harms (see again section 2.3(c)), one could 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 legislation²⁰ and automobile no-fault statutes have removed important categories of accident from the domain of tort law,²¹ New Zealand has barred suit for all cases of personal injury,²² and certain jurisdictions have shifted legal fees to losing plaintiffs to discourage litigation. On the other hand, legal aid programs have been employed to subsidize litigation, 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.²³

¹⁸For discussion of this fairness rationale, see, for example, Rowe 1982, 653-57.

¹⁹The proper consideration of notions of fairness from the perspective of welfare economics is discussed in chapters 26 and 27 below; the issue of relevance here, however, is the effect of fee-shifting on social welfare through its influence on litigation costs and the volume of suit.

²⁰See, for example, Dewees, Duff, and Trebilcock 1996, 387-94.

²¹See, for example, Joost 1992 and Carroll et al. 1991.

²²See, for example, Palmer 1994.

²³This is hardly to say that these policies have been adopted for the reasons that I have discussed. For example, it seems that worker's compensation legislation was enacted primarily because workers were not able to

Whether 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 deterrent 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 do so, 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 cases.

Note on the literature. The subject of the private versus the social incentive to bring suit was first developed in my article Shavell (1982b).²⁴

3. Settlement versus Trial

Assuming that suit has been brought, I now take up the question of whether parties will reach a settlement or instead go to trial. A settlement is a legally enforceable agreement, usually involving a payment from the defendant to the plaintiff, in which the plaintiff agrees not to pursue his claim further. If the parties do not reach a settlement, we will assume that they go to trial, that is, that the court will determine the outcome of their case.²⁵ I discuss here two different models describing whether settlement occurs and then consider the socially optimal versus the private decision of whether to settle.

3.1 Simple model. Let us suppose for simplicity that the plaintiff and the defendant each has somehow formed beliefs -- which may differ -- about the trial outcome. Then we can discuss settlement possibilities in terms of two quantities. Consider first the minimum amount that the plaintiff would accept in settlement. Assuming that the plaintiff is risk-neutral, this minimum acceptable amount equals his expected gain from trial less the cost of going to trial. For instance, if the plaintiff believes he will prevail with probability 70 percent, would obtain \$100,000 upon prevailing, and the trial would involve expenses to him of \$20,000, the minimum amount he would accept in settlement is $70\% \times \$100,000 - \$20,000 = \$70,000 - \$20,000 = \$50,000$; if he were offered anything less than this amount, he would be better off going to trial. The other quantity is the maximum amount that the defendant would be willing to pay in settlement; this is his expected loss from trial plus his expense of going to trial. If the defendant believes the odds of the plaintiff's winning are, say, only 50 percent, and the defendant's trial costs would be \$25,000, then he would pay at most $50\% \times \$100,000 + \$25,000 = \$50,000 + \$25,000 = \$75,000$ in settlement.

It is evident that *if the plaintiff's minimum acceptable amount is less than the defendant's maximum acceptable amount, a mutually beneficial settlement is possible* -- a settlement equal to any amount in between these two figures would be preferable to a trial for each party. Thus, if the plaintiff's minimum is \$50,000 and defendant's maximum is \$75,000, any amount in

obtain compensation under the tort system, not because the litigation costs of the tort system were felt to outweigh its deterrence benefits.

²⁴See also extensions and further developments in Kaplow 1986b, Menell 1983, Rose-Ackerman and Geistfeld 1987, and Shavell 1997, 1999; and see Ordover 1978 for related analysis.

²⁵The underlying assumption that there is either a settlement or a trial is to a degree an abstraction, because there can be both: A trial can proceed and before it ends, the parties can settle. Taking this possibility into account would not alter the qualitative nature of the conclusions to be reached, because it would remain true that settlement saves litigation costs and occurs before the full judicial conclusion is reached.

between, such as \$60,000, would be preferred by each to going to trial. But if the plaintiff's minimum exceeds the most that the defendant will pay, settlement cannot occur. If, for instance, the defendant thought the plaintiff's chances of winning were only 20 percent, the defendant's maximum amount would be $\$20,000 + \$25,000 = \$45,000$, so the most he would pay is less than the minimum \$50,000 that the plaintiff would be willing to accept, and settlement could not occur.

Can more be said about when a mutually beneficial settlement will exist? That is, under what conditions will the plaintiff's minimum acceptable demand be less than the defendant's maximum acceptable payment? It is clear that if the plaintiff and the defendant have the same beliefs about the trial outcome, then there should always exist mutually beneficial settlements, because they can each escape trial costs by settling. Suppose that they both believe that \$50,000 is the expected judgment the defendant will have to pay if there is a trial. Then the trial costs that the plaintiff would bear would lead to his willingness to accept a lower figure than \$50,000; if his trial costs would be \$10,000, he would accept \$40,000, and so forth. Conversely, any trial costs that the defendant would have to bear would increase above \$50,000 what he would be willing to pay; if his trial costs would be \$10,000, he would be willing to pay \$60,000. Thus, the settlement range would be from \$40,000 to \$60,000. For settlement possibilities to be eliminated, the plaintiff's minimum amount must rise far enough from \$40,000 and/or the defendant's fall far enough from \$60,000 such that the plaintiff's minimum turns out to exceed the defendant's maximum. That can occur only if they have different beliefs about the trial outcome. This line of thought suggests the following conclusion: A mutually beneficial settlement amount exists as long as the plaintiff's and defendant's estimates of the expected judgment do not diverge too much. Indeed, it can be shown that *a mutually beneficial settlement exists as long as the plaintiff's estimate of the expected judgment does not exceed the defendant's estimate by more than the sum of their costs of trial.*²⁶ Let us illustrate.

Example 3. Consider the situation that we have mentioned above in which the plaintiff's expected gain from suit is $70\% \text{ H } \$100,000 = \$70,000$, his costs of trial are \$20,000, the defendant's expected loss is $50\% \text{ H } \$100,000 = \$50,000$, and his costs of trial are \$25,000. Here, we observed that mutually beneficial settlements exist, for the plaintiff would accept as little as \$50,000 and the defendant would pay as much as \$75,000. Notice that it is also true that the difference between the plaintiff's estimate of the expected judgment and defendant's is $\$70,000 - \$50,000 = \$20,000$, and that that is less than the sum of their costs of trial, $\$20,000 + \$25,000 = \$45,000$. This is consistent with the italicized statement above. Moreover, we observed that if the defendant's estimate of the expected judgment is \$20,000, he would pay only \$45,000, so that no settlement exists. In this case, notice that the difference in the plaintiff's and the defendant's estimates is $\$70,000 - \$20,000 = \$50,000$, which exceeds the \$45,000 sum of litigation costs, so the nonexistence of a settlement amount is again consistent with the italicized statement above.//

3.2 Interpretation of the model. A number of comments may be made to help us interpret and understand the foregoing model.

²⁶To be explicit, let p_P be the plaintiff's estimate of the probability of winning w at trial and p_D be the defendant's probability estimate that the plaintiff will prevail; and let c_P and c_D be their respective trial costs. Then the plaintiff's minimum settlement amount is $p_P w - c_P$, and the defendant's maximum settlement amount is $p_D w + c_D$. Hence, a settlement is possible if and only if $p_P w - c_P \neq p_D w + c_D$, which is to say, if and only if, $p_P w - p_D w \neq c_P + c_D$.

(a) *Does the existence of a mutually beneficial settlement amount imply that settlement will occur?* Although we know that there cannot be a settlement when a mutually beneficial settlement amount does *not* exist, what can be said about the outcome when a mutually beneficial settlement amount does exist? The answer is that there may or may not be a settlement, depending on the nature of bargaining between the parties and the information they have about each other. That will be discussed below, in sections 3.3 and 3.4.

(b) *Parties' beliefs.* The effect of parties' beliefs on the existence of mutually beneficial settlement amounts, and thus on the tendency toward settlement, can be easily understood from the italicized conditions stated in the preceding section. Specifically, the greater the amount by which the plaintiff's estimate of the likelihood of winning exceeds the defendant's, the smaller the tendency toward settlement, as it is the excess of the plaintiff's expected judgment over the defendant's expected payment that leads to trial, according to the italicized condition. It is important to emphasize that what leads to trial is not that a plaintiff is confident of winning, but rather that he is *more* confident than the defendant thinks he has a right to be. A plaintiff's belief that he is very likely to win does not itself suggest that trial will occur, as might naively be thought. If the plaintiff is likely to win, it is true that he will ask for more in settlement from the defendant than otherwise, but it is also true that if the defendant agrees that the plaintiff is likely to win, the defendant will be willing to pay more in settlement. What makes for trial is a refusal of the defendant to pay what the plaintiff demands, and this will be the case when the defendant does not believe the plaintiff's demand is warranted, which is to say, when the defendant holds different beliefs about the expected trial outcome.

Of course, if the plaintiff assesses his chances of winning as lower than the defendant assesses the plaintiff's chances, there will definitely be a range of mutually beneficial settlements. If, for example, the plaintiff thinks his chances of winning are 30 percent and the defendant thinks the plaintiff's chances are 50 percent, then, given the figures mentioned above, the plaintiff would accept any amount over \$50,000, and the defendant would pay any amount up to \$80,000. (Note also that in such cases, the italicized condition automatically holds, as the difference in the plaintiff's expected gain and defendant's expected loss is negative and therefore definitely is less than the sum of trial costs.)

What would we expect the parties' beliefs about the likelihood of trial outcomes to be? The parties may, and often will, be in possession of different information about a case when it begins. For instance, the defendant may know more about whether he would be found liable than the plaintiff knows. Also, the two sides may initially have in mind different legal arguments that they can make. However, the parties may elect to share information or may be forced to do so (through the discovery process, on which see section 2 of chapter 18), and parties often can independently acquire information that the other side possesses (for instance, the plaintiff could interview witnesses to an accident and learn more about the defendant's behavior). To the degree that the parties do come to similar beliefs, settlement becomes more likely.

We would not expect, though, that beliefs will always tend to converge. For example, the parties might not want to share certain legal arguments with each other, believing that if a trial comes about, they would lose the advantage of surprise in making the arguments; yet in light of these arguments, each party might believe it is more likely to win than the other. Another factor standing in the way of convergence of beliefs is the element of natural optimism about one's

chances.²⁷ Additionally, a divergence of interests between the lawyer and the client (see section 7 of chapter 18) may lead the lawyer to tell the client less than the whole truth.

But, by and large, especially because parties have legal counsel who are likely to be familiar with the same body of law and because of opportunities for acquisition of information about facts material to disputes, we would expect beliefs about trial outcomes to be relatively close by the time of a trial, leading toward settlement.

(c) *Judgment amount.* If the size of the judgment rises, then the likelihood of trial rises, other things being constant. This is because the effect of a divergence in the assessments of the likelihood of winning is magnified if the judgment that would be awarded becomes larger. If there is, for instance, a 10 percent difference in beliefs and the judgment award would be \$10,000, the difference in expected judgments would be only \$1,000, perhaps not enough to exceed the sum of legal costs of trial and thus not enough to cause trial. But if the judgment would be \$1,000,000, the 10 percent difference in trial outcomes would signify a \$100,000 difference in expected awards, and thus be more likely to exceed the sum of trial costs, leading to trial.²⁸

Another point about the judgment amount is that, although we have assumed for simplicity that the parties agree on what the judgment amount would be, that might not be so. Differences in the parties' beliefs about the amount that would be won affect their expected judgments and thus the existence of mutually beneficial settlements, according to the italicized condition. If the plaintiff thinks he would win a larger amount than the defendant thinks, then again this would lead toward trial. Suppose, for instance, that although the plaintiff and the defendant agree that the plaintiff has a 60 percent chance of winning at trial, the plaintiff thinks the amount he would win would be \$200,000 and the defendant thinks the amount would be only \$100,000. Then the plaintiff's expected judgment is $60\% \times \$200,000 = \$120,000$ and the defendant's expected payment is $60\% \times \$100,000 = \$60,000$, a difference of \$60,000. Thus, the italicized condition indicates that unless the sum of their legal costs exceeds \$60,000, there exists no mutually beneficial settlement. This can be directly illustrated; if the plaintiff's trial costs are \$30,000 and the defendant's are \$20,000, the sum would be less than \$60,000, and the minimum acceptable settlement of the plaintiff would be \$90,000 and the maximum the defendant would pay would be \$80,000, thus leaving no room for settlement.²⁹

(d) *Legal expenses.* The larger are the legal expenses of either party, the greater are the chances of settlement, clearly, since the sum of legal costs will rise, and thus the greater will be the likelihood that the sum of legal costs will exceed any excess of the plaintiff's expectation over the defendant's expectation. One would expect legal expenses to rise with the size of the potential judgment. This factor tends to increase the chances of settlement for large stakes cases, and thus works opposite to the tendency just discussed in (c) toward litigation.

(e) *Risk aversion.* When we introduce risk aversion into the basic model, we see that it leads to a greater likelihood of settlement. The reason is simply that a trial is a risky venture

²⁷Loewenstein et al. 1993 and Mnookin 1993 suggest that overoptimism is plausible.

²⁸The condition for trial was seen to be that $p_P w - p_D w > c_P + c_D$. This is equivalent to $(p_P - p_D)w > c_P + c_D$, from which it is apparent that if w rises, trial is more likely.

²⁹If we let w_P and w_D denote the amount that plaintiff and defendant respectively believe that they would win at trial, then the condition for trial becomes $p_P w_P - p_D w_D > c_P + c_D$, from which the statements in text follow.

because its outcome is unknown. This means that to a risk-averse party, settlement is more attractive than it is to a risk-neutral party. Further, as the degree of risk aversion of either party increases, or as the amount at stake increases -- the size of the judgment or the size of legal fees - settlement becomes more likely, other things being equal.

3.3 Models with explicit bargaining and locus of information. The model that we have discussed is simple in two important respects, among others. First, the bargaining process is not explicit; the range of possible settlements is determined, but whether a bargain in the range will be reached, and where so, is not predicted within the model. Second, the origin of differences in beliefs is not explained by the model; it was assumed that the parties somehow come to their beliefs. More sophisticated models of settlement versus trial attempt to remedy these gaps and thus to provide additional insight into the settlement process (but, as will be remarked, they achieve less success than might at first appear).

The most basic version of such models is that in which bargaining consists of a single offer, and the party who makes the offer lacks knowledge about the opposing side. For concreteness, assume that the plaintiff makes a single offer to the defendant, but the plaintiff does not know the probability of defendant liability, whereas the defendant does know this (because, say, he has information about his own level of care). In this situation, we can determine the rational settlement offer for the plaintiff to make, and then whether or not it will be accepted. Consider the following illustration.

Example 4. If the plaintiff prevails, he will obtain a judgment of \$100,000. The cost of trial for the plaintiff would be \$10,000. There are three types of defendants, whom the plaintiff cannot tell apart: a minority group of 10 percent, who would lose for sure; a large group of 60 percent, who would lose with probability 50 percent; and a remaining 30 percent group, who would lose only with probability 20 percent. Thus, the expected gain from trial for a plaintiff depends on which type of defendant the plaintiff in fact faces: If success against a defendant is certain, the plaintiff's gross gain from trial would be \$100,000; if the likelihood of success is 50 percent, the expected gain would be \$50,000; and if the likelihood of success is 20 percent, the expected gain would be \$20,000. If the legal costs for a defendant would be \$10,000, the plaintiff could demand and obtain as much as \$110,000 from the first kind of defendant, \$60,000 from the second type, and \$30,000 from the third. If the plaintiff demands only \$30,000, this will be accepted by all types of defendants, so the plaintiff would obtain \$30,000. If the plaintiff demands \$60,000, this will be accepted by the first two types of defendants but rejected by the third, so that the plaintiff's expected gain would be $70\% \text{ H } \$60,000 + 30\% \text{ H } (\$20,000 - \$10,000) = \$45,000$. If the plaintiff demands \$110,000, this will be accepted only by the first type of defendant, so the plaintiff's expected gain would be $10\% \text{ H } \$110,000 + 60\% \text{ H } (\$50,000 - \$10,000) + 30\% \text{ H } (\$20,000 - \$10,000) = \$38,000$. Hence, the plaintiff will demand \$60,000 in settlement. A consequence of \$60,000 being the rational demand for the plaintiff to make is that in the 30 percent of cases where the defendant would lose only with a 20 percent chance, the defendant will spurn the offer and there will be a trial.//

Note that in this example, and in general, the rational settlement demand for the plaintiff to make may be such that there will be a chance of trial. In essence, the rational demand for the plaintiff to make may not be so low as to produce a yes answer from any and all possible types of defendants; to ask for so little is usually not in the expected interest of the plaintiff.³⁰

³⁰The model under discussion and illustrated in the example is developed in Bebchuk 1984. Its formal nature is roughly as follows. The plaintiff makes a single settlement demand x , knowing only the probability distribution over the

This feature of the outcome, that trial may result, may be considered to be due to asymmetry of information. For if the plaintiff knew the type of defendant he faced, he would ask for a different amount for each type, namely, the maximum amount that that type of defendant would be willing to pay rather than go to trial. It is therefore the asymmetry of information that leads the rational plaintiff to ask for more than some defendants are willing to pay, and thus to the possibility of trial.

3.4 Comments on and interpretation of the foregoing models. (a) *Variation of assumptions about the bargaining process.* If the bargaining process is different from that in the model just discussed, complications may arise, but the fundamental conclusion that trial may result, and that it is due to asymmetry of information is not altered. Important changes in assumptions that have been studied involve the making of offers by parties who possess private information (in the model just discussed, the party who did not possess private information made the offer), sequences of offers and counteroffers, and possession of private information by both the defendant and the plaintiff. In these more complex models, the main conclusion about trial does not change because, as in the model sketched in the previous section, a rational bargainer chooses his offer or demand to obtain the highest expected payoff, meaning that he may have the bad luck to be facing an opponent who will reject his offer. However, a new element that arises in the more complicated models is that parties may learn something about their opponents' information from their bargaining behavior.

(b) *Nature of private information.* Any kind of information that one side does not observe that could affect the other side's willingness to pay has the same general significance as private information about the trial outcome, and in particular, can lead to trial. The information might be about a party's costs of trial, his willingness to bear risk, or his need for funds. If, for instance, the plaintiff does not know what the defendant's trial costs will be and thinks they are probably high, when in fact they are low, the plaintiff may ask for too much and a trial might result.

(c) *Relationship to simple model.* The simple model of section 3.1 is roughly consistent with the explicit models of bargaining and asymmetric information now under discussion. In the simple model, trial resulted if parties' beliefs were sufficiently different, and in the explicit models of bargaining here, trial also results from differences in beliefs, where the differences are due to asymmetry of information. Moreover, the influence of trial costs and of other factors on trial outcomes is similar in both models.³¹

The main difference between the models is, as emphasized, that an explicit account of bargaining -- and thus of the magnitude of settlements and their probabilities -- and of the source of differences in beliefs is given in the explicit models. However, the definiteness of the account

probability p that he will prevail over a particular defendant, who does know his p . Thus, if $pw + c_D < x$, the defendant will reject the demand and the plaintiff will therefore obtain only $pw - c_P$; but if $pw + c_D \geq x$, the defendant will accept the demand and pay x . Let $z(x) = (x - c_D)/w$, the p at which defendants will just accept the offer x rather than go to trial. Therefore, the plaintiff's expected payoff as a function of x is

$$z(x) = \int_a^1 (pw - c_P)f(p)dp + (1 - F(z(x)))x,$$

where a is some minimum probability of plaintiff victory (it is assumed that $aw - c_P > 0$, so the plaintiff will always want to go to trial), f is the density of p , and F is the cumulative distribution of p . The plaintiff chooses x to maximize this expression.

