

This work is distributed as a Discussion Paper by the
STANFORD INSTITUTE FOR ECONOMIC POLICY RESEARCH

SIEPR Discussion Paper No. 05-05
Economic Analysis of Law

By
A. MITCHELL POLINSKY
and
STEVEN SHAVELL
November, 2005

Stanford Institute for Economic Policy Research
Stanford University
Stanford, CA 94305
(650) 725-1874

The Stanford Institute for Economic Policy Research at Stanford University supports research bearing on economic and public policy issues. The SIEPR Discussion Paper Series reports on research and policy analysis conducted by researchers affiliated with the Institute. Working papers in this series reflect the views of the authors and not necessarily those of the Stanford Institute for Economic Policy Research or Stanford University.

Economic Analysis of Law

A. Mitchell Polinsky and Steven Shavell*

Abstract: This entry for the forthcoming *The New Palgrave Dictionary of Economics* (Second Edition) surveys the economic analysis of five primary fields of law: property law; liability for accidents; contract law; litigation; and public enforcement and criminal law. It also briefly considers some criticisms of the economic analysis of law.

Keywords: law and economics; property law; liability for accidents; contract law; litigation; public enforcement; criminal law

JEL classification: D02; D23; D63; H23; J28; K11; K12; K13; K14; K32; K41; K42; P14; P48

* Stanford Law School and Harvard Law School, respectively. Both authors also are affiliated with the National Bureau of Economic Research. Polinsky's research was supported by the John M. Olin Program in Law and Economics at Stanford Law School, and Shavell's research was supported by the John M. Olin Center for Law, Economics, and Business at Harvard Law School.

Economic analysis of law seeks to identify the effects of legal rules on the behavior of relevant actors and whether these effects are socially desirable. The approach employed is that of economic analysis generally: the behavior of individuals and firms is described assuming that they are forward looking and rational, and the framework of welfare economics is adopted to assess the social desirability of outcomes. The field may be said to have begun with Bentham (1789), who systematically examined how actors would behave in the face of legal incentives (especially criminal sanctions) and who evaluated outcomes with respect to a clearly stated measure of social welfare (utilitarianism). His work was left essentially undeveloped until four important contributions were made: Coase (1960) on externalities and liability, Becker (1968) on crime and law enforcement, Calabresi (1970) on accident law, and Posner (1972) on economic analysis of law in general.

Our focus here will be on the analytical foundations of five basic legal subjects: property, torts, contracts, civil litigation, and crime and law enforcement (on these, see generally Cooter and Ulen, 2003; Posner, 2003; Miceli, 1997; and Shavell, 2004). We do not treat more particular areas of law, such as antitrust, corporate, and tax law, nor do we cite empirical work; for surveys of these and other areas of law and economics, including empirical studies, see Polinsky and Shavell (2006).

1. PROPERTY LAW. *Justification and emergence of property rights.* A beginning question is why there should be property rights in things. A number of factors have been stressed, especially by early writers. These include that property rights furnish incentives to work and to maintain durable things; that the rights make trade possible; and that if such rights were absent, individuals would spend effort trying to take things from each other and protecting their things.

Property rights would be expected to emerge when their advantages become sufficiently great. For example, Demsetz (1967) explains the development of property rights in land among Indians as a way of preventing overly intensive hunting of valuable animals. Umbeck (1981) shows that when gold was discovered in California in 1848, property rights in gold-bearing land and river beds developed, as this encouraged individuals to pan for gold and to build sluices; it also curbed wasteful efforts to grab land from one another. For a survey, see Libecap (1986).

Division of property rights. Property rights can be viewed as composed of *possessory* rights — rights of use — and rights to *transfer* possessory rights. Thus, what we commonly conceive of as ownership (say, of land) entails both a large swath of possessory rights (rights to build on land, plant on it, under most contingencies, and into the infinite future) and associated rights to transfer them. Property rights in things are generally held in substantially agglomerated bundles, but there is also significant partitioning of rights contemporaneously, over time and contingencies, and according to whether the rights are possessory or are for transfer. For example, an owner of land may not hold complete possessory rights, in that others may possess an easement giving them the right of passage upon his land, or the right to take timber, or the right to extract oil if found (thus a contingent right). A rental agreement constitutes a division of property rights over time. Trust arrangements, such as those under which an adult manages property for a child, divide possessory rights and rights to transfer.

The division of property rights may be valuable when different parties derive different benefits from them, because gains can then be achieved if rights are allocated to those who obtain the most from them. There may, however, be disadvantages to the division of rights, including that externalities may arise (a person with a right of passage might trample crops).

Public property and its acquisition; takings and compensation. An important class of property is that owned by the public. As is well known, the main justifications for public

property concern difficulty that private providers would obtain in charging for certain goods and services.