³¹See Bebchuk 1984.

of settlement versus litigation in these explicit models is to some extent misleading. First, the models make use of essentially arbitrary assumptions about whether it is the informed or the uninformed party who makes the offer and about the number of rounds of bargaining; these assumptions substantially influence the probability of settlement and the settlement offers.³² Second, the reasons for the differences in information between the parties are largely unexplained in the models, even though parties often have strong motives to share information (in order to reach settlement) and may be forced to do so (see sections 2 and 3 of chapter 18).

3.5 Actual amount of trial versus settlement. In fact, the vast majority of cases settle. Data on state courts show that over 96 percent of civil cases do not go to trial.³³ Similarly, recent data on federal courts demonstrate that, for fiscal year 2001, almost 98 percent of federal civil cases were resolved without trial.³⁴ These figures, however, may either overstate or understate the true rate: because cases that are not tried may have been dismissed by a court, 96 percent or 98 percent is the settlement rate plus the dismissal rate, not the settlement rate; but because many disputes are settled before any complaint is filed, 96 percent or 98 percent may understate the settlement rate. In any event, it is evident that the vast majority of cases settle.

Note on the literature.³⁵ The simple model used to explain settlement and trial, based on the minimum amount the plaintiff would accept and the maximum that the defendant would pay, was first put forward in Friedman (1969), Landes (1971), and Gould (1973), and further articulated in Posner (1972a) and Shavell (1982c). The explicit models of bargaining and asymmetric information were strongly influenced by Bebchuk (1984), and have been developed by, among others, Reinganum and Wilde (1986), Schweizer (1989) and Spier (1992a).³⁶

4. Divergence between the Private and the Socially Desirable Level of Settlement

4.1 In general.³⁷ The private and the social incentive to settle may diverge for a number of reasons that are related to those explaining the difference between the private and the social incentive to sue (see section 2).

First, the litigants may have a socially insufficient motive to settle because they do not take all of society's trial costs into account. In particular, the parties involved in litigation do not

³²However, Daughety and Reinganum 1993 attempt to explain what the nature of the bargaining process governing settlement negotiation will be.

³³See Ostrom, Kauder, and LaFountain 2001a, 29.

³⁴See U.S. Department of Justice, 2001a, 154, Table C-4.

³⁵For general surveys of the literature on settlement, see Cooter and Rubinfeld 1989, Daughety 2000, and Hay and Spier 1998.

³⁶See also Hay 1995. See as well Farmer and Pecorino 1996 for a survey of asymmetric information models of litigation, and Kennan and Wilson 1993 for a general survey of asymmetric information models of bargaining. See Farber and White 1991, Osborne 1999, and Sieg 2000 for empirical investigations of litigation emphasizing asymmetric information, and see also Ramseyer and Nakazato 1989, and Viscusi 1986a.

³⁷The subject of the social desirability of settlement as opposed to trial, and of possible divergence between private and social incentives to settle, has received very little attention (as opposed to the description of the private incentives to settle, the subject of section 3). However, a mainly informal discussion of the topic is contained in Shavell 1997, and some formal analysis is undertaken in Polinsky and Rubinfeld 1988, Shavell 1999, and Spier 1997.

bear the salaries of judges and of ancillary personnel, the value of jurors' time, the implicit rent on court buildings, and the like; the parties thus save less by settling than society does. For this reason, they may proceed to trial more often than would be socially desirable.

A second reason that the private incentive to settle may be socially inadequate concerns asymmetric information. As discussed in sections 3.3 and 3.4, asymmetric information leads parties to fail to settle because they may misgauge each others' situations. But that a party may incorrectly estimate the other's situation does not imply that social resources should be expended on trial.

A third factor suggesting that private incentives to settle may diverge from social incentives is that settlement affects deterrence. The parties themselves would usually not be thought to consider deterrence as an important factor in deciding between settlement and trial; for them, the event giving rise to a legal action has occurred and it might seem irrational to them to give deterrence of others any weight. (In the standard models, they would give it no weight.) Thus, if settlement were to reduce deterrence, and that were thought undesirable, we would conclude that there is a reason for parties to settle too much. But does settlement reduce deterrence? One would suppose that settlement exerts a diluting effect on deterrence because defendants want settlement. Yet the prospect of settlement is also an advantage to potential plaintiffs (because it means that they do not bear litigation costs or risk) and may, at least in principle, result in suits that would not otherwise occur; thus it seems possible that settlement might increase deterrence. A further complication is that even if we know in what direction settlement will affect deterrence, that does not tell us if the effect is socially desirable, for, as explained in section 1, the private incentive to bring suit could be socially excessive or socially inadequate. Still another complication is that the information brought out at trial may improve incentives as well as affect the expected judgment, and in different ways.³⁸ In the end, therefore, we can say that the effect of settlement on deterrence may result in a divergence between the private and the social motive to settle, but for a variety of reasons we cannot generalize about the nature of the divergence.

Thus, the conclusion to this point is that there are two underlying factors suggesting that parties have an inadequate incentive to settle -- they do not consider the full social costs of trial, and they may misgauge each other's situation -- and an underlying factor of unclear consequence -- that settlement affects deterrence.

4.2 The divergence between private and social incentives to settle continued. Let me review several other factors bearing on the comparison between private and socially desirable settlement incentives that go somewhat outside the type of model that we have been implicitly considering in our discussion.

One issue not so far mentioned is settlement as a means of securing privacy. When parties settle, various facts that would have emerged at trial do not come to the notice of the 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 light and may be willing to pay an extra amount for that reason in order to achieve settlement (the victim may not much care whether the information is revealed). A victim might be embarrassed by the facts of a dispute (suppose the suit is for sexual harrasment) or not want to acquire a reputation as a

³⁸I discuss aspects of this point in section 5.3 below and in section 2 of chapter 19.

troublemaker, and so 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.

Another issue is that valuable interpretations of the law may be made and new precedents established through trial. However, 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 not automatic. Furthermore, a natural conjecture is that the number of cases the courts need to hear in order to modify the law is trivial in relation to the total volume of cases that do go to trial.³⁹ Therefore, it would seem that in the great majority of cases, 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.

A similar issue is that social norms may be validated through trial in the sense that they may be publicly invoked 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 large.⁴⁰

Finally, it is often said that trial is socially desirable because it helps to keep peace in society and promotes solidarity by allowing individuals to air grievances and express their views about issues in dispute. However, this social benefit of trial appears often to be in alignment with the private benefit (when the social peace would be compromised if a person did not air his grievance, he would want to do so). Nevertheless, the social benefit under discussion is not coincident with the private benefit (social peace benefits more than just the aggrieved party), and suggests that the social incentive to have a trial may sometimes exceed the private incentive.

4.3 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 non-binding 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 impasses. 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

³⁹Moreover, 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 in order to interpret the law.

⁴⁰It seems that the potential for teaching morality lessons from trials can be well achieved through adjudication of a small, select group of cases.

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 a factor that 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 in which deterrence is likely to be compromised by the fact that in a settlement the identity of defendants, the fact of harm, and/or important aspects of the defendants' conduct do not become public knowledge.

4.4 Comment: A chief social purpose of the institution of trial is paradoxical -- to foster settlement. Finally, let me add a number of comments about the social purposes of maintaining our system of trial adjudication that deserve emphasis although they are restatements of what has been discussed above. First, the socially desirable fraction of trials relative to settlements is probably small, given the high costs of trials and the often limited and unclear benefits of trials over settlements. That the actual fraction of settlements is about 98 percent is consistent with this point. At the same time, there would be no settlements, and the great social benefits derived from them (notably, deterrence of unwanted behavior and compensation of victims), would not be achieved in the absence of the ability of parties to go to trial (as was discussed at length in section 3). Hence, it may be said that *an 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.*

5. Trial and Litigation Expenditure

5.1 Private incentives to spend on litigation. For a variety of reasons, expenditures will tend to increase a litigant's chances of prevailing at trial or will influence beneficially the magnitude or character of the judgment. A party will generally make a litigation expenditure as long as it costs less than the expected benefit it yields. To assess the expected benefit due to a particular legal step, a party will often have to consider not only the court's reaction to it, but also the reaction of the other litigant to it.⁴¹

5.2 Actual litigation expenditures. Amounts spent on litigation are substantial, as indicated by the statistic that about 1.25 per cent of our gross domestic product is spent on legal services.⁴² It should be noted that much of these expenditures are made before trial, in anticipation of possible trial (it is for expositional convenience that I often speak of them as trial expenditures).

5.3 Social versus private incentives to spend on litigation. There are several sources of divergence between social and private incentives to spend during litigation. First, litigants may well be spending in ways that fully offset each other, and thus that have no social value. A

⁴¹For analysis of expenditures of the two sides during litigation, see Braeutigam, Owen, and Panzar 1984, Katz 1988, and Posner 1972a; these writings are mainly descriptive in character.

⁴²See Statistical Abstract of the United States 2001, 418.

classic instance is when both parties devote effort to legal arguments of roughly equivalent weight but supporting opposite claims, or when both hire experts who produce equally convincing reports favoring opposite assertions.

Second, expenditures that are not offsetting may mislead the court rather than enhance the accuracy of outcomes. For example, a guilty defendant may be able to escape liability for harm for which he was responsible, and this possibility dilutes deterrence. Legal expenditures resulting in such outcomes have a negative social value even though they have positive private value.

Third, expenditures that are not offsetting and that do not mislead courts may not be socially optimal in magnitude. By analogy to what was stressed in section 1 about the bringing of suits, the parties decide on their level of expenditures based on how the expenditures will influence the litigation outcome, without regard to the effect (if any) on incentives to reduce harm. This could lead to expenditures that are too great or that are too small, relative to what is socially correct. Consider a victim's decision to hire an expert to produce a report on 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 impact of the victim's expert. 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. The victim will also fail to consider the effect of his expert's report on deterrence. Victims' ignoring the costs they impose on others may lead them to spend socially excessively on experts, while their not considering the deterrence produced by their experts' reports could result in their not spending enough on experts. Similar observations apply to virtually any type of legal expenditure, whether made by victims or by injurers. Thus, although non-offsetting 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.

An important instance of the possibility that expenditures could be socially excessive concerns the assessment of damages.⁴³ Suppose that the presently estimated harm deviates from the truth by \$100. Then one of the litigants will be willing to spend up to \$100 to prove the correct amount (the defendant will do so if the estimate exceeds the correct level, and the plaintiff will do so if the estimate is too low). It can be shown that the social value of the more accurate estimate tends, however, to be lower than \$100, because the social value of accuracy is based on its effects on incentives. Indeed, there will sometimes be no beneficial incentive effect from more accurate assessment of harm, such as when errors are unbiased and not predictable ex ante by potential injurers. 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 wages. But these expenditures will not improve the incentives of drivers to prevent accidents presuming that they do not know ahead of time 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 to what they are now if instead there is no real inquiry into the scale of harm, and damages are simply set equal to the

⁴³This point is developed in Kaplow and Shavell 1996a.

average harm that victims sustain in that kind of automobile accident.⁴⁴

5.4 Legal policy bearing on litigation expenditure. There are several means of influencing litigation expenditures that exist, given the basic form of legal rules and legal procedure. Expenditures can be encouraged through subsidy and discouraged through monetary disincentives such as 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 inducements.

In addition, litigation expenditures could be controlled through changes in substantive legal rules. A notable example of such a change is one under which damages would be based on a table rather than on an elaborate presentation of evidence. As was suggested above, it may well be that incentives toward desirable behavior would not be much affected were damages based on tables of expected values rather than on individual assessment of harm, so that this rule change could substantially reduce the private expenditures on litigation with little change in deterrence.

Finally, litigation expenditures can be modified through revision of legal procedure. A possibility that would be desirable in some circumstances is for certain types of evidence to be produced not by the parties but by court-appointed experts. Especially where private knowledge of the parties is not needed in order to develop evidence, court direction of the acquisition of information might be more beneficial than the parties', which might both mislead courts and result in duplication of efforts.

⁴⁴For 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, $(1/3)H \$1,000 + (1/3)H \$2,000 + (1/3)H \$3,000$). If instead damages are simply set equal to the average harm of \$2,000, injurers' expected liability given an accident will be the same. Therefore, injurers will behave identically.

Chapter 18

EXTENSIONS OF THE BASIC THEORY OF LITIGATION

I discuss here various extensions of the basic theory discussed in chapter 17. In particular, I investigate negative value suits, voluntary and required disclosure of information prior to trial, shifting of legal fees depending on the outcome at trial, inference from trial outcomes, elements of trial outcomes apart from the judgment, and the roles of lawyers and insurers in litigation. Consideration of these topics should help to round out our understanding of litigation.⁴⁵

1. Negative Value Suits

1.1 Definition and a puzzling aspect of negative value suits. It is often noted that plaintiffs sometimes bring suits that they would not in fact be willing to proceed to litigate. These suits are called *negative value suits* because, for the plaintiffs who bring them, the expected benefits of litigation would be outweighed by the costs of litigation. The reason that the suits are said to be brought nevertheless is that plaintiffs may be able to extract settlements from defendants. This claim about negative value suits raises an immediate question: Why would a defendant pay a positive amount in settlement to a plaintiff who would not actually go through with trial? That is, if the plaintiff does not have a credible threat to proceed to trial, why would a defendant rationally pay a settlement, and if the defendant would not be willing to pay, why would the plaintiff bring suit in the first place? Two basic answers to this question, resting on different assumptions, may be offered.⁴⁶

1.2 Explanations for negative value suits. The first answer concerns the possibility that *defendants do not know whether a plaintiff would or would not be willing to go to trial*. If a defendant lacks this information, then it may be rational for him to settle with a plaintiff because of the possibility that the plaintiff might be one who would be willing to go to trial. The defendant will be willing to pay a positive settlement if the proportion of plaintiffs who would be willing to go to trial (because they would obtain a positive expected return from trial) is sufficiently high relative to those plaintiffs who are masqueraders. In other words, the masqueraders ride on the coattails of those plaintiffs who are willing to go to trial. Let me illustrate.

Example 1. There are two groups of plaintiffs: a minority of 10 percent for whom losses are low, say 20, and the other 90 percent for whom losses are 100. Plaintiffs will collect their losses in judgments if they go to trial. The cost of trial is 30 for all plaintiffs. Thus, the low-loss plaintiffs are not willing to go to trial, for the cost of 30 outweighs their judgment of 20. The other plaintiffs, who would net 70 from trial, will be willing to go to trial.

Defendants are assumed to be unable to tell whether a plaintiff is a low-loss plaintiff or a high loss one -- and thus whether a plaintiff would be willing in fact to go to trial. Let us show that a defendant would rationally offer 70 in settlement in order to achieve settlements with the

⁴⁵However, there are obviously many other topics not discussed here that are of interest in respect to litigation, for example, class actions, and sequential versus joint adjudication of the issues in a case. My object is not to be comprehensive but rather to illustrate how the basic analysis of the last chapter can be elaborated to take into account additional factors of relevance and to treat some of the more important factors.

⁴⁶For a survey of literature on negative expected value suits, see Bebchuk 1998.
Chapter 18 – Page 1

High-loss plaintiffs. Assume that defendants' trial costs would be 25. If a defendant offers 70, he knows that all plaintiffs will settle, so his cost will be 70. A defendant will not offer a positive settlement amount less than 70, for any positive amount less than 70 will be refused by the high-loss plaintiffs, so would only benefit the low loss plaintiffs, who would not have gone to trial. If the defendant offers nothing, then all high-loss plaintiffs will go to trial, so his expected cost will be $90\% \times 125 = 112.5$, which exceeds 70. Hence, it is better for the defendant to offer 70, meaning that he pays the 10 percent of low-loss plaintiffs 70 even though they would not go to trial, in order to achieve settlement savings from the 90 percent of plaintiffs who would go to trial.

If, however, the fraction of low-loss plaintiffs who are masquerading as high-loss plaintiffs is sufficiently large (above 44 percent in fact), defendants will not make a positive settlement offer.⁴⁷//

More generally, the presence of masqueraders will tend to lower the settlement offers that defendants make, and this means that there will be more trials than otherwise, because more plaintiffs who would be willing to go to trial will refuse the lower settlement offer and go to trial.⁴⁸

A second explanation for negative value suits arises when the plaintiff would be able to win a judgment if the defendant does not spend an amount on defense (for instance, to develop evidence to defeat a groundless claim) -- even though the plaintiff would lose if the defendant does spend this amount (because the claim is groundless). In such a situation, the defendant will rationally pay a positive amount in settlement in order to avoid the defense cost. Thus, even though the plaintiff would not bring the suit if there were no chance of extracting a settlement, the plaintiff will bring the suit. An example will clarify this point.⁴⁹

Example 2. A plaintiff can bring suit at a cost of 10. If the plaintiff does bring suit, he will obtain a judgment of 100 if the defendant does not defend himself. If the defendant does defend himself, at a cost of 30, he will prevail and the plaintiff will obtain nothing. Note, therefore, that if there were no settlement negotiations, the plaintiff would not bring suit: For if he did so, he would spend 10, then the defendant would rationally defend, and the plaintiff would definitely lose.

However, once the plaintiff brings suit, he will be able to extract as much as 30 from the defendant in a settlement. The defendant would, for example, pay 25 in order for the plaintiff to drop the case, for otherwise the defendant would have to spend 30 to avoid paying the plaintiff 100. Knowing that he can extract a positive settlement from the defendant, the plaintiff may bring suit, and he will do so as long as the settlement he can extract exceeds the cost of bringing suit of 10.//

The key assumptions here that allow the plaintiff to extract a settlement, even though he

⁴⁷At 44 percent, the expected cost to a defendant who offers nothing is $56\% \times 125 = 70$, just equal to his cost if he offers 70.

⁴⁸In our example, this phenomenon did not occur because there was only one group of plaintiffs who would be willing to go to trial. But if there had been more than one group, or a continuum, then the presence of plaintiffs who are unwilling to go to trial can be shown generally to reduce the magnitude of any positive offer the defendant would otherwise make. This is shown in Bebchuk 1988, who first developed the point under discussion; see also Katz 1990a.

⁴⁹The point under discussion, as illustrated in the example, is developed in Rosenberg and Shavell 1985.

would lose if the defendant spent on defense, are two: first, that it is cheap for the plaintiff to bring suit and thereby create the threat to win, and second, that it is relatively expensive for the defendant to defend himself. This is the situation in a significant number of contexts, but is not the situation in many others, where the plaintiff would not win easily just because he brought suit and would have to go to considerable additional expense to prevail at trial.⁵⁰

1.3 Are negative value suits socially undesirable? Although negative value suits often are socially undesirable, it cannot be said that they are necessarily socially undesirable.

Consider first a suit that is without merit in the sense that the plaintiff definitely would not prevail if the facts were known. For example, imagine that the plaintiff did not really suffer injury. If such a plaintiff brings a negative value suit (as he might for either of the reasons discussed above), this is socially undesirable. Notably, it tends to distort incentives, for example, makes parties who might face negative value suits take excessive precautions or discourage them from engaging in activities for which the benefits exceed the true social costs. Such suits also generate litigation costs, and impose risk, to no social purpose.

Next consider suits that have possible merit, and perhaps suits that the plaintiff is certain to win, but which are nevertheless negative value suits because the trial costs exceed the expected judgment. These suits may or may not be socially undesirable. As was discussed in section 2 of chapter 17, the private incentive to bring suit may diverge from the social, and it is possible that suits that victims would not be inclined to bring, because their harm and thus their awards would be less than their trial costs, would be socially good for them to bring because of the beneficial incentives that they would create. If it so happens that such victims find that they can in fact bring suits because, say, they can pretend to be high-loss victims and obtain positive settlements, this may be socially desirable because it compensates implicitly for the underdeterrence of low harms that would ordinarily occur due to litigation costs. This illustrates that negative value suits might be socially desirable, even though the usual expectation is that they would not be.⁵¹

2. Sharing of Information Prior to Trial

2.1 Motive to share information. In the discussion of settlement versus trial of chapter 17, I assumed that the information of parties was somehow exogenously determined: Either information was in the background and influenced parties' perhaps disparate beliefs, or else information was explicitly presumed to be asymmetric. However, litigants in general have strong motives to share information with each other prior to trial, in order to foster settlement and to improve its terms.⁵² A plaintiff, for example, would want to show the defendant information

⁵⁰Although the two explanations given in this section are the major ones for negative value suits, there are others. Bebchuk 1996 shows that the fact that legal expenses are made in stages over time may lead to negative value suits. In particular, if a plaintiff brings suit and spends funds prior to trial, then just before trial, having already spent some funds, the *remaining* trial cost may be less than the expected judgment. Hence, at the point just before trial, he may have a credible threat to sue. This implies that, at the beginning, he may have a credible threat to sue and be able to extract a settlement even though, were he to go to trial, his total cost would exceed his expected judgment.

⁵¹For examination of how negative value suits might be discouraged, see Katz 1990a and Polinsky and Rubinfeld 1993, 1996.

⁵²The motive to share information prior to trial was first analyzed in Shavell 1989, which this section largely summarizes.

establishing that his losses were in fact higher than the defendant believes, so as to secure a settlement and a higher one than he could otherwise obtain. If a defendant thinks that a plaintiff who does not reveal information about his losses probably suffered losses in the neighborhood of \$10,000, but the plaintiff really experienced losses of \$20,000, the plaintiff will want to establish to the defendant that his losses are \$20,000 in order to induce the defendant to pay more in settlement and perhaps avoid an impasse leading to trial. (An essentially equivalent way of expressing this point is to observe that the plaintiff will want to reveal that his losses are \$20,000 in order to avoid a negative inference that the defendant would make from his failure to disclose information that his losses were relatively high.⁵³) Likewise, a defendant would want to show the plaintiff evidence pointing toward the defendant's lack of legal responsibility in an accident, in order to convince the plaintiff to accept a lower settlement offer.

The incentives that parties have to reveal information to one another tend to produce significant voluntary disclosure of information. Indeed, it can be shown that, due to rational voluntary disclosure of information, there will *always* be settlement in a benchmark model of litigation in which there is one-sided private information that that side can choose to disclose.⁵⁴ In other words, the fact that there is initial asymmetry of information will not lead to trial when the party with private information is able to share it.⁵⁵ This point means that the explanation for the occurrence of trial must rest on a deeper understanding of the litigants' situation than the observation that one side possesses private information.

⁵³The revelation of information to avoid a negative inference from silence is sometimes referred to as the unraveling phenomenon. On this, see the survey in Gertner 1998.

⁵⁴Suppose for concreteness that plaintiffs have private information about the magnitude of their losses. In the model, the assumptions are as follows. The plaintiff first decides whether or not to reveal his losses to the defendant. The defendant then makes a settlement offer. The plaintiff either accepts the offer or goes to trial. Of course, each side acts rationally in his self-interest. One aspect of rationality is that, when the defendant decides how much to offer to a silent plaintiff, the defendant takes into account the probability distribution of losses among silent plaintiffs (which is to say, makes a negative inference about their true losses -- why else did they remain silent?).

⁵⁵This is true even though, in general, the sharing of information is *not* complete. Consider the following illustrative example. Plaintiffs' losses are distributed between \$10,000 and \$50,000, plaintiffs' trial costs are \$3,000, and defendant's trial costs are \$5,000. Now it is true that a situation in which too large a range of plaintiffs are silent cannot occur, because the putatively silent plaintiffs with relatively high losses would choose to disclose their losses in order to obtain a higher amount. For instance, suppose that all plaintiffs with losses between \$10,000 and \$30,000 are silent and that silent plaintiffs receive an offer of \$18,000 (the defendant will choose his offer knowing the probability distribution of losses of silent plaintiffs). Then plaintiffs at the high end, such as those whose losses are \$30,000, would choose to reveal their losses, for if such a plaintiff revealed his losses, he could obtain at least \$30,000! \$3,000 (his litigation costs) or \$27,000 from the defendant. This "unraveling" effect, however, will not result in the topmost silent plaintiffs revealing their losses if the silent group is sufficiently small. For instance, suppose that the silent group is plaintiffs between \$10,000 and \$15,000 and that the defendant's offer to silent plaintiffs is \$12,000. Then the topmost plaintiff whose losses are \$15,000 would not benefit by revealing his losses -- for he would then receive an offer of \$15,000! \$3,000, or \$12,000. Moreover, it can be shown that the defendant might well not want to reduce his offer below \$12,000, because even though this would save him money from the majority of silent plaintiffs, he would lose more on account of the minority who would reject and go to trial, as he would then have to pay his own litigation costs.

2.2 Why some information is not shared and trial may result. In spite of the incentives to share information, some information will not be shared and this may lead to trial.⁵⁶ There are two principal reasons why information might not be shared. The first is simply that information may be difficult to share in a credible way, even though a party wants to do that. For instance, a plaintiff might know that his losses are \$20,000, but not be able to demonstrate this during settlement negotiations (because, say, experts will have to be hired by the time of trial to demonstrate the losses). Hence, trial could result because the defendant might make an offer based on his belief that the plaintiff's losses are only \$10,000.

A somewhat subtle coattails effect follows when some parties are unable to share information credibly: Certain parties who are able to reveal their information will decide to remain silent, so that they can be mistaken for those who are unable to reveal their information. For example, if the group of plaintiffs who are unable to demonstrate their information credibly have losses that average \$20,000, then a plaintiff whose losses are only \$5,000 and who is able to reveal them might well choose to remain silent, so that he will be treated more like the plaintiffs with losses that average \$20,000 and thus will be offered more in settlement than he would obtain if he revealed his low losses. Therefore, as a general matter, when some parties are unable to reveal information, it will become advantageous for other parties who could reveal their information to remain silent.⁵⁷

The second major reason why information might not be revealed is that revelation of information may reduce its value to a party because the opposing party may be able to counter it at trial once that party has foreknowledge of it. For example, the plaintiff might not reveal information showing that his losses are \$20,000 because the defendant will then have the time to find an expert who can cast doubt on the plaintiff's evidence so as to reduce what he can collect at trial to \$15,000. The plaintiff's withholding of this information could lead to trial.⁵⁸

3. Forced Disclosure of Information Prior to Trial: Discovery

The courts may require that a litigant disclose certain information to the other side; that is, one litigant may enjoy the legal right of *discovery* of information held by the other side.

3.1 Effect of discovery is only *in addition* to that flowing from voluntary disclosure.

It is commonly believed that the right of discovery significantly increases the likelihood of settlement because it reduces differences in parties' information. But, as just emphasized in section 2, there may well be substantial voluntary sharing of information, so that the influence on settlement of compulsory disclosure through discovery will not necessarily be significant. Indeed, the effect of discovery on settlement is nonexistent in the benchmark model in which the

⁵⁶The reasons to be given here for why parties do not reveal information are distinct from that discussed in the previous note, which was a reason for failure to disclose, but not one that leads to trial.

⁵⁷This in turn can increase the amount of trial. In the example of the low-loss plaintiffs who decide to remain silent and join the group who are unable to reveal their information, the rational offer that defendants make will be lower than otherwise (since defendants will know that some silent plaintiffs are silent just because their losses are low). And this reduced settlement offer will be rejected by more of the silent plaintiffs who cannot reveal their losses, so the likelihood of trial will rise.

⁵⁸Note in particular that if the plaintiff reveals the information, his threat becomes one of winning only \$15,000 at trial, so the defendant could offer him as little as \$12,000 (that is, \$15,000 less his \$3,000 of litigation costs). If the plaintiff is silent and goes to trial, he is better off because his net gain is \$17,000.

party with private information is able to disclose it, for in that situation settlement always occurs due to voluntary disclosure.⁵⁹

To understand how discovery affects outcomes, consider the reasons why information might not be voluntarily shared. The first reason that I noted above, in section 2.2, was that a party might be unable to demonstrate his information credibly, such as a plaintiff who is unable to establish his losses until an expert is hired. Discovery cannot have an effect on this type of plaintiff because, by hypothesis, he is unable to establish his losses (since he does not yet possess an expert's report). However, I also noted the coattails effect, that some parties who are able to disclose information would choose not to do so (such as low-loss plaintiffs who could reveal their losses). These parties would be forced by discovery to reveal their information. Moreover, just because these parties would reveal their information, there would be an indirect, complex effect on settlement offers that would tend to raise the settlement rate.⁶⁰

Next consider the more direct effect of discovery on parties who are able to disclose information but do not do so because it could be countered at trial. These parties will be forced to reveal their information by discovery, and that will clearly lead to more settlement.

3.2 Discovery as a costly threat. An important point about discovery that is wholly apart from its effects on information transmission is that obeying discovery requests is often expensive because significant time and resources may be needed to produce the desired information. This raises questions about the use of discovery requests as a threat, for the costs of compliance with discovery requests are, under our current system, generally borne by the side asked to comply.

3.3 Discovery and social optimality. The question of the social optimality of discovery involves a number of issues. As just noted, to the extent that the side requesting discovery does not bear its costs, there is reason to believe that the private incentive to engage in discovery may be socially excessive. Also, the private incentive to obtain discovery may diverge from the social benefits of discovery. The social benefits of discovery involve, among other elements, the promotion of deterrence through development of better information (which will affect both settlements and trial outcomes), and more effective and lower cost trials, owing to the parties' having relevant information in advance of trials. Because the private and social benefits generally diverge, the suspicion is that the private incentives to engage in discovery may often be socially undesirable, either excessive or insufficient, depending on the particulars of the situation.⁶¹

⁵⁹The first model of discovery was Sobel 1989, but his analysis does not determine behavior of parties who voluntarily disclose information, so the effect of discovery against the background of voluntary disclosure is not ascertained. The first model to determine the difference that discovery makes, given that there will be voluntary disclosure of information without discovery is Shavell 1989, the main points of which are noted here. See also subsequent related work of Cooter and Rubinfeld 1994, Hay 1994, Mnookin and Wilson 1998, Schrag 1999, and Shepherd 1999, and see the survey Rubinfeld 1998.

⁶⁰Because low-loss plaintiffs who can reveal their losses would be forced to do so, defendants would raise their settlement offers to plaintiffs who cannot supply information about their losses. That will raise the settlement rate because more of the plaintiffs who are unable to disclose their losses will accept the settlement offer rather than go to trial.

⁶¹One occasionally encounters the view that discovery is socially undesirable if its costs exceed the benefit to the party who obtains discovery. This is not a correct view, for the private benefits to the party seeking discovery, an improvement in settlement terms or the trial outcome, have only a weak connection to the social benefits of discovery.

4. Shifting of Legal Fees to the Loser at Trial

Thus far we have assumed that parties bear their own legal fees, which is often referred to as the American rule for the allocation of legal fees. By contrast, the bearing of legal fees can be made to depend on the outcome of trial.⁶² A major form of such fee-shifting is the so-called English rule, under which the loser at trial bears his opponent's legal fees as well as his own. We will restrict our attention here to this form of fee-shifting.⁶³

4.1 Effects of fee-shifting on the bringing of suit. Fee-shifting has implications for the incentive to sue. Fee-shifting leads to more suit in situations in which plaintiffs are likely to win, for this means that plaintiffs are unlikely to have to pay their legal expenses. Suppose that the probability of plaintiff victory is 70 percent, that the judgment amount would be \$100,000, and that the plaintiff's trial costs and the defendant's would each be \$80,000. Then under the American rule, the plaintiff would not bring suit because his costs of \$80,000 exceed his expected gain of \$70,000. Under the English rule, however, he would bring suit because his expected costs would be only 30% of \$160,000 or \$48,000.

Conversely, fee-shifting leads to less suit in situations in which plaintiffs are unlikely to win. Suppose that the plaintiff's probability of victory is now only 30 percent, that the judgment amount would be \$100,000, that the plaintiff's legal costs would be \$10,000, and that the defendant's legal costs would be \$80,000. Under the American rule, the plaintiff would sue, \$10,000 being less than the \$30,000 expected gain, but under the English rule the plaintiff would not sue because his expected expenses would be 70% of \$90,000 = \$62,000.

As a general matter, the range of high probabilities of prevailing for which fee-shifting increases suit (and the complementary range of low probabilities for which fee-shifting reduces suit) depends on the legal costs that the parties would bear at trial. If they would each bear equal trial costs, then 50 percent is the threshold probability: suit is more likely under the English rule when the likelihood of victory exceeds 50 percent. And if the plaintiff's costs are less than the defendant's, the threshold probability is higher.⁶⁴

The foregoing analysis presumes that plaintiffs are risk neutral. If plaintiffs are risk averse, then superimposed on the effects just mentioned is a disinclination to bring suit, because the risk of trial is increased by the fact that the sum of the legal fees depend on the trial outcome.

⁶²For a survey of economic literature on fee-shifting, see Hughes and Snyder 1998. For empirical study of fee-shifting, see, for example, Hughes and Snyder 1995, Kritzer 1984, and Snyder and Hughes 1990.

⁶³There are other forms of fee-shifting, and the analysis of them would in many respects follow in a straightforward way from what will be said here about the English rule. Under plaintiff-favoring fee shifting, the defendant has to pay the plaintiff's fees if the plaintiff prevails, but the plaintiff does not have to pay the defendant's fees if the defendant wins. Under defendant-favoring fee-shifting, the converse is true. Fees may also be shifted in a more complicated way, depending on whether the amount of the award at trial exceeded or fell below a settlement offer; for economic analysis of such "offer-of-settlement" rules, see Bebchuk and Chang 1999, Miller 1986, and Spier 1994a. For a general description of the actual use of fee-shifting, see Derfner and Wolf 1995.

⁶⁴To amplify, let p be the probability of plaintiff victory, x the judgment amount, and c_P and c_D the plaintiff's and the defendant's trial costs respectively (and let us abstract from issues relating to settlement). Under the American system, the plaintiff will sue if $c_P < px$ or, equivalently, if $x > c_P/p$. Under the English rule, the plaintiff will sue if $(1-p)(c_P + c_D) < px$, or equivalently, if $x > (1-p)(c_P + c_D)/p$. Hence, there will be more suit (a greater range of x will result in suit) under the English rule if and only if $(1-p)(c_P + c_D)/p < c_P/p$, which is to say, if and only if $p > c_D/(c_P + c_D)$. In particular, if $c_P = c_D$, there will be more suit under the English rule if $p > .5$. The effects of fee-shifting on the incentives to bring suit are examined in Shavell 1982a.

If the judgment amount is \$100,000 and the legal fees of each side are \$80,000, the plaintiff obtains \$100,000 if he wins and loses \$160,000 otherwise -- meaning a difference of \$260,000 in result under the English system. Under the American system, the difference in result is only \$100,000, a significantly smaller risk. Another factor (to be discussed below) is that the English system may turn out to increase legal fees, and this in turn also works to reduce the frequency of suit.

4.2 Effects on settlement versus trial. Given that suit has been brought, fee-shifting has an underlying tendency to increase the likelihood of trial. The reason, in essence, is that fee-shifting magnifies the effects of differences of opinion about the trial outcome. If a plaintiff is more optimistic about the chances of winning than the defendant thinks is correct, then fee-shifting will raise the plaintiff's demand relative to the defendant's offer because a plaintiff victory will now mean not only a judgment for him, but also escaping legal fees.

Example 3. Suppose that the plaintiff thinks his chances of winning \$100,000 are 80 percent, that his trial expenses would be \$50,000, that the defendant thinks the plaintiff's chances of prevailing are only 30 percent, and that his trial expenses would be \$40,000. Then under the American rule, there would be room for settlement: The plaintiff's minimum acceptable amount is $\$80,000 - \$50,000 = \$30,000$, which is less than the defendant's maximum acceptable amount of $\$30,000 + \$40,000 = \$70,000$. Under the English rule, however, there would be a trial, for the plaintiff's minimum acceptable amount is $\$80,000 - 20\% \times \$90,000 = \$62,000$ (note that \$90,000 is the sum of the parties' legal costs), whereas the defendant's maximum acceptable amount is $\$30,000 + 30\% \times \$90,000 = \$57,000$.

This example illustrates that when the plaintiff's estimate of the probability of winning is higher than the defendant's estimate, fee-shifting means that each side believes it is relatively likely to lay off its legal expenses on the other side, which widens the difference between them and may lead to trial.⁶⁵

However, the factor of risk aversion counters the tendency toward trial. Because fee-shifting makes trial a more risky proposition, fee-shifting tends to promote settlement. Also, the likelihood that fee-shifting leads to larger legal fees also makes settlement more likely. Therefore, the effect of fee-shifting in the end might not be to increase trial, even though that is its influence on risk-neutral parties.

4.3 Effects on trial expenditures. Fee-shifting will generally lead parties to spend more on legal fees than they otherwise would, for two reasons. First, fee-shifting means that a party will not necessarily have to pay the bill for legal services that he orders, making legal services

⁶⁵To show this result in the simple model of settlement versus litigation, let p_P be the plaintiff's estimate of the probability of winning w at trial and p_D the defendant's probability estimate, and c_P and c_D their respective trial costs. Under the American system, the plaintiff's minimum settlement amount is $p_P w - c_P$, and the defendant's maximum amount is $p_D w + c_D$. Hence, a trial occurs if $p_P w - c_P > p_D w + c_D$, which is to say, if $p_P w - p_D w > c_P + c_D$. Under the English system, the plaintiff's minimum settlement amount is $p_P w - (1 - p_P)(c_P + c_D)$ and the defendant's maximum is $p_D w + p_D(c_P + c_D)$, so that trial occurs if $p_P w - p_D w > (1 - p_P)(c_P + c_D) + p_D(c_P + c_D)$. We want to show that trial occurs more often under the English rule. Hence, assume that trial occurs under the American rule. Then it must be that $p_P > p_D$, which implies that $c_P + c_D > (1 - p_P)(c_P + c_D) + p_D(c_P + c_D)$, from which the conclusion follows. This conclusion was demonstrated by Posner 1973a and Shavell 1982a. The same conclusion holds in the asymmetric information bargaining model for closely related reasons; see Bebchuk 1984. Similar conclusions hold for fee-shifting favoring defendants and fee-shifting favoring plaintiffs, as shown in Shavell 1982a.

effectively cheaper. If the plaintiff has a lawyer spend \$1,000 more of time and expects to win with a probability of about 70 percent, the odds that he will have to pay for the extra \$1,000 of services are only 30 percent, so their effective cost to him is only \$300. Hence, if the expected payoff from the services is \$500, he would spend the \$1,000 under fee-shifting, but he would not have done so in its absence.