When it is desirable for the state to acquire property for public use, the state can purchase it or take it, through the exercise of the power of *eminent domain*. In the latter case, the law typically provides that the state must compensate property owners for the value of what has been taken from them.

A difference between purchase and compensated takings is that the amounts owners receive is determined by negotiation in the former case but unilaterally by the state in the latter. Because of errors in state determination of value, as well as concern about the behavior of government officials, purchase would ordinarily be superior to compensated takings. When, however, the state needs to assemble many contiguous parcels, such as for a road, acquisition by purchase might be stymied by hold-out problems, making the power to take socially advantageous.

Assuming that there is a reason for the state to take property, a requirement to pay compensation may curb problems of overzealousness or abuse of authority by public officials, yet it may also exacerbate potential problems of too little public activity, because public authorities do not directly receive the benefits of takings (Kaplow, 1986). Payment of compensation also may lead property owners to invest excessively in property; see Blume et al. (1984).

Acquisition of property in unowned things. The law has to determine under what conditions a person will become a legal owner of previously unowned things, such as wild animals, fish, and mineral and oil deposits. Under the *finders-keepers rule*, incentives to invest in capture (such as to hunt for animals or explore for oil) are optimal if only one person is making the effort. However, if many individuals seek unowned things, they will invest a

socially excessive amount of resources in search: one person's investment usually will come, at least partly, at the expense of other persons' likelihood of finding unowned things. Various aspects of the law ameliorate this problem of excessive search effort. Notable examples are that regulations may limit the quantities that can be taken of fish and wild animals, the right to search for minerals on the ocean floor may be auctioned, and oil extraction rights may be assigned to one party.

Acquisition of good title when property is sold. A basic difficulty associated with sale of property that a legal system must solve is establishing validity of ownership or *title*. Good title is important for trade, since buyers want to be assured that they have property rights in what they purchase. But if any sale gives a buyer good title, theft is encouraged, since thieves could then easily sell stolen goods. Under a *registration system*, good title means that one's name is listed in the registry as the owner, and title passes at the time of sale by an authorized change in the registry. Hence, buyers can clearly determine whether they are obtaining good title by checking the registry, and a thief could not easily sell stolen property by claiming that he has good title. Registries, however, are expensive to establish and maintain.

In the absence of registries, the law may employ the *original ownership rule*, under which the buyer does not obtain good title if the seller did not have good title. Alternatively, under the *bona fide purchase rule*, a buyer acquires good title as long as he had reason to think that the sale was legitimate, even if the item sold was in fact wrongfully obtained. This rule makes theft more attractive because thieves will often be able to sell their property to buyers who will be motivated to "believe" that sales are bona fide.

Adverse possession. The legal doctrine of adverse possession allows involuntary transfer of land: a person is deemed to become the legal owner of land if he takes possession of it and uses it openly for at least a prescribed period, such as ten years. It may appear that this rule

could be desirable because it encourages productive use of idle land. But this overlooks the possibility that there may be good reasons for allowing land to remain idle, plus a prospective adverse possessor could always bargain with the owner to rent or buy the land. Additionally, the rule induces owners to expend resources policing incursions and potential adverse possessors to attempt possession. A historical justification for the rule is that, before reliable land registries existed, it allowed a seller of land to establish good title to a buyer relatively easily: the seller need only show that he was on the land for the prescribed period.

Constraints on sale of property. Legal restrictions are often imposed on the sale of goods and services. One standard justification is externalities. For example, the sale of fireworks might be banned because of the externality that their ownership creates, namely, putting others at risk of injury. The other standard justification for legal restrictions on sale is lack of consumer information. For instance, a drug may not be sold without a prescription because of fear that otherwise buyers would not use it properly. Rather than restrict sales, however, the government could supply relevant information to consumers, for example, indicating that the drug has dangerous side effects, or that it should only be taken with the advice of a medical expert.

Externalities. When individuals use property, they may cause externalities, namely, harm or benefit to others. As a general matter, it is socially desirable for individuals to do more than is in their self-interest to reduce detrimental externalities and to act so as to increase beneficial externalities. The socially optimal resolution of harmful externalities often involves the behavior of victims as well as that of injurers. If victims can do things to reduce the amount of harm more cheaply than injurers (say install air filters to avoid pollution), it is optimal for victims to do so. Moreover, victims can sometimes alter their locations to reduce their exposure to harm.