The second reason that fee-shifting leads to greater expenditures is that fee-shifting increases the payoff from winning, because the fees themselves are at stake in the trial outcome. This can be a significant factor when the fees are substantial in relation to the judgment. Recall the example where the judgment amount is \$100,000 and the legal fees of each side are \$80,000. Here, because the plaintiff obtains \$100,000 if he wins and loses \$160,000 otherwise, there is a \$260,000 difference between winning and losing under the English system, as opposed to only a \$100,000 difference under the American system. Hence, the value of an expenditure to increase the chance of winning is 2.6 higher than its value under the American system, reinforcing the other effect that leads to higher spending under the English system.⁶⁶

4.4 Social desirability of fee-shifting. It is not possible to draw general conclusions about the social desirability of fee-shifting, although statements can be made in particular circumstances. To illustrate that fee-shifting could be beneficial, suppose that victims of low magnitude harms are not bringing suit even though they would win, because legal fees exceed the damages they would collect; and suppose that because injurers escape suit, they fail to take cheap precautions that would be socially desirable. Then fee-shifting might well be socially beneficial, because it would result in victims bringing suit, knowing that they would not have to bear legal fees, and this threat would induce injurers to exercise precautions. To show that fee-shifting might be socially undesirable, suppose that in an area of litigation, the major effect of fee-shifting would be to increase the volume of suit but that this would not affect injurers' incentives because there is little more that they can do to reduce risk. Then fee-shifting would be socially undesirable because it would increase the litigation costs that society incurs without producing a social benefit. It is apparent, therefore, that the social desirability of fee-shifting will depend very much on the context.⁶⁷

It may be noted that a commonly mentioned argument that is made about fee-shifting is that if a party prevails at trial, he has been revealed to be right by the court, and therefore, being right, should not be penalized by having to pay legal fees. This argument, note, makes no reference to the effects of fee-shifting on behavior -- on suit, settlement, or trial expenditures -- so that it is manifestly incomplete.⁶⁸

5. Difficulty of Statistical Inference from Trial Outcomes

A question of interest is whether cases that go to trial are representative of the underlying population of cases, and notably, whether the likelihood of plaintiff victory at trial or the amounts won are similar to those that characterize the cases that settled. This question is

⁶⁶The effects of fee-shifting on expenditures at trial are discussed in Braeutigam, Owen, and Panzar 1984, Hause 1989, and Katz 1987.

⁶⁷This point is emphasized by Gravelle 1993.

⁶⁸The relevance of such arguments as this, based on notions of fairness of outcomes, will be discussed generally in chapters 26 and 27.

important because, often, the most readily available data is on cases that go to trial, whereas the great majority of cases settle. Unfortunately, as I am about to explain, the cases that go to trial may be very different from the population of cases that settle, so that generalizing from trial cases is difficult and may be misleading.

Consider two examples. First, suppose that 99 percent of defendants are guilty of negligently causing \$100,000 of harm, and that 1 percent are not and have exonerating evidence, but the plaintiffs cannot tell who is guilty and who is not. Then in settlement negotiations, plaintiffs' settlement demands would reflect their belief that virtually all defendants would be liable, and let us say the demands would approximate the \$100,000 judgments that the 99 percent of liable defendants would have to pay. Thus, we would expect that essentially all of the 1 percent of defendants who would not be found liable would go to trial rather than accept the \$100,000 demands of plaintiffs, whereas essentially all of the 99 percent of defendants who would be found liable would accept plaintiffs' settlement demands to avoid litigation costs. In this situation, then, only 1 percent of cases go to trial, and all of them are won by defendants. Obviously, it would be a gross mistake to extrapolate from the uniformity of trial results in favor of defendants to conclude that, in the entire population of cases, defendants are usually not negligent, for the truth is that 99 percent of them are negligent.

Second, suppose that all defendants are liable, that 99 percent of plaintiffs suffered losses of \$100,000, and that 1 percent of plaintiffs suffered losses of \$200,000, but defendants cannot tell them apart at the time of settlement negotiations. Then, by similar logic to that of the previous paragraph, we might expect defendants to offer around \$100,000 as a settlement, that the 99 percent of plaintiffs with \$100,000 losses will settle, but that the other plaintiffs will reject the settlement offer, go to trial, and obtain judgments of \$200,000. In this situation, it would again be a mistake to infer from the \$200,000 judgments made in all the trials that this amount is the typical loss of plaintiffs, when the truth is that \$100,000 is the typical loss.

These examples illustrate how dramatically trial outcomes can diverge from the actual statistics characterizing the population of all cases, both in terms of who wins and in terms of the amounts awarded. The differences between trial outcomes and the population of cases have to do with the fact that the cases that go to trial are not a random sample from the population of cases, but rather are those cases in which there was a reason -- something not anticipated by one of the sides to the litigation, and thus something atypical -- that led to a bargaining impasse.⁶⁹

6. Elements of Trial Outcomes Apart from the Judgment

We have assumed that the only outcome of a trial is a judgment paid by the defendant and received by the plaintiff, but there are other possibilities that affect litigation. First, a trial outcome may have implications for a litigant beyond the immediate judgment. For example, a firm may believe that a loss at trial would invite a string of future lawsuits; thus, a loss would be

⁶⁹The general point that cases that go to trial are not representative of the underlying population of cases was first emphasized by Priest and Klein 1984. (They also suggested that cases that go to trial would be won by plaintiffs approximately 50 percent of the time, regardless of the underlying population of cases. This somewhat surprising conclusion of theirs is correct given their assumptions; but it is not borne out in fact, and it does not hold theoretically (as, for example, the text illustrates) under general assumptions about the population of cases and bargaining over settlement and trial.) See also Eisenberg 1990, Eisenberg and Farber 1997, Hylton 1993, Shavell 1996, Waldfogel 1995, and Wittman 1985. For a survey of literature, see Waldfogel 1998.

more costly for it than the judgment.⁷⁰ This would tend to make settlement more likely than otherwise, as it would raise the amount the defendant firm would be willing to pay in settlement. Another possibility is that a trial might prevent future lawsuits, by revealing the defendant to be tough-minded. This would make trial more likely than we had suggested.

Second, a litigant may care whether a trial is held per se: A plaintiff might, say, wish the defendant to be exposed to public scrutiny. This would make both suit and trial rather than settlement more likely. Or a party might want to avoid a trial because it would result in the airing of embarrassing facts or the disclosure of valuable business information, which would tend to make trial less likely. In all, it is apparent that the consequences of trial apart from the literal judgment may affect the tendency to sue and the decision over settlement versus litigation.

7. Role of Lawyers in Litigation

Although we have so far spoken of “the” plaintiff and “the” defendant, each litigant is typically represented by a lawyer, and the lawyer and his client are distinct individuals. Let me briefly indicate how recognition of the lawyer as distinct from the client affects the prior analysis of the bringing of suits, of settlement versus trial, and of trial effort. I will assume that the lawyer has superior information about the law and about the likely outcome of a case than does the client, and, to begin with, that parties are risk neutral. Two basic payment arrangements will be considered: hourly fees for services, and an outcome-based or *contingency fee* arrangement, under which the plaintiff’s lawyer’s fee equals a percentage of the judgment or settlement amount (and the lawyer receives no fee for time worked).

7.1 Description of lawyer’s decisions. Regarding the incentive to sue, it is clear that under an hourly fee arrangement, lawyers will have incentives to accept cases in which the expected gains are less than the legal costs, for lawyers will be paid regardless of the outcome. Under contingency fees, this effect will be dulled because, if a case is not promising, the lawyer will not want to spend his time on it for a low expected fee. However, the lawyer might be unwilling to take a case for which the expected gain exceeds the legal costs, as the lawyer obtains only a fraction of the gains but bears the full costs. In summary, relative to the benchmark of maximizing the expected gain minus legal costs, the incentive to take cases under the hourly fee is excessive, and the incentive to take cases under the contingency, outcome-based arrangement is inadequate.

Next, consider settlement negotiations. In this context, the lawyer who obtains hourly fees will have an incentive to go to trial if the lawyer wants more work, so that he will be inclined to go to trial more often than he would were the goal to maximize expected gains minus legal costs. (However, if trial work is less attractive than other available work -- perhaps because trial is a more intense type of work -- then the incentive would be in the opposite direction and the lawyer would push the client to settle.) Under contingency fee arrangements, the lawyer’s incentives are different, leading the lawyer to press for settlement more often than when the settlement offer exceeds the expected judgment net of litigation costs, as the lawyer bears all the litigation costs but obtains only a percentage of the settlement. Thus, relative to the benchmark of maximizing the expected gain minus legal costs, the incentive to go to trial under the hourly fee is excessive, and the incentive to go to trial under the contingency fee contract is too low.

For similar reasons, if a case has gone to trial, then under the hourly fee arrangement, the

⁷⁰See Che and Yi 1993.

lawyer will want to work more than may be worthwhile in terms of the expected judgment. And under the contingency fee arrangement, this is not likely; the lawyer will tend to work less than would be worthwhile in terms of the expected judgment, for, again, the lawyer obtains only a fraction of that for himself but bears the entire litigation costs.

7.2 Mutually beneficial lawyer-client arrangements. What can be said from the foregoing description about mutually beneficial contracts between lawyers and clients? We would expect contracts to gravitate toward those that result in maximization of the expected gains less litigation costs, for that is the pie that lawyers and clients have available to share. From this perspective, it appears that either hourly fee or outcome-based contingency arrangements could be superior to the other, because either could lead to a higher expected gain minus litigation costs. In particular, the hourly fee arrangement would be superior if the excess incentives under it -- to take cases, and then to go to trial and work more during trial -- is not as important a drawback as the inadequacy of incentives under contingency fees.

However, it should be noted that the incentive problems facing clients are reduced by their ability to obtain information about lawyer behavior. Clients can seek second opinions, and sometimes lawyer behavior will become apparent and harm a lawyer's reputation or even result in suit for malpractice. These factors mitigate the incentive problems under different fee arrangements.

When risk-bearing is taken into account, additional considerations become relevant. Under hourly fees, lawyers bear little risk relative to what they bear under outcome-based contingency fees. Hence, if lawyers are more risk averse than clients, the mutual appeal of hourly fees will be increased, whereas if lawyers are less risk averse than their clients, the mutual appeal of contingency fee arrangements will rise. With regard to the latter, it is often emphasized that, when clients' assets are limited, having to bear legal fees (or the inability to pay them) would discourage suits under an hourly fee arrangement, and only under contingency fee arrangements will they be brought.

In any event, in fact, hourly fee arrangements seem to be the dominant form of compensation, and contingency fees seem to be limited mainly to tort actions. Moreover, contingency fee arrangements are regulated and seem to be frowned upon.⁷¹ Thus, the relative absence of contingency fees, based on trial outcomes, may not really reflect the desires of lawyers and clients.⁷²

7.3 Socially desirable arrangements. We can also inquire about the social welfare implications of lawyer-client arrangements. Of course, the general argument that contracts raise the well-being of the contracting parties suggests that it is socially desirable to allow lawyers and clients to make the fee arrangements that they wish.

However, as we also know, it may be socially desirable to limit private contracting when contracts have external effects (see section 9 of chapter 13).⁷³ Hence, if lawyer-client fee

⁷¹See, for example, Brickman, Horowitz, and O'Connell 1994, Rubinfeld and Scotchmer 1998, 415-16, and Wolfram 1986, 526-42.

⁷²Literature on the effects and the mutual desirability of lawyer-client fee arrangements includes Dana and Spier 1993, Danzon 1983, Emons 2000, Hay 1996, Miller 1987, Rubinfeld and Scotchmer 1993, and is generally described in the survey by Rubinfeld and Scotchmer 1998.

⁷³Another major reason for intervention in private contracting is lack of information by contracting parties. This applies only if disclosure requirements and/or public provision of information about fee arrangements do not function

arrangements have external effects, it might be advantageous to regulate them. Suppose that contingency fee arrangements lead to more suit because they enable risk-averse clients to bring cases that they would not otherwise have brought, and suppose also that the level of suit is felt to be excessive. Under these assumptions, it might be thought desirable to regulate contingency fee arrangements. Yet such a conclusion overlooks the possibility of using other policy tools to cure externality problems. If the volume of suit is excessive, that can be corrected by imposing fees for bringing suit. This kind of approach, being tailored to a specific externality problem, may often be superior to intervention in contractual relations, for that is likely to deny parties benefits of a broader nature (contingent fee arrangements will help parties for many types of suit, not just those that are excessive).

8. Role of Insurers in Litigation

Insurers tend to have interests in the outcome of litigation. First, plaintiffs may be insured against the losses for which they are suing, and their insurance policies often have clauses allowing the insurers to be reimbursed from payments made by defendants. For example, a plaintiff who sustains losses of \$100,000 due to a negligently caused fire may have a fire insurance policy providing \$60,000 of coverage and thus will be able to collect this amount immediately from his insurer. But the policy may also stipulate that, if the plaintiff wins a judgment or obtains a settlement, the insurer will have the right to up to \$60,000 of that amount (this feature of fire insurance policies would lower insurer costs and thus premiums). Accordingly, the insurer may have an interest in the plaintiff's suit. Second, defendants frequently own liability insurance policies, and defendants' insurers therefore obviously have an incentive to defeat plaintiffs in litigation or to pay them less rather than more in settlements.

8.1 Description of effects. In discussing the effects of insurance, let us consider the example just mentioned, of a plaintiff's policy giving him \$60,000 of coverage against his \$100,000 loss, and allowing the insurer to collect up to \$60,000 from suit or settlement. We also will assume for concreteness that the defendant's liability insurance coverage is in the amount \$70,000. Further, we will suppose that the insurers will bear all legal costs and make litigation decisions. Examination of this example will reveal important aspects of the effects of insurance, and I will offer some comments later about how the conclusions would be altered were one to relax various assumptions of the example. The example in consideration, note, has the feature that the insurance coverage is not full.⁷⁴ This case is significant because, as will be evident, it means that the interests of the insurer and of the insured are each positive and generally different.

Consider first the effects of plaintiffs' insurance on incentives to sue. If the plaintiff's losses are \$100,000 and he collects \$60,000 from his insurer, his gain from suit will only be \$40,000, not \$100,000. Likewise, his insurer's gain from suit will be only \$60,000. Thus, the gain of the plaintiff and of his insurer from suit is, for each separately, only partial when the insurance coverage is less than the plaintiff's losses. Because of the assumption that the insurer bears all the litigation costs, the insurer will have a reduced incentive to bring suit, relative to the

well.

⁷⁴As I have discussed in chapter 5, coverage is often partial, due to moral hazard, the judgment-proof possibility, and administrative costs of insurance.

standard case of a single plaintiff who would gain \$100,000 from suit. However, the plaintiff would very much want to bring the suit to win the extra \$40,000, for he does not have to bear any of the costs of suit. This creates a conflict between the plaintiff and the insurer.⁷⁵ In any event, the incentive to bring the suit is dulled relative to the standard case because the assumption is that the insurer decides whether the suit goes forward.

There is also an important, and opposite, effect of defendants' liability insurance on the plaintiff's incentive to sue: The existence of liability insurance is a spur to suit, for otherwise many defendants would be unable to pay judgments and thus not be sued -- plaintiffs and their insurers will not bring suit if they know that defendants do not have the assets or liability insurance to pay the judgments that they seek.

With regard to the decision about settlement versus trial once a suit has been brought, the situation is complicated, as the following illustration shows.

Example 4. Assume that litigation costs for the plaintiff's side would be \$15,000, that he and his insurer believe the chance of victory is 50 percent, that his insurer would bear litigation costs, and that the insurer would receive the first \$60,000 of any settlement or judgment and the plaintiff the remainder. What would be the plaintiff's minimum acceptable settlement? If there is a trial, the plaintiff's expected return would be $50\% \times \$40,000 = \$20,000$, as he does not bear legal costs. Thus, the settlement would have to be at least \$80,000 to satisfy him (as the insurer would obtain the first \$60,000 of a settlement). Now consider the insurer's minimum acceptable settlement offer. The insurer would net $50\% \times \$60,000 - \$15,000 = \$15,000$ from trial, so it would be willing to accept any amount over \$15,000 in settlement. Hence, the insurer would accept a lower settlement than the plaintiff, owing to its bearing the legal expenses and also to its obtaining only the first \$60,000 of any settlement. Note also that, were the plaintiff a single uninsured party, the plaintiff's minimum settlement offer would be $50\% \times \$100,000 - \$15,000 = \$35,000$, which is in between the insured plaintiff's and the insurer's minimum acceptable settlement offers.

Now consider the defendant's side. As the defendant's liability insurance coverage is for \$70,000, his exposure is for \$30,000. Assume that the defendant's legal expenses would be \$10,000 and would be paid by the liability insurer, and also that the first \$70,000 of any settlement or judgment is to be paid by the insurer. The defendant's expected loss from trial would thus be $50\% \times \$30,000 = \$15,000$. Therefore, the defendant would agree to any settlement amount up to \$85,000 (recall that the liability insurer pays the first \$70,000 of any settlement). The liability insurer would lose $50\% \times \$70,000 + \$10,000 = \$45,000$ at trial, so would pay up to \$45,000 in settlement. Thus, the liability insurer would be less willing to settle than the defendant. If there were a single uninsured defendant, that defendant would pay up to $50\% \times \$100,000 + \$10,000$ or \$60,000 in settlement, which also exceeds what the liability insurer would be willing to pay.//

We have seen that each litigant and his insurer have disparate interests in this example. The plaintiff's insurer is more likely to want to settle than the plaintiff, and the defendant's

⁷⁵Such conflict can sometimes be resolved through bargaining between insurer and insured. For example, if the insurer does not want to press a suit and the plaintiff does, it may be that if the plaintiff makes a contribution toward the cost of suit, then the insurer would agree to go forward with it. Of course, bargaining costs and asymmetry of information may hinder such renegotiation, so that conflicts often remain even where mutually beneficial changes exist. (Comments similar to this one will apply to most of what is said below about differences between insureds and their insurers.)

insurer is less likely to want to settle than the defendant -- the plaintiff's insurer will demand less than the plaintiff would, and the defendant's insurer will pay less than the defendant would.⁷⁶ The same relationship holds between the insurer's demands and offers and those in the standard case of plaintiffs and defendants who do not possess insurance. Thus, it is not evident, in the circumstances of the example, whether settlement will be more likely or less likely than in the standard case.