Legal intervention can ameliorate problems of externalities. A major form of intervention that has been studied is *direct regulation*, under which the state restricts permissible

behavior, such as requiring factories to use smoke arrestors. Closely related is the *injunction*, whereby a potential victim can enlist the power of the state to force a potential injurer to take steps to prevent harm or to cease his activity. Society can also make use of financial incentives to induce injurers to reduce harmful externalities. Under the *corrective tax*, a party pays the state an amount equal to the expected harm he causes, for example, the expected harm due to a discharge of a pollutant into a lake. There is also *liability*, a privately-initiated means of providing financial incentives, under which injurers pay for harm done if sued by victims. These methods differ in the information that the state needs to apply them, in whether they require or harness information that victims have about harm, and in other respects, such that each may be superior to the other in different circumstances (Shavell, 1993).

Parties affected by externalities will sometimes have the opportunity to make mutually beneficial agreements with those who generate the externalities, as Coase (1960) stressed. But bargaining may not occur for many reasons: cost; collective action problems (such as when many victims each face small harms); and lack of knowledge of harm (such as from an invisible carcinogen). If bargaining does occur, it may not be successful, owing to asymmetric information. These difficulties often make bargaining a problematic solution to externality problems and imply that liability rules are needed, as discussed by Calabresi and Melamed (1972).

Property rights in information. The granting of property rights in information, notably the award of *patents* for inventions and *copyrights* for written works and certain other compositions, involves a major social benefit — provision of incentives to create intellectual works — but also a social disadvantage — creation of power to price above marginal cost. Patent and copyright law have been examined to ascertain how they reflect the tradeoff between this benefit and disadvantage. A distinct form of legal protection is *trade secret law*, comprising

various doctrines of contract and tort law that serve to protect a range of commercially valuable information that is not (or cannot be) protected by patent or copyright, such as customer lists. On property rights in information, see generally Landes and Posner (2003).

An alternative to property rights in information is for the state to offer *rewards* to creators of information and for information that is developed to be made available to all who want it. Thus, an author of a book would receive a reward from the state for the writing of the book, possibly based on sales of the book, but anyone who wanted to print it and sell it could do so. This system would create incentives for the creation of information without distorting prices, but requires the state to choose the magnitude of rewards.

Property rights in labels. Many goods and services are identified by labels, which have substantial social value because the quality of goods and services may be hard for consumers to determine directly. Labels enable consumers to purchase goods and services on the basis of product quality without requiring consumers to independently determine quality; a person who wants to stay at a high-quality hotel in another city can choose such a hotel merely by its label, such as “Ritz Hotel.” In addition, sellers who label their output will have an incentive to produce goods and services of quality because consumers will recognize quality through sellers’ labels. This basic reasoning is used to justify property rights in *trademarks*, as discussed by Landes and Posner (1987b).

2. LIABILITY FOR ACCIDENTS. Legal liability for accidents, which is governed by tort law, is a means by which society can reduce the risk of harm by threatening potential injurers with having to pay for the harms they cause. Liability is also frequently viewed as a device for compensating victims of harm, but we will emphasize that insurance can provide compensation more cheaply than the liability system. There are two basic rules of liability.

Under *strict liability*, an injurer must always pay for harm due to an accident that he causes.

Under the *negligence rule*, an injurer must pay for harm caused only when he is found negligent, that is, only when his level of care was less than a standard of care chosen by the courts, often referred to as *due care*. (There are various versions of these rules that depend on whether victims' care was insufficient.) In practice, the negligence rule is the dominant form of liability; strict liability is reserved mainly for certain especially dangerous activities. On economic analysis of liability for accidents, see generally Calabresi (1970), Landes and Posner (1987a), and Shavell (1987a).

Incentives to take care. In order to focus on how liability affects the incentive to prevent harm, assume first that parties are risk neutral and that accidents are *unilateral* — only injurers (not victims) influence risk — by their choice of *care* x . Let $p(x)$ be the probability of an accident that causes harm h , where p is declining in x . Assume that the social objective is to minimize total expected costs, $x + p(x)h$, and let x^* denote the optimal x .

Under strict liability, injurers pay damages equal to h whenever an accident occurs, and they naturally bear the cost of care x . Thus, they minimize $x + p(x)h$; accordingly, they choose x^* .

Under the negligence rule, suppose that the due care level is set equal to x^* , meaning that an injurer who causes harm will have to pay h if $x < x^*$, but will not have to pay anything if $x \geq x^*$. Then it can be shown that the injurer will choose x^* : clearly, the injurer will not choose x greater than x^* ; and he will not choose $x < x^*$, for then he will be liable (in which case the analysis of strict liability shows that he would not choose $x < x^*$). Thus, under both forms of liability, injurers are led to take optimal care. Note that to apply the negligence rule, courts need sufficient information to calculate x^* and to observe x , whereas under strict liability they only have to observe x .

The analysis of incentives and liability has been undertaken as well for *bilateral* accidents, in which victims also take care, and when there is uncertainty in the determination of negligence (such as due to imperfect observation of x). On incentives and liability for unilateral and bilateral accidents, see originally Brown (1973) (and see also Diamond, 1974).

Level of activity. An important extension allows for injurers to choose their *level of activity* z , which is interpreted as the (continuously variable) number of times they engage in their activity (or if injurers are firms, their output). Let $b(z)$ be the benefit (or profit) from the activity, and assume the social objective is to maximize $b(z) - z(x + p(x)h)$; here $x + p(x)h$ is assumed to be the cost of care and expected harm each time an injurer engages in his activity. Let x^* and z^* be optimal values. Note that x^* minimizes $x + p(x)h$, so x^* is as described above, and that z^* is determined by $b'(z) = x^* + p(x^*)h$, which is to say, the marginal benefit from the activity equals the marginal social cost.

Under strict liability, an injurer will choose both the level of care and the level of activity optimally, as his objective will be the same as the social objective, to maximize $b(z) - z(x + p(x)h)$, because damage payments equal h whenever harm occurs. Under the negligence rule, an injurer will choose optimal care x^* as before, but his level of activity z will be socially excessive. In particular, because an injurer will escape liability by taking care x^* , he will choose z to maximize $b(z) - zx^*$, so that z will satisfy $b'(z) = x^*$. The injurer's cost of raising his level of activity is only his cost of care x^* , which is less than the social cost, which also includes $p(x^*)h$. On liability and the level of activity, see Shavell (1980b).

The failure of the negligence rule to control the level of activity arises because negligence is defined here (and also generally in practice) in terms of care alone. A justification for this restriction is the difficulty courts would face in determining the optimal activity level z^* and the actual z . The failure of the negligence rule to control the injurer's level of activity is

applicable to any aspect of injurer behavior that would be difficult to regulate directly (including, for example, research and development activity). If, however, courts were able to incorporate all aspects of injurer behavior into the definition of due care, the negligence rule would result in optimal behavior in all respects. (Note that the variable x in the original problem could be interpreted as a vector, with each element corresponding to a dimension of behavior.)

Product liability. Another extension of the model of liability and incentives concerns product liability, the liability of firms for harms suffered by their customers. Here the degree to which liability creates incentives to reduce risk depends on customer knowledge of risk. If their knowledge is perfect, liability does not affect incentives since customers will recognize risky products and pay appropriately less for them. If their knowledge is imperfect, there is a role for liability, in many respects similar to what has been discussed above.

Risk-bearing and insurance. In addition to affecting incentives to reduce harm, the socially optimal resolution of the accident problem involves the spreading of risk to lessen risk-bearing by risk-averse parties. Risk-bearing is relevant not only because potential victims may face the risk of accident losses, but also because potential injurers may face the risk of liability. The former risk can be mitigated through so-called first-party insurance that covers losses suffered in accidents, and the latter through liability insurance.

Because risk-averse individuals tend to purchase insurance, the incentives associated with liability do not function in the direct way discussed above, but instead are mediated by the terms of insurance policies. To illustrate, consider strict liability in the unilateral accident model with care alone allowed to vary, and assume that insurance is sold at actuarially fair rates. If injurers are risk averse and liability insurers can observe their levels of care, injurers will purchase full liability insurance coverage and their premiums will depend on their level of care; their premiums will equal $p(x)h$. Thus, injurers will want to minimize their costs of care plus

premiums, or $x + p(x)h$, so they will choose the optimal level of care x^* . In this instance, liability insurance eliminates risk for injurers, and the situation reduces to the previously analyzed risk-neutral case. (Victims do not bear risk either because, in the present case, they are fully compensated for their losses.)

If, however, liability insurers cannot observe levels of care, insurance policies with full coverage would create severe moral hazard, and so would not be purchased. Instead, as we know from the theory of insurance, the typical amount of coverage purchased will be partial, for that leaves injurers with an incentive to reduce risk. In this case, therefore, the liability rule results in some direct incentive to take care because injurers are left bearing some risk after their purchase of liability insurance. But levels of care will still tend to be less than first-best.

This last observation raises the question of whether the sale of liability insurance is socially desirable. (We note that because of concern about diluted incentives, liability insurance was delayed for decades in many countries and is sometimes forbidden today, such as against punitive damages.) Notwithstanding the moral hazard problem, the sale of liability insurance is socially desirable, at least in basic models of accidents and some variations of them. This is because, if the liability insurer and the injurer together have to pay for the harm caused, the insurance policy will appropriately balance the social desire to reduce harm and the social desire to reduce risk-bearing.

Parallel observations apply under the negligence rule, where the focus of concern is on the bearing of risk by victims since injurers generally will take due care and not be liable. Risk-averse potential victims will tend to purchase first-party accident insurance.