The effect of insurance on incentives to spend at trial is generally to reduce it. This is for the simple reason that the insurers have less at stake than the full \$100,000, yet bear the entire legal costs. Of course, the plaintiff and the defendant would each like their insurers to spend more than the insurers wish, as the plaintiff and defendant each have some stake in the outcome.

To summarize, we can see that, given our assumption that insurers bear legal expenses and make litigation decisions, insurance coverage leads to conflicts between insurers and insureds, a reduced incentive to sue, and a lower incentive to spend at trial. The specific conclusions about the direction of effects may well be altered if we were to consider different assumptions about the nature of the insurance contract, especially about who bears legal costs and about who makes litigation decisions or about whether agreement about litigation decisions between insurer and insured is necessary.⁷⁷

An important assumption that we implicitly made should be mentioned: that insurers and insureds do not renegotiate during litigation. This assumption might be justified where the costs of renegotiation are significant, or when asymmetry of information between insured and insurer (especially about the likelihood of the trial outcome) would often result in failure to come to an agreement. In any event, if the parties are able to renegotiate and come to an agreement, they would tend to act approximately in the same way that a single litigant would, so that the conclusions of the standard case would be maintained. For instance, it was stated above that the the plaintiff's insurer has a lower incentive to sue than in the standard case because the insurer will gain only \$60,000 of the \$100,000 judgment, whereas the insurer bears all the legal costs. If, though, the plaintiff cooperates and contributes toward legal costs, his interest in obtaining \$40,000 will make their joint willingness to bring suit similar to that in the standard case.⁷⁸

8.2 Mutually desirable insurance contract terms regarding litigation. The parties will tend to elect a type of insurance contract that, in its terms regarding litigation, is jointly beneficial to them. Such contracts will reflect several factors. One is that it is often efficient to have just one party control the litigation and bear the litigation costs, for that reduces the need for costly negotiation and coordination during litigation. An additional factor is that the insured

⁷⁶In our example, it so happens that there will be settlement, for the plaintiff's insurer will accept any amount over \$15,000, and the defendant's insurer will pay up to \$45,000. Of course, if beliefs were different, or if information were asymmetric, settlement might not occur.

⁷⁷For example, if the legal fees are divided pro rata between insurer and insured and they make a joint decision on litigation, their incentive to bring suit would be the same as in the standard case. A systematic account of the conclusions as a function of the type of insurance policy is, however, beyond our scope, and has not been undertaken in the literature to my knowledge. Only a few papers examine litigation-related conflicts of interests between insureds and insurers from an economic perspective; see Meurer 1992 and Sykes 1994. These papers are suggestive, but are limited in nature, focusing on aspects of liability insurer incentives to settle.

⁷⁸It might not be identical to that in the standard case because the plaintiff is risk averse and thus would be less likely to want to engage in the risky venture of suit than the risk-neutral party of the standard case.

will, being risk averse, want coverage not only against loss and liability, but also against legal expenses; this favors insurers bearing legal expenses and controlling litigation. Another factor is that the party with the most experience with litigation is the natural one to assume the responsibility of decisionmaking about litigation and thus to bear litigation costs; this party is generally the insurer, not the insured. These factors all suggest the mutual desirability of insurer control and expenditure on litigation. A factor working in a different direction is that it tends to be jointly beneficial for the litigation decisions to maximize joint expected returns for the plaintiff and his insurer, or to minimize joint payments for the defendant and his insurer. If, however, the insurer is in control of the litigation and bears all the litigation costs, then, as was seen in the last section, the incentives of the insurer generally diverge from those of the standard case in which joint net returns are maximized for plaintiffs or joint costs are minimized for defendants. The character of the mutually desirable terms of the insurance contract thus reflects concerns that are competing in some respects.

8.3 Social desirability of the foregoing contractual terms. The main point to be made about the social desirability of the contractual terms that insurers and insureds select is the analogue of that of section 7.3: Because contractual terms of insurance policies raise the well-being of insureds and insurers, these terms are socially desirable unless they have unwanted external effects (such as on the volume of suit) that cannot be alleviated through some other policy (typically they can be).⁷⁹

It is sometimes suggested that the fact that insurers often display different interests from their insureds is socially undesirable and calls for legal intervention. Consider the situation in which a plaintiff would like the insurer to sue for \$100,000, but the insurer hesitates to do so because it would gain only \$60,000 and would pay the entire legal fees. The reluctance of the insurer to sue when the insured wants suit, and when the expected total gain exceeds the cost of suit, might be seen as justifying a legal response to force the insurer to do what the insured wants.⁸⁰ Such a view overlooks the point that the insurance contract is elected by the insured and the insurer, and that the conflict of interests at the time of possible suit is a byproduct of their contractual terms, which were chosen for their beneficial joint purposes. In the example in question, the terms that give the insurer control over litigation decisions may be warranted because the insurer is best suited to control legal expenditures efficiently (given the scale of its business), and the terms that limit coverage of the insured to \$60,000 (which is what gives rise to the conflict) may be best because they combat moral hazard and fraud. For the courts to intervene, such as to prevent the insurer from maintaining control over litigation, is effectively to hurt the insured *ex ante* by preventing him from obtaining certain types of contractual terms, which would tend to raise insurance prices and lower insureds' expected utility.

⁷⁹Again, lack of information by contracting parties, here by insureds about the nature of insurance contracts, may also justify intervention in contracts. That may be so if requirements to disclose clearly the nature of terms in contracts and publicly provided information do not function well.

⁸⁰Indeed, in many jurisdictions, liability insurers have a duty in settlement negotiations not to act in their own narrow interests but effectively to act as if they were covering the entire liability, not just an amount up to the coverage limit of the policy. See Sykes 1994.

Chapter 19

GENERAL TOPICS ON THE LEGAL PROCESS

By the *legal process*, I refer to the set of rules governing the actual exercise of legal rights of parties and the manner in which opposing parties defend themselves. Thus, the legal process comprehends the rules of procedure, evidence, lawyer conduct, and other factors bearing on the application and adjudication of legal rights. The legal process is often contrasted to the *substantive law* that lays out the underlying legal rights of parties (such as a rule of tort liability giving a victim of harm the right to collect from an injurer) but that does not state how these rights are to be exercised.⁸¹ In this chapter, we will address several important general issues concerning the legal process and certain related aspects of the substantive law.

1. Public versus Private Legal Systems

1.1 Introduction. A basic question about the design of the legal system is whether the state-authorized legal system should be the sole system of substantive law and of the legal process or whether, and in what circumstances, private legal systems should be allowed a role.⁸²

In fact, we observe that the public legal system provides the default system of substantive rules and adjudication but that private systems, notably arbitration and trade association regimes, have substantial importance.

1.2 Socially optimal legal system. To analyze the desirable relationship between public and private legal systems, it will be helpful to have in mind the socially optimal legal system, that is, the system with the socially optimal substantive legal rules and also the socially optimal system of adjudication. The socially optimal legal system will also have to be enforced to be effective, of course, and that may require use of state power to collect funds, to transfer property, and the like.

Of importance for us is that this optimal system -- both the substantive legal rules and the method of adjudication -- will vary with the characteristics of the involved parties. For example, the best set of legal rules and adjudication for breach of contract may be different for businesses in some industry from what they are for other types of parties.

1.3 Assumptions about public and private legal systems. Let us make the following simplifying assumptions about the public and private legal systems. First, the state's goal, as reflected in the public legal system, is to maximize social welfare. Second, the state can employ its various powers to enforce the law. Third, however, the state may suffer from lack of

⁸¹The definition of the substantive law, as opposed to the legal process, is not always clear. For example, would we say that joint and several liability -- allowing a plaintiff to collect from any of the defendants when more than one injured him -- is an element of the substantive law or of the legal process? The question of rigorous definition of the two notions need not detain us, however, for what is of interest, of course, is the analysis of the rules of the legal system, not the headings under which its rules are placed.

⁸²One of the first economically-oriented articles on this subject is Landes and Posner 1975. This section builds on their article and on Shavell 1995a, which emphasizes the distinction that will be made here between ex ante and ex post choice of private judicial systems. For economic literature focusing on private arbitration (but not on it as an alternative to public adjudication), see, for example, Ashenfelter 1998, Ashenfelter and Bloom 1984, and Benson 2000.

information about the best legal system, either about the best rules or the best method of adjudication, for certain groups of parties. Fourth, private parties themselves will have at least as good -- and sometimes better -- information about what legal system is best for them than the state will have. And fifth, private legal systems will typically have less power to enforce their decisions than the state enjoys, because they lack the power to appropriate funds and to imprison, but they may still sometimes possess methods of enforcement, such as expulsion from a group.

1.4 Public legal system as the sole system. If the public system is the sole legal system, the outcome will deviate from the optimal because of the state's lack of information of what is best for certain groups of parties. Such deviations from what is best suggests that if the state can harness parties' superior knowledge about the best legal system for them and allow the public system to be appropriately modified, the altered system will be superior to the public system alone. We now turn to consider the conditions under which the public system would and would not be expected to be changed in a beneficial way.

1.5. Choice of private system by a single party is undesirable. We can first dispense with the possibility that unilateral choice of a private legal system by a party to a dispute would tend to be socially desirable. It would not, for the party would select a system that favored him, reducing or eliminating the law of its ability to channel behavior or to remedy loss desirably. For example, a rule of tort law requiring injurers to pay for harm, which could be beneficial due to the incentives it provides to take care, would be robbed of force if defendants could select their own legal system. They would choose a different rule that allowed them to escape responsibility, or if they were permitted only to elect the method of adjudication, they would select a tribunal that watered down the ability of plaintiffs to collect. Converse problems, involving excessive liability, would arise if plaintiffs could unilaterally choose liability rules or methods of adjudication. It is plain, therefore, that permitting unilateral modification of the public system of law is socially undesirable.

1.6 Choice of private system by affected parties is desirable if made ex ante. By contrast, it will be socially desirable to allow modification of the public legal system in many circumstances where the decision is made by the parties who are affected by legal system. An important example is where the parties to a contract stipulate that they want a private system to govern contractual problems that may arise. The reasons that allowing them to choose a private system is socially desirable are twofold. First, if each of the contracting parties agrees to the private system, it must make each of them better off. Second, no one else will be made worse off, presuming that the parties to the contract are the only people affected by it.

This argument applies more generally than just to contracts -- it applies whenever all the parties who are affected by some type of behavior make an ex ante agreement about a private legal system. Suppose that the trucks owned by a company present the risk of accidents to citizens of a community, and the company and the community agree to their own legal system to govern accidents. Perhaps they choose strict liability with agreed-upon formulas for damages, rather than the negligence rule and the complicated and costly methods of adjudication that we generally employ, in order to give truckers good incentives to take care in many dimensions and also to streamline dispute resolution. Here, as in the contracting context, the use of the private legal system will be socially beneficial. For if the parties agree to a private system as an alternative to the public system, they must be made better off, and others are not made worse off, as by hypothesis the group in question includes all potentially affected parties.

Two comments may be made about the foregoing. First, the private system that the parties elect may differ from the public system in its substantive legal rules (such as employing strict liability rather than negligence for truck accidents) or in its methods of dispute resolution (such as in the manner of determining damages). Second, the private system may or may not use its own system of enforcement. Typically, private parties will not have a good way to enforce the decisions of a private system, such as of arbitrators, and will wish to rely on the state to accomplish this. Thus, it will be desirable for the state to lend its enforcement apparatus to the private system; otherwise the private system would be rendered ineffective.

1.7 Choice of private system by affected parties may or may not be desirable if made ex post. The parties to a dispute may mutually desire to choose a private system of adjudication over the public system. Notably, they may want to do this to lower the costs of dispute resolution and to reduce risk. However, allowing them to choose a private system does not necessarily raise social welfare.⁸³ Consider an accident between an individual and a firm. They may decide that it is in their joint interests to elect arbitration for its simplicity and speed, but that may mean that the firm escapes with inadequate liability or that the firm's fault is never properly investigated and made known to the public. And if firms anticipate often being able to reach such agreements to arbitrate, they may not be properly deterred. In other words, because the ex post incentives of parties to use private adjudication are naturally divorced from considerations of deterrence, their use of private systems may be undesirable.⁸⁴ By contrast, when parties make agreements ex ante, and all potentially affected parties are involved in the agreements, deterrence is not overlooked; in the case of truck accidents in a community, if the citizens make an agreement to use a private legal system, it can only be because the system does not result in too much dilution of deterrence.

Of course, ex post agreements to use private legal systems may also be socially desirable, because their potentially undesirable effect on deterrence is outweighed by their value in

⁸³The argument to be made is in some respects similar to the point made in section 2.6 of chapter 17 about the possibility that settlement may be socially undesirable; here the decision to use a private legal system (often because it is less costly than the public one) is analogous to the decision to settle.

⁸⁴An example may clarify this point. Suppose that if firms exercise care at a cost of \$10, accidents causing harm of \$15 will be prevented, so that the exercise of care is desirable. Suppose too that liability is strict and that the probability that a firm will be sued is only 50 percent -- there is difficulty in proving causation; that the costs of use of the public legal system would be \$6 each for the plaintiff and the defendant, and only \$2 each for use of a streamlined private arbitration system. Then, if an accident occurs for which causation is clear (and for which there is thus a threat of suit under the public legal system), the parties would elect the private system, in order to save \$4 each in adjudication costs. This, however, means that the expected liability of the firm if it does not take care will be $50\% \text{H}(\$15 + \$2) = \$8.50$, so it will not be led to spend \$10 on care. But if the public system were employed, the firm's expected liability if it does not take care would be $50\% \text{H}\$21 = \10.50 , so it would be led to take care (in which case, note, there would never be any litigation expenses, as no accidents would occur). Thus, social welfare, and the well-being of the victims of accidents, are lowered because of the ability of the victims and the firms to make ex post agreements that reduce firms' incentives to take care.

The reader might wonder whether the point of the example depends on the fact that not all of the affected parties participate in the decision to employ the private system of adjudication -- for the 50 percent of victims who are not able to prove causation are left out of negotiations. However, this feature of the example is not intrinsic to it. For instance, one could alter the example by assuming that all individuals sue but that a fraction of losses cannot be proved due to difficulties in establishing causation, and the same point could be made even though all victims would be participating in the decision to use a private legal system.

reducing the costs of dispute resolution and perhaps in increasing its accuracy. Therefore, it is possible, if not plausible, that ex post agreements to employ private legal systems are often socially desirable.

An observation that should be made about ex post agreements to use private legal systems is that such agreements will generally be limited to the method of dispute resolution and will not alter the basic thrust of the legal rules of the state, for that would be against the interests of one or the other of the parties. For example, if the substantive legal rule is the negligence rule, then a non-negligent injurer would be unlikely to agree to a private system in which liability would be strict. However, were victims and injurers to choose a private system ex ante, they might opt for strict liability.

1.8 The notion that selection of legal rules and adjudication are natural functions of the state. The foregoing analysis helps to answer the question of whether the choice of substantive legal rules and the system of adjudication are natural functions of the state, as some seem to think is so.⁸⁵ The answer, we have seen, depends importantly on whether the parties that are affected by some type of behavior are likely to bargain with one another ex ante about the choice of a private legal regime. When they are likely to do so, and when they have superior information to the state about the legal rules or manner of adjudication that best serves their purposes, it is best for them to select it. In this circumstance, the idea that the state is naturally the entity that should choose the legal system is incorrect (although the state may still play a role in enforcement).

When affected parties cannot bargain with each other ex ante, which is the case when they are strangers to one another, the state will play a natural role in the selection of substantive legal rules. Thus, in the usual tort context, for instance, the state must make the substantive legal rules in order for there to be any rules that will influence parties' behavior. However, there still may be scope for private parties to beneficially elect a private system of adjudication after harm is done, as discussed in section 1.7.

1.9 Actual practice in the light of the theory. Actual practice conforms broadly to what has been described above as socially desirable. In particular, we observe that private parties often opt out of the state-authorized legal system, as noted at the outset. There is significant use of arbitration and of other private means of adjudication, such as those of trade associations, employers, and religious organizations.⁸⁶ Also, as stated above, the private systems that are chosen ex ante often involve procedural and substantive legal rules that depart significantly from those of the public system.⁸⁷ However, there does not appear to be a tendency in practice to distinguish between ex ante agreements to use private systems and ex post agreements, even though the case in theory for respecting agreements made ex post is less strong.

One point to note is that, at least in the United States today, there is a tendency of commentators, and of the state itself, to encourage use of alternatives to the public system of

⁸⁵See, for example, Couture 1999.

⁸⁶See for example Benson 2000, 159-62, Bernstein 1998, Cooley and Lubet 1997, 20-23, and Goldberg, Sander, and Rogers 1999, 10-11.

⁸⁷See, for example, Bernstein 1998, Goldberg, Sander, and Rogers 1999, 233-34, and Ware 2001, 21, 80-86, and for case studies, Bernstein 1992, 2001.

adjudication, and even to subsidize it or require private parties to employ it. The only apparent justification for encouragement or subsidy of private alternatives to the public system is that parties do not pay legal fees that cover the actual costs of use of public courts, so that the public system may appear cheaper relative to private systems than it should; fostering the use of private systems might correct this distortion.⁸⁸

2. Accuracy of the Legal Process

2.1 Introduction. By the accuracy of the legal process is meant the absence of error. Error may arise in the determination of whether or not a person is liable -- an innocent person may mistakenly be found liable or a guilty person not found liable -- or in the determination of the level of sanctions to be imposed on a liable person.

The accuracy of the legal process is influenced by its design; indeed, the entirety of the rules of legal procedure and of evidence bear on the incidence of error. And the accuracy of the legal process is also affected by the actions of litigants; their gathering of information, selection of evidence, and use of legal procedure influence the likelihood of error.

Here I will discuss first the social value of accuracy of the legal process: the increase in social welfare that greater accuracy brings about. The social value of accuracy inheres in its effects on outcomes, and notably in its effects on the behavior of parties who anticipate greater accuracy of the legal process; accuracy is not valued in itself. As will become evident, the social value of accuracy depends in significant ways on context.

Given the social value of accuracy, the socially optimal level of accuracy can be determined. As a general matter, increasing accuracy is socially costly, as it requires a lengthier and higher quality legal process. Therefore, the level of accuracy that maximizes social welfare will reflect a compromise between the value of increasing accuracy and the cost of achieving it.

I will then turn to examine the private value of accuracy, for example, the value to a plaintiff of establishing that the defendant is the party who harmed him. It will be seen that the private value of accuracy generally diverges from the social value, calling in principle for the legal process to include features that remedy the divergence.⁸⁹

2.2 Social value of accuracy. The social value of accuracy can usefully be divided into three components.

Improved control of behavior. It is intuitively clear that greater accuracy of the legal process should tend to bring about improved control of behavior, but why exactly should this be true?

Consider first accuracy in the determination of liability. Greater accuracy in imposition of liability on truly guilty parties clearly leads to an enhanced incentive to obey the law, for it means that the expected sanction for violations is higher. Not so obvious, however, is that greater accuracy in respect to exonerating innocent parties enhances deterrence. The reason is that the incentive to obey the law is not simply equal to the expected sanction if one violates the law, but

⁸⁸But a better answer to this distortion in relative prices would seem to be to price the public system appropriately, not to subsidize the private system, for that may lead to excessive litigation. Yet this whole issue is complicated by the concern, discussed at length in chapter 17, that the private incentive to use the legal system differs from the social, so that pricing legal services at cost is not generally optimal in the first place.