The presence of insurance implies that the liability system cannot be justified primarily as a means of compensating risk-averse victims against loss. Rather, the justification for the liability system must lie in significant part in the incentives that it creates to reduce risk. To

amplify, although both strict liability and the insurance system can compensate victims, the liability system is much more expensive than the insurance system (see below). Accordingly, if there were not a social need to create incentives to reduce risk, it would be best to dispense with the liability system and to rely on insurance to accomplish compensation. On liability and insurance, see Shavell (1982a).

Administrative costs. The administrative costs of the liability system — the legal costs and effort of litigants involved in suit, settlement, and trial — are substantial, generally exceeding the amounts received by victims. Consideration of administrative costs affects the comparison of liability rules, but it is not clear which rule involves greater expense: more cases are brought under strict liability than under the negligence rule (victims will not sue under the negligence rule if they believe the injurer was not negligent), but the cost of resolving a case should be greater under the negligence rule (because due care and the injurer's care level need to be ascertained). The existence of administrative costs raises the question of whether the incentive benefits of the liability system justify incurring these costs, as well as whether the private incentive to sue is socially optimal. These questions will be discussed in section 4.

3. CONTRACTS. A contract is a specification of the actions that named parties are supposed to take at various times, as a function of the conditions that then obtain. A contract is said to be completely-detailed, or simply *complete*, if the contract provides *explicitly* for all possible conditions. An incomplete contract may well cover all conditions by implication. A contract stating merely that a specified price will be paid for a bushel of wheat is incomplete because it does not mention many contingencies that might affect the parties. Note that such an incomplete contract has no *gaps*, as it stipulates what the parties are to do in all circumstances. Typically, incomplete contracts do not include conditions which, were they easy to include,

would allow both parties to be made better off in an expected sense.

Contracts are here assumed to be enforced by a tribunal, which will usually be interpreted to be a state-authorized court, but it could also be another entity, such as an arbitrator or the decisionmaking body of a trade association or a religious group. (Reputation and other non-legal factors may also serve to enforce contracts, but these will not be discussed.) Enforcement refers to actions taken by the tribunal when one or more of the parties to the contract decide to come before it.

General reasons for contracts. At a broad level, parties make contracts when they have a need to make plans. They also want contracts enforced to prevent opportunistic behavior that otherwise might occur during the course of the contractual relationship and stymie fulfillment of their plans. There are two basic contexts in which parties make enforceable contracts.

The first concerns virtually any kind of financial arrangement. The necessity of contract enforcement here is transparent. In financial arrangements, there is often a party who extends credit to another for some time period, and contract enforcement prevents his credit from being appropriated, which otherwise would render the arrangements impossible. For example, if borrowers are not forced to repay loans, loans would be unworkable. In addition, financial contracts that allocate risk would generally be useless without enforcement because, once the risky outcome became known, one of the parties would not wish to honor the contract.

The second context in which parties make enforceable contracts involves the supply of customized or specialized goods and services, which cannot be purchased on a spot market in a simultaneous exchange for money. The need for enforcement of agreements for supply of customized goods and services inheres in several advantages: averting problems of holdup, which might distort incentives to invest in the contractual enterprise; allocation of risk; and prevention of inappropriate breach or performance, which can result from imperfect bargaining

due to sheer cost or asymmetric information.

Contract formation. The formation of contracts is of interest in several respects. One issue concerns search effort (Diamond and Maskin, 1979). Parties expend effort in finding contractual partners, and it is apparent that their search effort will not generally be socially optimal. On one hand, they might not search enough: because the joint gain from contracting will generally be divided between the parties through the bargaining process, the private return to search may be less than the social return. On the other hand, parties might search more than is socially desirable because of a negative externality associated with discovery of a contract partner: when one party finds and contracts with a second, other parties are thereby prevented from contracting with that party.

A basic question that a tribunal must answer is at what stage of interactions between parties does a contract become legally recognized? The general legal rule is that contracts are recognized if and only if both parties give a clear indication of assent, such as signing their names on a document. This rule allows parties to make enforceable contracts when they so desire, and it also protects parties from becoming legally obligated against their wishes, such as from one party's reliance on the other's statements (Bebchuk and Ben-Shahar, 2001; Wils, 1993). Mutual assent sometimes is not simultaneous; one party will make an offer and time will pass before the other agrees. An issue that this raises is how long, and under what circumstances, the offeror will want to be held to his offer, and whether he should be held to it. If an offeror is held to his terms, offerees will often be led to invest effort in investigating contractual opportunities. Otherwise, offerees might be taken advantage of by offerors if the offerees expressed serious interest after costly investigation (the offeror could change to less favorable terms). The anticipation of such offeror advantage-taking would reduce offerees' incentive to engage in investigation and thus diminish mutually beneficial contract formation.

(See, for example, Craswell, 1996; Katz, 1990, 1996.)

Another issue of note is disclosure of information at the time of contract formation. Disclosure may be socially beneficial because the disclosed information may be desirably employed by one of the parties; suppose, for instance, that a buyer of a house learns from the seller that the basement leaks and thus decides not to store valuables there. However, a disclosure obligation discourages parties from investing in acquisition of information (Kronman, 1978). For example, an oil company contemplating buying land might decide against conducting a geological analysis of it to determine its oil-bearing potential if the company would be required to disclose its findings to the seller of the land, as the seller would then demand a price reflecting the value of the land. The social welfare consequences of the effect of a disclosure obligation on the motive to acquire information depend on whether the information is socially valuable or mere foreknowledge, on whether the party acquiring information is the buyer or the seller, and on inferences that would be made from silence (Shavell, 1994).

Even if both parties have given their assent, a contract will not be recognized if it was made when one of the parties was put under undue pressure, for example, if he was physically or otherwise threatened by another. This legal rule has virtues similar to those of laws against theft; it reduces individuals' incentives to expend effort making threats and to defend against them.

In addition, contracts may not be legally recognized if they are made in emergency situations, such as when the owner of a ship in distress promises to pay an exorbitant amount for rescue. Nonenforcement in such situations beneficially provides potential victims with implicit insurance against having to pay high prices, but it also reduces incentives for rescue.

Incomplete nature of contracts and their less-than-rigorous enforcement. Contracts are commonly observed to be significantly incomplete, leaving out all manner of variables and contingencies that are of potential relevance to contracting parties. Moreover, contracts are not

enforced with high sanctions, and breach is not an uncommon event.

There are three reasons for the incompleteness of contracts. The first is the cost of writing more complete contracts. The second is that some variables (effort levels, technical production difficulties) cannot be verified by tribunals. The third is that the expected consequences of incompleteness may not be very harmful to contracting parties. Incompleteness may not be harmful because a tribunal might interpret an imperfect contract in a desirable manner. Also, as will be seen, the prospect of having to pay damages for breach of contract may serve as an implicit substitute for more detailed terms. Furthermore, the opportunity to renegotiate a contract often furnishes a way for parties to alter terms in the light of circumstances for which contractual provisions had not been made.

Interpretation of contracts. Contractual interpretation, which includes a tribunal's filling gaps, resolving ambiguities, and overriding literal language, can benefit parties by easing their drafting burdens or reducing their need to understand contractual detail. For example, if it is efficient to excuse a seller from having to perform if his factory burns down, the parties need not incur the cost of specifying this exception in their contract, assuming that they can trust the tribunal to interpret their contract as if the exception were specified. A method of interpretation can be viewed formally as a function that transforms the contract individuals write into the effective contract that the tribunal will enforce. Given a method of interpretation, parties will choose contracts in a constrained-efficient way. Notably, if the parties are concerned that an aspect of their contract would not be interpreted as they want, they could either bear the cost of writing a more explicit term that would be respected by the tribunal, or they could simply accept the expected loss from having a less-than-efficient term. The socially optimal method of interpretation will take this reaction of contracting parties into account and can be regarded as minimizing the sum of the costs the parties bear in writing contracts and the losses resulting from

inefficient enforcement. (See Ayres and Gertner, 1989; Hadfield, 1994; Schwartz, 1992; and Shavell, 2006.)

Damage measures for breach of contract. When parties breach a contract, they often have to pay damages in consequence. The damage measure, the formula governing what they should pay, can be determined by the tribunal or it can be stipulated in advance by the parties to the contract. One would expect parties to specify their own damage measure when it would better serve their purposes than the measure the tribunal would employ, and otherwise to allow the tribunal to select the damage measure. In either case, we now examine the utility of different damage measures to contracting parties, assuming initially that there is no renegotiation of contracts.

Clearly, the prospect of having to pay damages provides an incentive to perform contractual obligations, and thus generally promotes enforcement of contracts and the goals of the parties. Under the commonly employed *expectation measure*, damages equal the amount that compensates the victim of breach for his losses. Under this measure, a seller contemplating breach will be induced to perform if the cost of performance to the seller is less than the value of performance to the buyer, and to breach otherwise. Because the expectation measure leads to maximization of joint value, it would be chosen by the parties (setting aside investment incentives and risk-bearing), as emphasized in Shavell (1980a). Another commonly employed measure of damages is the *reliance measure*, damages equal to the amount spent by the victim relying on contract performance, such as expenditures on advertising an entertainer who has contracted to appear at one's nightclub.

The point that the expectation measure of damages induces efficient performance of parties sheds light on the view of many legal commentators that breach is immoral. This view fails to account for the fact that contracts that are breached are generally incomplete, and that

breach constitutes behavior that the parties truly want and would have provided for in a complete contract.