⁸⁹This section is largely a synthesis of Kaplow and Shavell 1994a, 1996a, and Kaplow 1994.

rather to the *difference* between the expected sanction if one violates the law and the expected sanction if one does not. If the expected sanction suffered by the innocent, due to error, is 20, and the expected sanction experienced by the guilty is 60, then the effective sanction for a violation is 40, not 60, for 40 is the added sanction brought about by a violation. Hence, the incentive to obey the law is enhanced by reducing errors that penalize the innocent (if, for instance, the expected sanction suffered by the innocent fell from 20 to 10, the effective sanction for violations would rise from 40 to 50).⁹⁰

An additional social benefit from increased accuracy in the determination of liability is improved decisions about whether to engage in an activity (such as to operate a motor vehicle) that would expose innocent parties to the risk of liability. Greater accuracy implies that fewer parties will be undesirably discouraged from engaging in such an activity because of mistakenly imposed liability. It also means that fewer parties will be undesirably encouraged to engage in the activity because of mistakenly escaping liability.

Another, more particular social benefit from increased accuracy applies when liability is based on whether or not a party's level of care falls below a fault standard. In these situations, as was discussed in section 1.1 of chapter 10, error in the assessment of care may lead to the exercise of excessive care (such as defensive medical practices) to reduce the chance of mistakenly being found liable. Here greater accuracy in the assessment of care will lead to less excess in precautions.⁹¹

Next, consider accuracy in the assessment of damages. Here accuracy tends to improve behavior. The reason is that, if an actor knows that the harm he might cause will be accurately assessed, he will tend to take steps to prevent harm commensurate with its magnitude; he will appropriately do more to prevent harm the larger its magnitude. However, it is important to note that the actor can only take these harm-appropriate steps if he knows what the harm will be at the time he decides on precautions. And in many contexts, actors will have relatively little knowledge about the particular harm they would cause in an accident when they exercise precautions. For instance, a driver will typically have relatively little knowledge of the harm he would cause in a collision (the harm would depend on the speed of the other vehicle, where it was struck, the number of occupants in it, and so forth). This possibility reduces the social value of accuracy in the assessment of damages.⁹²

Reduced social costs from litigation and from imposition of sanctions. Another social advantage of accuracy is a higher frequency of settlement, and thus of savings in litigation costs.

⁹⁰The point under discussion may be stated algebraically. Let s_I be the expected sanction if a person is innocent, obeys the law, and let s_G be the expected sanction if a person is guilty and violates the law. Then if b is the benefit from violating the law, a person will violate the law if $b > s_G - s_I$, that is, if $b > s_G - s_I$, so that the incentive to obey the law is the difference between the expected sanctions. This point was initially emphasized in Png 1986.

⁹¹There is no conflict between this point and the point above that greater accuracy increases deterrence, for the presumed contexts are different. It was assumed above that a party chooses between two actions, obeying the law or not, whereas here the assumption is that a party chooses a level of care that is continuously variable (or has many levels), and it is possible for the party to choose a higher level than is called for and thereby to lower the chance of erroneous imposition of liability.

⁹²See Kaplow and Shavell 1996a, but see also Spier 1994b for a qualification to this point: that there is social value in accuracy in assessing damages even when an actor does not know the magnitude of harm in advance, if the level of his precautions lowers the probability distribution of harm given that an accident occurs.

Specifically, if greater accuracy of the legal process means that litigants are more likely to agree on their estimates of trial outcomes, settlement will be promoted. Additionally, accuracy may reduce the need to impose sanctions and thus the costs of so doing.⁹³ Greater accuracy leads to reduced imposition of sanctions when the accuracy results in better identification of the innocent. (Of course, greater accuracy in identification of the guilty results in more frequent imposition of sanctions, and thus greater costs from imposing sanctions.)

Lowered costs of risk-bearing. Greater accuracy in outcomes tends to lower risk, for it means that outcomes as to whether parties are liable and the amount that liable parties pay are more predictable. The reduced risk is socially valuable in itself, to the extent that parties are risk-averse and uninsured against the risks in question.

2.3 Socially optimal degree of accuracy. As indicated above, the socially optimal level of accuracy will depend on a tradeoff between its social value and its cost. For example, where potential injurers have relatively little knowledge of the precise magnitude of harm when they choose their levels of care (plausibly the case with drivers, as suggested above), the social value of accuracy in assessing harm will be low. Thus, the best policy might call for little to be spent on accuracy, meaning that streamlined methods (such as simple tables) to ascertain damages might be best. The optimal degree of accuracy will depend generally on context, as the social value and the costs of accuracy will depend on the particulars of the area of behavior at issue.

2.4 Traditional view of accuracy versus the economic view. The traditional view of legal scholars about accuracy has several features. One is that accuracy is of intrinsic value; this is inconsistent with the economic view. A second strand of traditional thinking is that accuracy is necessary in order to maintain the legitimacy of the legal process. The economic view is not inconsistent with this point (although I have not mentioned it above), for if individuals respect the legal system and cooperate in its application, it will work more effectively to further social ends. A third element of traditional writing on accuracy is that it serves instrumental purposes. This, of course, is entirely consistent with the economic view, but the instrumental purposes of accuracy are rarely analyzed in a sustained way by traditional scholars, whereas these purposes are the focus of economic analysis.

2.5 Private versus social value of accuracy, and implications. Let us now turn to the value of accuracy to private parties themselves, as opposed to its social value.

Perhaps the first point to note is simply that accuracy may not be desired -- may have negative value -- to a private party. A plaintiff may want to conceal facts in order to prevail against a defendant who is in fact innocent or may want to exaggerate his losses; a defendant may have similarly perverse incentives to prevent the truth from becoming known.

Of course, private parties may also have incentives to prove the truth. A plaintiff will want to show that he is really the victim of harm and that the named defendant is the liable party, or a defendant will want to establish that the plaintiff was truly contributorily negligent, and the like. However, the incentives of private parties to establish the truth will tend to diverge from the socially optimal, and in either direction. The private reason to spend in order to establish a fact is, for the plaintiff, to increase his expected judgment, and for the defendant, to reduce it. From the social perspective, however, the justification for expenditures lies in improved control of

⁹³These costs are the various administrative costs associated with settlements and judgments, which were noted to be large (see section 1.3 of chapter 17). In the criminal context, which will be considered later in the book, the costs of imposing sanctions also include the expenses of incarceration.

behavior and the other factors mentioned in section 2.3; these social benefits may be quite different from the private ones. For example, it may be that increased accuracy in determining liability would be socially very valuable in deterring negligent behavior, but that the private value of establishing negligence is too low to induce a plaintiff to do so.⁹⁴ Or it might be that increased accuracy in ascertaining liability has small social value, because there is little that potential injurers can do to reduce risk, yet the private value of establishing liability may be quite high, because of the damages that the plaintiff can collect, so that more would be spent establishing accuracy than is socially desirable.⁹⁵ In general, the reasons for the private-social divergence in the value of accuracy are analogous to those discussed in chapter 17 on the private-social divergence in the value of bringing suit. Indeed, a factor stressed there, that each side generally fails to take into account the litigation costs that it induces the other side to incur, is also relevant in the present context.

An additional point relating to the private incentive to establish the magnitude of harm should be recalled (see section 5.3 of chapter 17). Namely, there is a systematic tendency for private parties to value accuracy in determining the magnitude of harm above its social value. This is suggested by the observation that a plaintiff will be willing to spend up to \$100 to prove that his losses are in fact \$100 higher (or a defendant to spend up to \$100 to prove that the losses he caused are in fact \$100 lower), but the social value of establishing that losses are \$100 higher is generally less than \$100.

What are the implications of the divergences between the social and the private values of accuracy? When private parties have incentives to withhold information, the state should attempt to prevent that if the information's social value outweighs its cost. Of course, much of the law of procedure is designed to address this problem. When private parties have incentives to supply information rather than to conceal it, the state should encourage its provision if parties' incentives are inadequate, but limit its provision if parties' incentives are excessive. The latter may be accomplished through the use of procedural rules that, for instance, restrict the number of witnesses or that simplify the calculation of damages. A conjecture is that the problem of excessive private incentives to establish facts is considerable -- especially in respect to the proof of the magnitude of damages -- and that it is not properly appreciated.

3. Appeals and the Legal System

⁹⁴Suppose that defendants can reduce the risk of harm of \$1,000 from 100 percent to 1 percent by exercising a precaution at a cost of \$100, and that proving negligence -- that the precaution was not taken -- would cost a plaintiff \$2,000. Then a harmed plaintiff would not spend the \$2,000 to prove negligence, and since defendants would anticipate this, they would not take the \$100 precaution; thus, social costs would be \$1,000. However, it would be socially desirable that plaintiffs spend \$2,000 to prove negligence whenever accidents occur, for then defendants would expect that negligence would be found out and would be led to spend the \$100. As a consequence, if, say, plaintiffs were given free legal services costing \$2,000 to investigate possible defendant negligence, social costs would be only $\$100 + 1\% \text{ H } \$2,000 = \$120$ (and thus lower than \$1,000).

⁹⁵Suppose that liability is strict, that losses of \$1,000 occur with probability 10 percent, that nothing at all can be done to prevent the losses, and that plaintiffs must spend \$100 to prove causation, and thus to establish liability accurately, if a loss occurs. Clearly plaintiffs will do that, as \$100 is less than the \$1,000 they will collect; yet socially that is undesirable because social costs would be only $10\% \text{ H } \$1,000 = \100 if suit is not brought, but social costs are $10\% \text{ H } \$1,100 = \110 if suits are brought (actually, they would be higher if we take into account the defendant's litigation costs and the state's).

3.1 Introduction. An important feature of our legal system is the inclusion of an appeals process, whereby a disappointed litigant can make a request for reconsideration of the initial trial result.⁹⁶ It should be emphasized, however, that a legal system might not include an appeals process -- whether it does so is an element of the design of the system -- and in some contexts there is no appeals process.⁹⁷ Notably, when parties elect binding arbitration, they are usually choosing to forgo an appeals process.⁹⁸

Here I will ask why the appeals process may be socially desirable. In answering the question, one must explain, among other things, why society (or an organization) may find the appeals process superior to the alternative of enhancing the quality of the trial process. Society enjoys the option, after all, of investing in more skilled trial court judges, of lengthening trial proceedings to allow for more evidence and argument to be considered, and the like. Moreover, one must say why, if society does decide to employ a tribunal that supersedes the trial courts, it should wish to grant disappointed litigants the right to instigate action by the higher tribunal -- rather than, say, permit the higher tribunal to reconsider trial outcomes on its own initiative.

The theme to be advanced here is that the appeals process may function to correct errors in an economic way (other functions of appeals will be noted in sections 3.3 and 3.4). The kernel of the argument to be developed is that if litigants tend to possess information about the occurrence of error and appeals courts can frequently verify it, then litigants may tend to bring appeals when errors are likely to have been made but not otherwise. Under these circumstances, not only may the appeals process result in error correction, it may also do so cheaply, for the legal system will be burdened with reconsidering only the subset of cases in which errors were more probably made. This may render society's investment in the appeals process economical in comparison to an investment in improving the accuracy of the trial process -- an approach that, by its nature, would require extra expenditure in every case rather than only in a subset of cases. The appeals process, in other words, may allow society to harness information that litigants have about erroneous decisions and thereby to reduce the incidence of mistake at relatively low cost.⁹⁹

3.2 Basic model: the appeals process and error correction. Suppose in this section that litigants are able to recognize when error occurs at trial, that the state chooses whether or not to establish an appeals court, and that the state can increase the accuracy of adjudication at both the trial court level and the appeals court level by making appropriate expenditures. If there is an appeals court, then, by definition, a disappointed litigant may bring an appeal if he chooses. To bring an appeal, the litigant must pay his private legal costs. In addition, the litigant may have to pay a fee, or possibly will receive a subsidy. Thus, the total cost to a litigant for bringing an

⁹⁶That all developed legal systems generally recognize some form of appeal process is discussed, for example, in chapter 8 of Herzog and Delmar 1982 and Platto 1992.

⁹⁷A number of legal systems, for example, those of France and Italy, do not allow appeal of cases for which the amount at stake is below a threshold; see Byrd and Barbier 1992, 160, and Beltramo 1995, 470.

⁹⁸Decisions of arbitrators generally may not be appealed to the legal system; see, for example, Goldberg, Sander, and Rogers 1999, 235, and Ware 2001, 21. Moreover, the arbitration process itself ordinarily does not allow for appeal to other arbitrator; see, for example, the website of the American Arbitration Association at www.adr.org.

⁹⁹The analysis that follows is based on Shavell 1995b. For studies of the appeal process emphasizing factors other than error correction, see Daughety and Reinganum 2000 and Spitzer and Talley 2000.

appeal is his private cost plus a potential fee or minus a possible subsidy.

The state's objective is assumed to be minimization of total social costs: the sum of the social costs of adjudication -- the costs of trial together with the expected costs of appeal, if there is an appeals process -- plus the social harm from erroneous decisions.

Consider first the best that the state can do if it does not allow appeals. In this situation, the state's problem is simply to select the level of trial court accuracy so as to minimize trial costs plus expected harm from trial court error. The optimal level of accuracy will be dictated by the effectiveness of legal resources in promoting accuracy and by the magnitude of harm from error.¹⁰⁰

Now suppose that the state establishes an appeals court and, provisionally, take its accuracy as given, but assume that it is minimally accurate in the sense that it is more likely to reverse trial court errors than to reverse correct decisions. Hence, a disappointed litigant's expected gross return from an appeal will be higher if an error occurred than if it did not, because the reversal probability is greater under the former circumstance than under the latter. For example, if the reversal probability following a mistake is 80 percent but is only 30 percent following a correct decision, then if a litigant's gain from reversal would be \$100,000, the expected gross return from an appeal would be \$80,000 after a mistake but only \$30,000 after a correct adverse decision.

Accordingly, there may be *separation* of disappointed litigants, wherein those who are the victims of error find it worthwhile to bring appeals and those who are not victims of error do not find it worthwhile to bring appeals. There will be such separation of disappointed litigants if the private cost of an appeal is less than the expected return given mistakes but exceeds the expected return given correct decisions. In the example just mentioned, there will be separation if the cost of an appeal is, say, \$50,000; \$50,000 is less than \$80,000, meaning that there will be appeals after mistakes, but \$50,000 exceeds \$30,000, meaning that there will not be appeals after correct adverse decisions.

If separation of disappointed litigants would not occur naturally, due to whatever happens to be the private costs of appeal, the state can ensure that separation occurs by selecting the right fee or subsidy. If appeals would be made even after correct decisions, because the private cost of an appeal is lower than the expected return, the state can impose a fee in order to achieve separation.¹⁰¹ Conversely, if appeals would not be made even after mistakes, because the private cost of an appeal is higher than the expected return from appeal, the state can grant a subsidy to induce separation.¹⁰²

¹⁰⁰Let x be the investment in trial court accuracy, $p(x)$ the probability of trial court error, where p is decreasing and convex in x , and h the social cost of error. Then, in the absence of an appeals process, the state should choose x to minimize $x + p(x)h$; let the optimal x be x^* . (This notation will be used in later notes without further comment). Observe that, in a formal sense, the state's problem is identical to that of choosing the optimal level of precautions in the model of accidents considered in chapter 8; the level of precautions to prevent accidents in that model corresponds here to the level of investment in the accuracy of trial courts to prevent errors.

¹⁰¹Appeals would be made in the example even after correct decisions if the private cost of an appeal were \$10,000, for this is less than the \$30,000 expected return from an appeal. Hence, if a fee of, for instance, \$40,000 were imposed, the total cost of making an appeal would become \$50,000, and appeals after correct decisions would be discouraged (but appeals after mistakes would still be brought, as the return from them is \$80,000).

¹⁰²In the example, appeals would not be made after mistakes if the private cost of an appeal were \$100,000

Moreover, it is clear that the state would want to achieve separation of disappointed litigants when it has an appeals process. On one hand, if litigants never bring appeals, the appeals process can hardly achieve good. On the other, if litigants bring appeals even when correct decisions are made, society incurs needless costs in the appeals process and also, to its detriment, finds that a certain number of these correct decisions are reversed.

Knowing that separation of disappointed litigants is socially desirable and that the state can, if need be, accomplish separation through an appropriate fee or subsidy scheme, we can easily determine whether the appeals process will be socially helpful. If the appeals process is not used, then when an error is made, the associated social harm will definitely be suffered. But if the appeals process is utilized, an error will result in an appeal, so that the social costs incurred will instead be those of the appeal and the expected harm due only to the possible failure to reverse error. It follows that *the appeals process will be socially desirable to establish if and only if the social harm from certain error exceeds the social cost of an appeal plus the expected harm from failure to reverse error*, the latter being the probability of failing to reverse error multiplied by the harm from error.¹⁰³ If the social harm from error is \$500,000, the social costs of an appeal are \$150,000, and the probability of reversal of error is 80 percent, then the appeals process will be advantageous because it will reduce the certain harm from error of \$500,000 to \$150,000 + 20% H \$500,000 or \$250,000. In general, therefore, the appeals process is more likely to be socially desirable the lower the cost of the appeals process, the greater the chance of reversing error, and the greater the social harm from error. In particular, and other things being equal, the appeals process will be desirable if the social harm exceeds a certain threshold and will not be desirable if the harm lies below this threshold.

To this point, we have taken the cost and accuracy of the appeals process as given, as well as those of the trial process, but as the reader knows, these are variable. What can be said about their optimal choice? With regard to the appeals process, it is socially desirable to invest in accuracy as long as the increase in costs is outweighed by the increase in the expected gain, that is, the increase in the probability of reversal of error multiplied by the social harm from error.

With regard to the trial process, it is socially advantageous to invest in accuracy as long as the increase in costs is outweighed by the increase in the expected gain from a lower probability of error. But the social harm from error at trial, it should be emphasized, is *less* than the harm flowing from a sure error. The social harm from error at trial is instead measured by what follows trial court error, namely, the cost of the appeals process plus the expected harm from failure to reverse error; this amount is lower than the sure harm from error (in the example

because this exceeds \$80,000. If, however, a subsidy of, say, \$50,000 were employed, appeals would be made after mistakes (but not after correct decisions).

More generally, to show that separation of disappointed litigants is always possible, let g be the gain to a litigant from winning an appeal, a be the litigant's private cost of making an appeal, and b be the fee (a positive b) or subsidy (a negative b). Also, let q be the probability of reversal of an error and r the probability of reversal of a correct decision, and assume that $q > r$. Then there will be separation of disappointed litigants if $rg < a + b < qg$. It is clearly always possible to find a b satisfying that inequality because $q > r$.