Damage measures not only affect performance, they also influence the ex ante motive to make investments in reliance on contract performance. Under the expectation measure, reliance investments tend to exceed efficient levels: the buyer will treat an investment (like advertising an entertainer) as one with a sure payoff, since he either will receive performance or expectation damages, whereas the actual return to the investment is uncertain, due to the possibility of breach (advertising will be a waste if the entertainer does not appear); see Shavell (1980a). This tendency toward overreliance stands in contrast to the problem of inadequate reliance investment associated with lack of contract enforcement.

Damage measures affect risk-bearing as well as incentives. Notably, because the expectation measure compensates the victim of a breach, the measure might be mutually desirable as a form of insurance if the victim is risk averse (Polinsky, 1983). However, the prospect of having to pay damages also constitutes a risk for a party who might commit breach (such as a seller whose costs suddenly rise), and he might be risk averse as well. The latter consideration may lead parties to want to lower damages or to employ damages less frequently by writing more detailed contracts (for instance, the parties could go to the expense of specifying in the contract that a seller can be excused from performance if his costs are unusually high).

Specific performance as a remedy for breach. An alternative to use of a damage measure for breach of contract is specific performance: requiring a party to satisfy his contractual obligation. Specific performance can be accomplished with a sufficiently high threat or by exercise of the state's police powers, such as by a sheriff removing a person from the land that he promised to convey. (Note that if a monetary penalty can be employed to induce performance, then specific performance is equivalent to a damage measure with a high level of damages.)

It is apparent from what has been said about incomplete contracts and damage measures that parties should not want specific performance of many contracts that they write, for they do not wish their incomplete contracts always to be performed. It is therefore not surprising that, in fact, specific performance is not used as the remedy for breach for most contracts for production of goods and for provision of services. Additionally, specific performance might be peculiarly difficult to enforce in these contexts because of problems in monitoring and controlling parties' effort levels and the quality of production.

However, specific performance does have advantages for parties in certain contexts, such as in contracts for the transfer of things that already exist, like land, and specific performance is the usual legal remedy for sellers' breaches of contracts for the sale of land.

Renegotiation of contracts. Parties often have the opportunity to renegotiate their contracts when problems arise. Indeed, the assumption that they will do this has appeal because, having made an initial contract, the parties know of each other's existence and of many particulars of the contractual situation. For this reason, much of the economics literature (as opposed to law and economics literature) on contracts assumes that renegotiation always occurs and that, due to symmetric information between the parties, it always results in efficient performance. Hence, damage measures for breach of contract, or more generally, the mechanisms that the parties stipulate in their contracts, establish the threat points for renegotiation. If properly designed, the mechanisms can foster beneficial incentives to invest ex ante, on the part of both parties. On this extensive literature, see, for example, Rogerson (1984), Hart (1987), Hart and Moore (1988), and Bolton and Dewatripont (2005).

Legal overriding of contracts. A basic rationale for legislative or judicial overriding of contracts is the existence of externalities. Contracts that are likely to harm third parties are often not enforced, including, for example, agreements to commit crimes, price-fixing compacts,

liability insurance policies against fines, and certain sales contracts (such as for machine guns).

Another general rationale for nonenforcement of contracts is to prevent a loss in welfare to one or both of the parties to a contract. This concern may justify nonenforcement when a party is incompetent, lacks relevant information, or is in an emergency situation. The rationale also applies in the context of contract interpretation by tribunals. As noted, contract interpretation may amount to the overriding of a written contractual term, and this practice may promote the welfare of contracting parties by allowing them to save writing costs, given that courts will step in and correct inefficient terms.

Additionally, contracts sometimes are not enforced because they involve the sale of things said to be inalienable, such as human organs, babies, and voting rights. In many of these cases, the inalienability justification for lack of enforcement can be recognized as involving externalities or the welfare of the contracting parties.

4. LITIGATION. We here consider the bringing and adjudication of lawsuits: the choice of a party who has suffered a loss whether to sue; if suit is brought, the choice of the litigants whether to settle with each other or instead go to trial; and the choice of litigants, before or during trial, of how much to spend on litigation.

Suit. As a general matter, a party who has suffered loss, the plaintiff, will sue when the cost of suit c_P is less than the expected benefits from suit. The expected benefits from suit incorporate potential settlements or trial outcomes, but assume for simplicity that if suit is brought, the plaintiff obtains for sure a judgment equal to harm suffered, h . Thus the plaintiff will sue when his litigation cost, c_P , is less than h . (Obviously, if there is only a probability p of winning this amount, a risk-neutral plaintiff would sue when $c_P < ph$; and a risk-averse plaintiff would be less likely to sue.)