¹⁰³ To amplify, given that the state can ensure that disappointed litigants sue if and only if an error was made, social costs under an appeals process will be $x + p(x)[y + (1 - q(y))h]$, where y is the state's expenditure on the appeals process. If there is no appeals process, social costs will be $x + p(x)h$. Hence, it will be desirable to append an appeals process to a trial process if and only if $y + (1 - q(y))h < h$, which corresponds to the italicized statement in the text.

of the appeals process above, the cost of that process plus the expected harm from failure to reverse was \$250,000, much less than the \$500,000 harm from error in the absence of the appeals process). Hence, *the optimal investment in, and accuracy of, the trial process is less than it would be if there were no appeals process and no opportunity to correct errors.*¹⁰⁴ In other words, we have argued that it is optimal for trial courts to be less accurate on account of the existence of appeals courts than the trial courts would otherwise need to be.

3.3 Qualifications and extensions to the basic error correction model. Let me briefly consider a number of qualifications to, and modifications of, the foregoing model.

Factors limiting error correction under the appeals process. Two important factors reduce the ability of the appeals process to correct errors as described in the model. One is that the degree of separation will in reality be less than perfect, and the appeals process will thus function less well, to the degree that separation of disappointed litigants does not naturally occur and the state does not employ fees or subsidies to accomplish separation. Another limiting factor is that litigants may not be able to tell when errors were made at trial, so that the likelihood that errors result in appeals is lowered, and the likelihood that correct decisions are appealed is raised.

Multiple levels of appeal. The basic model can be extended in a straightforward manner, whereby the state chooses the number of levels of appeal and the resources and accuracy of the process at each level. The conclusions about this extended model are natural generalizations of those for the basic model. In particular, given any number of levels of appeal, an additional level of appeal will be desirable if, for some investment in the accuracy of the next level, the cost of that level of appeal plus the expected harm from failure to reverse error at that level is less than the certain harm that would be incurred if that level appeal were not allowed. Additionally, under certain conditions, the optimal level of investment in, and the accuracy of, appeals courts increases with their level. This reflects the point that the higher the level of appeal, the fewer the number of opportunities that remain to correct error, so the more valuable is accuracy.

Judges' incentives to avoid reversal. In the basic model, the appeals process increased accuracy through correcting errors in trial court decisions, but the appeals process may also increase the accuracy of trial courts by influencing the behavior of judges who dislike being reversed (because reversal may harm their reputation, lower their salaries, and the like). In particular, judges who fear reversal will have a greater incentive to avoid error the more likely erroneous decisions are to be appealed and reversed than are correct decisions. (Note, therefore, that if the errors were no more likely to be reversed than correct decisions, judges would have no incentive (on these grounds) to decide cases accurately; hence improving judges' incentives cannot alone justify the appeals process but is instead a byproduct of the error correction function of the appeals process.)

Inference from the fact that an appeal is brought. In the basic model, because litigants bring appeals if and only if errors are made at trial, an appeals court can infer that everyone who

¹⁰⁴To demonstrate this, denote the x that minimizes $x + p(x)z$ by $x^*(z)$ and observe that $x^*(z)$ is increasing in z (this is readily shown to follow, given convexity of p). Let y^* be the optimal investment in the appeals process, so that social costs under the appeals process are $x + p(x)[y^* + (1 - q(y^*))h]$. Hence, the optimal x given that there is an appeals process, denoted by x^{**} , must equal $x^*(y^* + (1 - q(y^*))h)$. But the optimal x without the appeals process, x^* , equals $x^*(h)$ (see note 20). And because $x^*(z)$ is increasing in z , and $y^* + (1 - q(y^*))h < h$ (this must be true, by note 23 for the appeals process to be desirable to employ), it thus follows that $x^{**} < x^*$.

comes before it ought to obtain a reversal. But I implicitly assumed that appeals courts do not employ this knowledge; rather, they use their usual rules of decision, and they thus fail to reverse with some positive frequency. If appeals courts were to reverse all decisions, on the basis of their inference that all appellants are the victims of error, then disappointed litigants who are not the victims of error clearly would have an incentive to bring appeals, for they could obtain sure reversals. Thus, the separation of disappointed litigants would unravel, and the utility of the appeals process in error correction would be diminished. This suggests that it is not socially desirable for appeals courts to use inferences from the fact that appeals are brought in their decisionmaking. Appeals courts can (and apparently do) refrain from using inferential knowledge by following certain legal procedures, notably, by restricting evidence considered on appeal to the trial record.

3.4 Functions of the appeals process other than error correction: harmonization of the law; and law-making. Two purposes of the appeals process apart from error correction that are frequently emphasized are harmonization of the law -- reconciliation of conflicting interpretations of the law by different trial courts -- and law-making -- amplification of the law through new interpretations. The main observation to be made about these two functions is that, although they in fact are carried out by appeals courts, the appeals process does not seem necessary to accomplish them. A higher level court could readily accomplish harmonization on its own initiative, for conflicts among trial courts are self-evident in nature, meaning that there is no reason to give disappointed litigants the right of appeal in order to have conflicts brought to the attention of a higher court. Similarly, it would seem that the need for amplification of the law is generally fairly clear to higher level courts, so that these courts would not require that litigants bring cases to their attention in order to know where amplification is in the social interest; higher level courts could by themselves choose where to amplify the law and do that in declaratory rulings.

4. Legal Advice and the Role of Lawyers in the Legal System

4.1 Introduction. A basic aspect of the legal system is that lawyers provide clients with advice. The advice may be about the nature of legal rules, about the probability and magnitude of sanctions for violations, and about litigation and legal procedure.¹⁰⁵ I have already mentioned some specific aspects of legal advice, concerning the bringing of suit and settlement decisions (see section 7 in chapter 18), and here I consider the topic more generally.¹⁰⁶

A client may obtain legal advice *ex ante* -- when he is contemplating an action -- or he may secure it *ex post* -- after he has acted or someone has been harmed, which is to say, at the stage of possible or actual litigation. I will consider these two types of advice separately because of their distinctive aspects. A notable difference between these types of advice is that *ex ante* advice can channel behavior directly in conformity with law, whereas *ex post* advice comes too late to accomplish that, although, as will be seen, *ex post* advice has indirect effects on

¹⁰⁵The provision of legal advice is not taken to be coextensive with the provision of legal services in general, however. Some legal services (such as the making of arguments in court) are better regarded as the performance of specialized tasks than as the provision of information to clients.

¹⁰⁶Indeed, most of what is said here will apply to the context of public law enforcement as well as civil litigation.

behavior.¹⁰⁷

Several assumptions will be maintained in most of the discussion in this section: that legal advice is not purposely subversive of the law (for instance, that advice is not intended to enable a person to perpetrate a fraud), that lawyer-client communications and legal work product are confidential, and that lawyers are truthful to clients and endeavor to provide them with good advice. However, each of these assumptions will be examined in section 4.4.

4.2 Ex ante advice: advice about contemplated acts. Advice will have private value to a party who is considering taking some action with a possible legal consequence if the advice might lead him to alter his decision. Suppose that a firm is deciding whether to release waste from a holding tank into a river, rather than to spend on transporting the waste to a dump site, but the firm does not know whether a discharge of this particular waste would constitute a violation of law allowing some class of victims to sue for harm. One possibility is that, without advice, the firm would elect to discharge the waste into the river (suppose the firm thinks discharge probably would not be a violation). In such situations, advice would have private value if it might lead the firm instead to transport the waste to the dump site (the advice might be that a discharge would in fact constitute a violation), because advice would then enable the firm to avoid liability. Also, advice might have value to the firm if, in the absence of advice, the firm would decide to transport the waste to the dump site (suppose the firm believes that a discharge probably would constitute a violation). Here advice would have value to the firm if it might lead the firm instead to discharge the waste into the river (the advice might be that a discharge would not constitute a violation), because advice would then save the firm the transport costs. In general, the private value of legal advice is the expected value of the private gain from possible changes in a party's decisions. This notion of the private value of legal advice is, it may be noted, just an application of the conventional definition of the expected value of information to a decisionmaker.¹⁰⁸

The social, as opposed to the private, value of ex ante legal advice inheres in the social desirability, or lack thereof, of advice-induced changes in parties' behavior. Suppose that it is socially desirable for the waste not to be discharged, and thus discharging it would constitute a violation (because the harm from a discharge would exceed the cost of transport to the dump site). Then if advice would result in the firm deciding against discharging the waste, the advice would be socially desirable. But if advice would lead the firm to discharge the waste (say, because the firm would learn that the probability of suit is low), the advice would not be socially desirable. The social value of advice is the expected value of the potential social gains and losses produced by the advice.

The comparison between the social and the private values of legal advice depends, among other factors, on the form of liability: whether it is strict or based on the negligence rule. In the consideration of these rules, it will initially be assumed that suit always occurs and that a sanction equal to harm is imposed whenever parties are supposed to be liable for harm.

When parties are strictly liable, the private value of legal advice is the same as its social

¹⁰⁷ Ex ante legal advice was first studied from an economic theoretical perspective in Shavell 1988 and Kaplow and Shavell 1992; ex post legal advice was initially investigated from this standpoint in Kaplow and Shavell 1989, 1990. Legal advice is further studied from an economic viewpoint in Bundy and Elhauge 1991 and in Fischel 1998.

¹⁰⁸ See, for example, Raiffa 1968.

value. This basic and important conclusion follows essentially because a party's liability burden will equal the harm he causes. If a party learns through advice that taking some precaution will reduce his liability by \$10,000, say from \$15,000 to \$5,000, this also means that the precaution will lower harm by \$10,000. Hence, it should not be surprising that the private and social values of advice are equivalent.

Under the negligence rule, however, the private value of legal advice tends to exceed its social value. The explanation is suggested by two points. First, if a party avoids negligence because of advice, the party's liability saving will generally be larger than the reduction in expected harm he accomplishes. Suppose that, without advice, the party just mentioned would not take the precaution and would be found negligent and be liable for the harm of \$15,000. And suppose that, with advice, the party would take the precaution and thereby avoid liability for negligence. Thus the advice would lead to a reduction in liability for the party of \$15,000 -- an amount exceeding the \$10,000 reduction in harm. The reason that the private liability saving from advice is larger than the reduction in harm is that, under the negligence rule, a party escapes having to pay for any harm caused when he acts nonnegligently (the party escapes having to pay for the \$5,000 of harm he still generates if he takes the precaution). The second point is similar. If a party would learn from advice that he can relax somewhat his level of precautions and will still not be found negligent, his saving will be the full amount of the reduction in precautionary costs. However, society will not save as much as the party, for when precautions decline, expected harm rises.

It should be emphasized that, whatever may be the difference between the private and the social value of ex ante legal advice, such advice does have positive social value because it can only change private decisions in socially desirable ways, to better conform with the law (or with lawyers' understanding of the law, which is assumed to be superior to clients'). This statement presumes that the law is properly enforced and that decisions that conform with the law are socially desirable. However, what if the law is underenforced or is not socially desirable?

When legal rules are underenforced, that is, when the probability of having to pay for harm is less than 100 percent or when the level of damages is less than harm, legal advice might have negative social value because it may lead parties to disobey the law when they otherwise would not. Likewise, when legal rules are incorrectly formulated, legal advice tends to have negative social value. Suppose, for example, that an environmental authority mistakenly omits a truly harmful waste from its list of substances for which discharge would be considered a negligent act. Then legal advice that gives a party foreknowledge of this error might cause the party to discharge the harmful waste, when it would not otherwise have done so.¹⁰⁹ Thus, the social value of advice is negative -- it would be best for parties not to obtain advice -- even though its private value is positive.

4.3 Ex post advice: advice about acts already committed. The private value of ex post legal advice, advice provided after acts have been committed, is analogous to the private value of ex ante legal advice. It resides in the possibility that the advice will lead a party to change his decisions, but now about whether to sue or how to conduct litigation (including settlement negotiation) rather than about the party's earlier, substantive behavior. Ex post legal advice can affect not only what legal arguments to pursue, but also how to develop evidence, what evidence to present and not to present, and how to challenge false arguments. It is virtually inevitable that

¹⁰⁹The relevance of this case is mainly hypothetical, for if a social authority understands that the law is undesirable, it would make more sense for the authority to alter the law than to seek to influence the giving of legal advice.

ex post legal advice will have substantial private value because of the complicated nature of legal procedure and the unlikelihood that potential litigants will know the law in real detail.

Advice about the bringing of suit. In considering the social value of ex post advice and comparing it to the private value of ex post advice, let us begin with advice about whether a harmed party should bring suit. The social value of this advice derives principally from the effect of suit on the prior behavior of parties who might be sued, that is, on their precautions and participation in potentially harm-producing activity. This incentive effect of suit could be small or large, and it could be either greater or less than the expected private gain from suit, as was explained in section 2 of chapter 17. Hence, the private and social value of advice about suit could diverge from each other in either direction.

Advice about the course of litigation. Now consider the social value of legal advice that defendant parties obtain during litigation. As noted in the introduction to this section, because such advice is, by its nature, imparted to parties only after they have acted, it cannot have aided them in conforming with the law, in choosing how to act if they were uncertain about the law. (The firm that does not know whether discharging a waste into a river will constitute a violation of law cannot be led to behave appropriately by learning what the law is after it decides about discharging the waste.) This simple but fundamental observation means that ex post advice to defendants does not raise social welfare in the direct way that ex ante advice to potential defendants does. Nonetheless, ex post advice certainly may influence prior behavior and social welfare.

One way that ex post advice may affect prior behavior and social welfare is by diluting sanctions and thus deterrence of undesirable conduct. Lawyers may lower expected sanctions by their advantageous use of legal strategy and, importantly, by counseling defendants on the selection of evidence to present and to suppress. Given that individuals anticipate that expected sanctions for causing harm will be reduced due to the subsequent availability of legal advice, fewer individuals will be deterred from engaging in undesirable behavior. Thus, legal advice may have negative social value.¹¹⁰ (In principle, though, a partial remedy for this problem would be for the state to raise sanctions overall to offset the dilution of deterrence due to legal advice.)

However, ex post advice may also enhance social welfare by raising otherwise inadequate sanctions. Advice may raise expected sanctions because lawyers may help plaintiffs to obtain higher judgments, better reflecting the harms they have sustained, than they would receive if they did not have legal advice.

Additionally, ex post advice may raise social welfare by lowering sanctions for defendants who did not violate the law, or who face higher sanctions than they should. If parties anticipate that, if they ever incorrectly face a legal sanction, advice will help them to avoid that sanction, they will not be undesirably discouraged from engaging in many useful activities or be led to take expensive and inordinate precautions.

In sum, ex post advice may be either socially desirable or socially undesirable, depending on context. Moreover, when advice is socially desirable, its social value could be different from its private value.

4.4 Other factors bearing on advice. Let me now consider several additional issues bearing on the effects and value of legal advice.

Subversion of the law. It has been assumed for the most part that legal advice is

¹¹⁰The point of this paragraph was early emphasized by Bentham 1827.

informational in character, conveying knowledge about the law and legal sanctions but not altering expected sanctions. Yet lawyers are sometimes able to subvert the law by effectively lowering sanctions or their probability. As mentioned above, lawyers may inappropriately reduce expected sanctions by suppressing or destroying unfavorable evidence to present, by helping clients to hide assets, and the like. Such legal assistance is to be distinguished from advice that lowers expected sanctions for bona fide reasons, for example, by demonstrating that an asserted harm was not a true harm. Of course, lawyers are not supposed to thwart law enforcement, but they have an economic incentive to do so and can fairly easily avoid punishment for it (lawyers give advice in private and can phrase their advice in hypothetical but readily understood terms). From the social perspective, legal advice that frustrates law enforcement is obviously undesirable.

Confidentiality of advice. The legal system protects the confidentiality of communications between lawyers and their clients under wide circumstances, and this protection has been implicitly assumed in the above discussion. Confidentiality of legal advice will benefit clients when there is positive probability that disclosure of advice would lower its value. This would usually be true of advice about the selection of evidence to present in litigation: Such advice generally would be robbed of effectiveness if it were disclosed to the opposing side and the court. Confidentiality is also of obvious importance to those obtaining advice subversive of the law. By contrast, confidentiality often should not matter to parties obtaining ex ante advice about the legality of an act or about magnitude or likelihood of sanctions, because disclosure of such advice will usually not disadvantage them. For example, disclosure of the ex ante advice that a party obtains about what is considered negligent behavior ordinarily should not matter to the party.

Still, whatever is the character of legal advice, maintaining the confidentiality of much information about clients that is revealed to lawyers in the course of their dealings with clients will frequently be important to the clients. For instance, a firm would usually not want information pertaining to its business plans revealed to others, and an individual would ordinarily not want information of a personal nature disclosed to outsiders.

Because protection of confidentiality can benefit clients (and never is a disadvantage to them), it encourages clients to consult with and reveal information to their lawyers. This in itself is sometimes thought to imply that confidentiality is socially desirable. That reasoning, however, is mistaken: Confidentiality is socially desirable only if the legal advice that confidentiality encourages is socially desirable, and as has been explained above, that may not be the case.

Protection of legal work product. The legal system also protects the confidentiality of legal work product -- documents and other records of lawyers' effort -- that they generate on behalf of clients in expectation of litigation. The protection of work product is accomplished principally by denying opposing litigants the legal right to discover work product (that is, the right to order the party with work product to produce it). The effect of work product protection is similar to that of protection of confidentiality, so it can be considered very briefly. Protection of work product encourages lawyers to engage in research on and development of their clients' cases, for much of the value of such research and development would be lost if it became immediately known to the other side.¹¹¹ (Thus, this protection functions analogously to

¹¹¹On this point, see for example Easterbrook 1981.

copyright.) Because protection of work product raises the value and quality of legal advice, it inures to clients' benefit. But whether protection of work product is socially desirable is not evident a priori, for it depends on whether or not the advice that the work product supports is socially desirable. A further complication is that, even when the advice is socially desirable, the private value of advice, and thus the amount of work product, may be socially excessive.

Quality and truthfulness of advice. The issue of lawyers' incentives to supply good advice and to be truthful to clients has not been addressed above. To the degree that poor or dishonest advice would be discovered and that lawyers would suffer penalties for having provided such advice, they will have reason not to give bad advice. There are two basic types of penalties that lawyers face for furnishing unsound legal advice: loss of business because of damage to reputation; and legal sanction, in the form of a damage judgment arising from a malpractice action, a fine assessed by a court, or a punishment imposed by a professional association.¹¹²

¹¹²For a general treatment of these ways of regulating lawyer conduct, see Wilkins 1992.

References

- Ashenfelter, Orley. 1998. Arbitration. In *The New Palgrave Dictionary of Economics and the Law*, edited by Peter Newman, 1:88-93. London: Macmillan.
- Ashenfelter, Orley, and David E. Bloom. 1984. Models of Arbitrator Behavior: Theory and Evidence. *American Economic Review* 74:111-24.
- Bebchuk, Lucian A. 1984. Litigation and Settlement under Imperfect Information. *Rand Journal of Economics* 15:404-15.
- 1988. Suing Solely to Extract a Settlement Offer. *Journal of Legal Studies* 17:437-50.
- 1996. A New Theory Concerning the Credibility and Success of Threats to Sue. *Journal of Legal Studies* 25:1-25.
- 1998. Suits with Negative Expected Value. In *The New Palgrave Dictionary of Economics and the Law*, edited by Peter Newman, 3:551-54. London: Macmillan.
- Bebchuk, Lucian A., and Howard Chang. 1999. The Effect of Offer-of-Settlement Rules on the Terms of Settlement. *Journal of Legal Studies* 28:489-513.
- Beltramo, Mario. 1995. Italy. In *International Civil Procedures*, edited by Christian T. Campbell, 419-94. London: Lloyd's of London Press.
- Benson, Bruce L. 2000. Arbitration. In *Encyclopedia of Law and Economics*, edited by Boudewijn Bouckaert and Gerrit De Geest, 5:159-93. Cheltenham: Edward Elgar.
- Bentham, Jeremy. 1827. *Rationale of Judicial Evidence*. Vol. 5. London: Hunt and Clarke.
- Bernstein, Lisa. 1992. Opting Out of the Legal System: Extralegal Contractual Relations in the Diamond Industry. *Journal of Legal Studies* 21:115-57.
- 1998. Private Commercial Law. In *The New Palgrave Dictionary of Economics and the Law*, edited by Peter Newman, 3:108-14. London: Macmillan.
- Braeutigam, Ronald, Bruce Owen, and John Panzar. 1984. An Economic Analysis of Alternative Fee Shifting Systems. *Law and Contemporary Problems* 47:173-85.
- Brickman, Lester, Michael Horowitz, and Jeffrey O'Connell. 1994. *Rethinking Contingency Fees*. New York: Manhattan Institute.
- Bundy, Stephen McG., and Einer R. Elhauge. 1991. Do Lawyers Improve the Adversary System? A General Theory of Litigation Advice and Its Regulation. *California Law Review* 79:313-420.
- Byrd, Robert W., and Marion Barbier. 1992. France. In *Civil Appeal Procedures Worldwide*, edited by Charles Platto, 158-68. London: Graham and Trotman.

- Carroll, Stephen J., James S. Kakalik, Nicholas M. Pace, and John L. Adams. 1991. *No-fault Approaches to Compensating People Injured in Automobile Accidents*. Santa Monica: Rand Institute for Civil Justice.
- Che, Yeon-Koo, and Jong Goo Yi. 1993. The Role of Precedents in Repeated Litigation. *Journal of Law, Economics, & Organization* 9:399-424.
- Cooley, John W., and Steven Lubet. 1997. *Arbitration Advocacy*. South Bend, Ind.: National Institute for Trial Advocacy.
- Cooter, Robert D., and Daniel L. Rubinfeld. 1989. Economic Analysis of Legal Disputes and Their Resolution. *Journal of Economic Literature* 27:1067-97.
- 1994. An Economic Model of Legal Discovery. *Journal of Legal Studies* 23:435-63.
- Couture, Tony. 1999. State. In *The Philosophy of Law*, edited by Christopher Berry Gray, 2:834-37. New York: Garland Publishing.
- Daughety, Andrew F. 2000. Settlement. In *Encyclopedia of Law and Economics*, edited by Boudewijn Bouckaert and Gerrit De Geest, 5:95-158. Cheltenham: Edward Elgar.
- Daughety, Andrew F., and Jennifer F. Reinganum. 1993. Endogenous Sequencing in Models of Settlement and Litigation. *Journal of Law, Economics, and Organization* 9:314-48.
- 2000. Appealing Judgments. *Rand Journal of Economics* 31:502-25.
- Derfner, Mary Francis, and Arthur D. Wolf. 1995. *Court Awarded Attorney Fees*. New York: Matthew Bender.
- Deweese, Don, David Duff, and Michael Trebilcock. 1996. *Exploring the Domain of Accident Law: Taking the Facts Seriously*. New York: Oxford University Press.
- Easterbrook, Frank H. 1981. Insider Trading, Secret Agents, Evidentiary Privileges, and the Production of Information. *Supreme Court Review* 1981:309-65.
- Eisenberg, Theodore. 1990. Testing the Selection Effect: A New Theoretical Framework with Empirical Tests. *Journal of Legal Studies* 19:337-58.
- Eisenberg, Theodore, and Henry S. Farber. 1997. The Litigious Plaintiff Hypothesis: Case Selection and Resolution. *Rand Journal of Economics* 28:S92-S112.
- Farber, Henry S., and Michelle J. White. 1991. Medical Malpractice: An Empirical Examination of the Litigation Process. *Rand Journal of Economics* 22:199-217.
- Farmer, Amy, and Paul Pecorino. 1996. Issues of Informational Asymmetry in Legal Bargaining. In *Dispute Resolution: Bridging the Settlement Gap*, edited by David A. Anderson, 79-105. Greenwich, Conn.: JAI Press.
- Fischel, Daniel R. 1998. Lawyers and Confidentiality. *University of Chicago Law Review* 65:1-33.

- Friedman, Alan E. 1969. An Analysis of Settlement. *Stanford Law Review* 22:67-100.
- Gertner, Robert H. 1998. Disclosure and Unravelling. In *The New Palgrave Dictionary of Economics and the Law*, edited by Peter Newman, 1:605-08. London: Macmillan.
- Goldberg, Stephen B., Frank E. A. Sander, and Nancy H. Rogers. 1999. *Dispute Resolution: Negotiation, Mediation, and Other Processes*. Third edition. Gaithersburg, Md.: Aspen Law and Business.
- Gould, John P. 1973. The Economics of Legal Conflicts. *Journal of Legal Studies* 2:279-300.
- Gravelle, H. S. E. 1993. The Efficiency Implications of Cost-Shifting Rules. *International Review of Law and Economics* 13:3-18.
- Hause, John C. 1989. Indemnity, Settlement, and Litigation, or I'll Be Suing You. *Journal of Legal Studies* 18:157-79.
- Hay, Bruce L. 1994. Civil Discovery: Its Effects and Optimal Scope. *Journal of Legal Studies* 23:481-515.
- 1995. Effort, Information, Settlement, Trial. *Journal of Legal Studies* 24:29-62.
- Hay, Bruce L., and Kathryn E. Spier. 1998. Settlement of Litigation. In *The New Palgrave Dictionary of Economics and the Law*, edited by Peter Newman, 3:442-51. London: Macmillan.
- Hensler, Deborah R., Mary E. Vaiana, James S. Kakalik, and Mark A. Peterson. 1987. *Trends in Tort Litigation*. R-3583-ICJ. Santa Monica: RAND Institute for Civil Justice.
- Herzog, Peter E., and Delmar Karlen. 1982. Attacks on Judicial Decisions. Chapter 8 in volume 16, *Civil Procedure*, edited by Mauro Cappelletti. *International Encyclopedia of Comparative Law*. Tübingen: J. C. B. Mohr.
- Hughes, James W., and Edward A. Snyder. 1995. Litigation and Settlement under the English and American Rules: Theory and Evidence. *Journal of Law and Economics* 38:225-50.
- 1998. Allocation of Litigation Costs: American and English Rules. In *The New Palgrave Dictionary of Economics and the Law*, edited by Peter Newman, 1:51-56. London: Macmillan.
- Hylton, Keith N. 1993. Asymmetric Information and the Selection of Disputes for Litigation. *Journal of Legal Studies* 22:187-210.
- Joost, Robert H. 1992. *Automobile Insurance and No-fault Law*. Second edition. Deerfield, Ill.: Clark Boardman Callaghan.
- Kakalik, James S., and Nicholas M. Pace. 1986. *Costs and Compensation Paid in Tort Litigation*. Report R-3391-ICJ. Santa Monica: RAND Institute for Civil Justice.
- Kaplow, Louis. 1986b. Private versus Social Costs in Bringing Suit. *Journal of Legal Studies*

15:371-85.

----- 1994. The Value of Accuracy in Adjudication: An Economic Analysis. *Journal of Legal Studies* 23:307-401.

Kaplow, Louis, and Steven Shavell. 1989. Legal Advice About Information to Present in Litigation: Its Effects and Social Desirability. *Harvard Law Review* 102:565-615.

----- 1990. Legal Advice about Acts Already Committed. *International Review of Law and Economics* 10:149-59.

----- 1992. Private versus Socially Optimal Provision of Ex Ante Legal Advice. *Journal of Law, Economics, & Organization* 8:306-20.

----- 1994a. Accuracy in the Determination of Liability. *Journal of Law and Economics* 37:1-15.

----- 1996a. Accuracy in the Assessment of Damages. *Journal of Law and Economics* 39:191-210.

----- 1996b. Property Rules versus Liability Rules: An Economic Analysis. *Harvard Law Review* 109:713-90.

Katz, Avery. 1987. Measuring the Demand for Litigation: Is the English Rule Really Cheaper? *Journal of Law, Economics, & Organization* 3:143-76.

----- 1988. Judicial Decisionmaking and Litigation Expenditure. *International Review of Law and Economics* 8:127-43.

----- 1990a. The Effect of Frivolous Lawsuits on the Settlement of Litigation. *International Review of Law and Economics* 10:3-27.

Kennan, John, and Robert Wilson. 1993. Bargaining with Private Information. *Journal of Economic Literature* 31:45-104.

Kritzer, Herbert M. 1984. Fee Arrangements and Fee Shifting: Lessons from the Experience in Ontario. *Law and Contemporary Problems* 47:125-38.

Landes, William M. 1971. An Economic Analysis of the Courts. *Journal of Law and Economics* 14:61-107.

Landes, William M., and Richard A. Posner. 1975. The Private Enforcement of Law. *Journal of Legal Studies* 4:1-46.

Lee, Rex E. 1985. The American Courts as Public Goods: Who Should Pay the Costs of Litigation? *Catholic University Law Review* 34:267-76.

Loewenstein, George, Samuel Issacharoff, Colin Camerer, and Linda Babcock. 1993. Self-Serving Assessments of Fairness and Pretrial Bargaining. *Journal of Legal Studies* 22:135-59.

- Menell, Peter S. 1983. A Note on the Private versus Social Incentives to Sue in a Costly Legal System. *Journal of Legal Studies* 12:41-52.
- Meurer, Michael J. 1992. The Gains from Faith in an Unfaithful Agent: Settlement Conflicts between Defendants and Liability Insurers. *Journal of Law, Economics, & Organization* 8:502-22.
- Miller, Geoffrey P. 1986. An Economic Analysis of Rule 68. *Journal of Legal Studies* 15:93-125.
- Mnookin, Robert H. 1993. Why Negotiations Fail: An Exploration of Barriers to the Resolution of Conflict. *Ohio State Journal on Dispute Resolution* 8:235-49.
- Mnookin, Robert, and Robert Wilson. 1998. A Model of Efficient Discovery. *Games and Economic Behavior* 25:219-50.
- Ordover, Janusz A. 1978. Costly Litigation in the Model of Single Activity Accidents. *Journal of Legal Studies* 7:243-61.
- Osborne, Evan. 1999. Who Should Be Worried about Asymmetric Information in Litigation? *International Review of Law and Economics* 19:399-409.
- Ostrom, Brian J., Neil B. Kauder, and Robert C. LaFountain. 2001a. *Examining the Work of State Courts, 1999-2000*. Williamsburg, Va: National Center for State Courts.
- 2001b. *Examining the Work of State Courts, 2001*. Williamsburg, Va: National Center for State Courts.
- Palmer, Geoffrey. 1994. New Zealand's Accident Compensation Scheme: Twenty Years On. *University of Toronto Law Journal* 44:223-73.
- Platto, Charles, editor. 1992. *Civil Appeal Procedures Worldwide*. London: Graham and Trotman.
- Png, Ivan P. L. 1986. Optimal Subsidies and Damages in the Presence of Judicial Error. *International Review of Law and Economics* 6:101-05.
- Polinsky, A. Mitchell, and Daniel L. Rubinfeld. 1988. The Deterrent Effects of Settlements and Trials. *International Review of Law and Economics* 8:109-16.
- 1993. Sanctioning Frivolous Suits: An Economic Analysis. *Georgetown Law Journal* 82:397-435.
- 1996. Optimal Awards and Penalties When the Probability of Prevailing Varies Among Plaintiffs. *RAND Journal of Economics* 27:269-80.
- Posner, Richard A. 1972a. *Economic Analysis of Law*. Boston: Little, Brown.
- 1973. Strict Liability: A Comment. *Journal of Legal Studies* 2:205-21.

- 1985a. An Economic Theory of the Criminal Law. *Columbia Law Review* 85:1193-1231.
- 1985b. *The Federal Courts: Crisis and Reform*. Cambridge, Mass.: Harvard University Press.
- Priest, George L. 1988. Products Liability Law and the Accident Rate. In *Liability: Perspectives and Policy*, edited by Robert Litan and Clifford Winston, 184-222. Washington, D.C.: Brookings Institution.
- Priest, George L., and Benjamin Klein. 1984. The Selection of Disputes for Litigation. *Journal of Legal Studies* 13:1-55.
- Raiffa, Howard. 1968. *Decision Analysis: Introductory Lectures on Choices Under Uncertainty*. Reading, Mass.: Addison-Wesley.
- Ramseyer, J. Mark, and Minoru Nakato. 1989. The Rational Litigant: Settlement Amounts and Verdict Rates in Japan. *Journal of Legal Studies* 18:263-90.
- Reinganum, Jennifer F., and Louis L. Wilde. 1986. Settlement, Litigation, and the Allocation of Litigation Costs. *Rand Journal of Economics* 17:557-66.
- Rose-Ackerman, Susan, and Mark Geistfeld. 1987. The Divergence Between Social and Private Incentives to Sue: A Comment on Shavell, Menell, and Kaplow. *Journal of Legal Studies* 16:483-91.
- Rosenberg, David, and Steven Shavell. 1985. A Model in Which Suits Are Brought for Their Nuisance Value. *International Review of Law and Economics* 5:3-13.
- Rowe, Thomas D., Jr. 1982. The Legal Theory of Attorney Fee Shifting: A Critical Overview. *Duke Law Journal* 1982:651-80.
- Rubinfeld, Daniel L. 1998. Discovery. In *The New Palgrave Dictionary of Economics and the Law*, edited by Peter Newman, 1:609-15. London: Macmillan.
- Rubinfeld, Daniel L., and Suzanne Scotchmer. 1993. Contingent Fees for Attorneys: An Economic Analysis. *RAND Journal of Economics* 24:343-56.
- 1998. Contingent Fees. In *The New Palgrave Dictionary of Economics and the Law*, edited by Peter Newman, 1:415-20. London: Macmillan.
- Schrag, Joel L. 1999. Managerial Judges: An Economic Analysis of the Judicial Management of Legal Discovery. *RAND Journal of Economics* 30:305-23.
- Schweizer, Urs. 1989. Litigation and Settlement under Two-Sided Incomplete Information. *Review of Economic Studies* 56:163-77.
- Shavell, Steven. 1982a. On Liability and Insurance. *Bell Journal of Economics* 13:120-32.
- 1982b. The Social versus the Private Incentive to Bring Suit in a Costly Legal System. *Journal of Legal Studies* 11:333-39.

- 1982c. Suit, Settlement, and Trial: A Theoretical Analysis Under Alternative Methods for the Allocation of Legal Costs. *Journal of Legal Studies* 11:55-81.
- 1985b. Uncertainty Over Causation and the Determination of Civil Liability. *Journal of Law and Economics* 28:587-609.
- 1988. Legal Advice about Contemplated Acts: The Decision to Obtain Advice, Its Social Desirability, and Protection of Confidentiality. *Journal of Legal Studies* 17:123-50.
- 1989. Sharing of Information Prior to Settlement or Litigation. *RAND Journal of Economics* 20:183-95.
- 1995a. Alternative Dispute Resolution: An Economic Analysis. *Journal of Legal Studies* 24:1-28.
- 1996. Any Frequency of Plaintiff Victory at Trial Is Possible. *Journal of Legal Studies* 25:493-501.
- 1997. The Fundamental Divergence Between the Private and the Social Motive to Use the Legal System. *Journal of Legal Studies* 26:575-612.
- 1999. The Level of Litigation: Private versus Social Optimality of Suit and of Settlement. *International Review of Law and Economics* 19:99-115.
- Shepherd, George B. 1999. An Empirical Study of the Economics of Pretrial Discovery. *International Review of Law and Economics* 19:245-63.
- Sieg, Holger 2000. Estimating a Bargaining Model with Asymmetric Information: Evidence from Medical Malpractice Disputes. *Journal of Political Economy* 108:1006-21.
- Sloan, Frank A. 1998. Automobile Accidents, Insurance, and Tort Liability. In *The New Palgrave Dictionary of Economics and the Law*, edited by Peter Newman, 1:140-44. London: Macmillan.
- Snyder, Edward A., and James M. Hughes. 1990. The English Rule for Allocating Legal Costs: Evidence Confronts Theory. *Journal of Law, Economics & Organization* 6:345-80.
- Sobel, Joel. 1989. An Analysis of Discovery Rules. *Law and Contemporary Problems* 52(no. 1):133-59.
- Spier, Kathryn E. 1992a. The Dynamics of Pretrial Negotiation. *Review of Economic Studies* 59:93-108.
- 1994a. Pretrial Bargaining and the Design of Fee-Shifting Rules. *RAND Journal of Economics* 25:197-214.
- 1994b. Settlement Bargaining and the Design of Damage Awards. *Journal of Law, Economics & Organization* 10:84-95.
- 1997. A Note on the Divergence between the Private and the Social Motive to Settle under a

- Negligence Rule. *Journal of Legal Studies* 26:613-21.
- Spitzer, Matthew, and Eric Talley. 2000. Judicial Auditing. *Journal of Legal Studies* 29:649-83.
- Statistical Abstract of the United States 2001*. 2001. Economics and Statistics Administration, U.S. Census Bureau. Washington, D.C.: U.S. Department of Commerce.
- Sykes, Alan O. 1994. "Bad Faith" Refusal to Settle by Liability Insurers: Some Implications of the Judgment-proof Problem. *Journal of Legal Studies* 23:77-110.
- Tillinghast-Towers Perrin. 2002. *U.S. Tort Costs 2000: Trends and Findings on the Costs of the U.S. Tort System*. February 2002. Chicago: Tillinghast-Towers Perrin.
- U. S. Department of Justice. 2001a. *Judicial Business of the United States Courts, 2001 Report of the Director*. Administrative Office of the United States Courts, Statistics Division. Washington, D.C.: U.S. Department of Justice.
- Viscusi, W. Kip. 1986a. The Determinants of the Disposition of Product Liability Claims and Compensation for Bodily Injury. *Journal of Legal Studies* 15:321-46.
- Waldfogel, Joel. 1995. The Selection Hypothesis and the Relationship between Trial and Plaintiff Victory. *Journal of Political Economy* 103:229-60.
- 1998. Selection of Cases for Trial. In *The New Palgrave Dictionary of Economics and the Law*, edited by Peter Newman, 3:419-24. London: Macmillan.
- Ware, Stephen J. 2001. *Alternative Dispute Resolution*. St. Paul, Minn.: West Group.
- Wilkins, David B. 1992. Who Should Regulate Lawyers? *Harvard Law Review* 105:799-887.
- Wittman, Donald. 1985. Is the Selection of Cases for Trial Biases? *Journal of Legal Studies* 14:185-214.