The private incentive to sue is fundamentally misaligned with the socially optimal incentive to sue, as emphasized in Shavell (1982b, 1997). The deviation could be in either direction. On one hand, there is a divergence between private and social costs that can lead to socially excessive suit: 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. On the other hand, there is a difference between the private and social benefits of suit that can either lead to a socially inadequate level of suit or reinforce the cost-related tendency toward excessive suit. Specifically, the plaintiff considers his private benefit from suit, the gain he would obtain from prevailing, but not the social benefit, the deterrent effect on the behavior of injurers generally. The private gain could be larger or smaller than the social benefit.

To illustrate, suppose that liability is strict. As stated, victims will sue if and only if $c_P < h$. Let x be the precaution expenditures that injurers will be induced to make if there is suit, q the probability of harm if suit is not brought, and q' the probability of harm if suit is brought. (Thus, q' will be less than q if x is spent on precautions.) Suit will be socially worthwhile if and only if $q'(c_P + c_D + c_S) < (q - q')h - x$, where c_D is the defendant's litigation cost and c_S is the state's cost. In other words, suit is socially worthwhile if the expected litigation costs are less than the deterrence benefits of suit net of the cost of precautions. The condition for victims to sue and the condition for suit to be socially optimal 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 what the victim will receive as a damages award, h ; in contrast, the social benefit is the harm weighted by the reduction in the accident probability, $q - q'$, net of the cost of precautions, x . It is evident, therefore, that victims might sue when suit is not socially desirable, or that victims might not sue even when suit would be socially beneficial.

The main implication of the private and social divergence is that state intervention may

be desirable, either to correct a problem of excessive suit, notably by taxing suit or barring it in some domain, or a problem of inadequate suit, by subsidizing suit in some way. For the state to determine optimal policy, however, requires it to estimate the effects of suit on injurer behavior and weigh them against the social costs of suit.

The importance of the private-social divergence in incentives to sue may be substantial. This is suggested by the high costs of using the legal system; indeed, legal costs may on average actually equal the amounts received by those who sue. Hence, the incentives created by the legal system must be significant to justify its use. Regardless of whether the legal system creates valuable incentives, however, the private motive to bring suit may be great, giving rise to a reason for social intervention. Conversely, in some domains the incentive to sue may be low (say damages per plaintiff are not great) even though the value of deterrence is significant. This might justify the state's encouraging litigation.

Settlement versus trial. Assuming that a suit has been brought, we now consider whether parties will reach a settlement or go to trial. A settlement is a legally enforceable contract, 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 assume that they go to trial, that is, that some tribunal determines the outcome of their case. In fact, the vast majority of cases settle.

One model of the settlement-versus-trial decision presumes that the parties have somehow each come to a belief about the probability of the trial outcome (Posner 2003, ch. 21; Shavell, 2004, ch. 17); let p_P represent the plaintiff's opinion about his probability of the plaintiff prevailing, and let p_D be the defendant's opinion about that same probability. Let w be the amount that would be won (for simplicity assume that they agree about w). Assume also that the parties are risk neutral. The plaintiff's expected gain from trial, net of his litigation costs, is $p_P w$

$-c_P$. The defendant's expected loss from trial, including his litigation costs, is $p_DW + c_D$. Hence, a settlement is possible if and only if $p_{PW} - c_P \leq p_{DW} + c_D$, in which case the settlement amount will be in the settlement range $[p_{PW} - c_P, p_{DW} + c_D]$. Note that if the parties agree on the plaintiff's probability of prevailing, a settlement is feasible. A settlement range does not exist, and therefore trial will occur, if $p_{PW} - p_{DW} > c_P + c_D$. Risk aversion of the parties increases the size of the settlement range and thus, one presumes, makes settlement more likely: if the plaintiff is risk averse, he will be willing to settle for less than $p_{PW} - c_P$; and if the defendant is risk averse, he will be willing to pay more than $p_{DW} + c_D$.

The model just discussed does not explain the origin of the parties' beliefs and does not include a description of rational bargaining between them. Subsequently, standard asymmetric information models of settlement versus litigation were examined (Bebchuk, 1984; Reinganum and Wilde, 1986; Schweizer, 1989; Spier, 1992; Hay and Spier, 1998; and Daughety, 2000). In a simple model of this type, there is one-sided asymmetry of information and the party without private information makes a take-it-or-leave-it settlement proposal. For example, the plaintiff makes a demand x to the defendant, who has private information about the probability p that he will lose at trial. 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 and pay x . The plaintiff chooses x to maximize his expected payoff from settlement or trial. The higher his demand x , the more he will obtain if it is accepted, but the greater the likelihood of rejection and thus of his bearing trial costs. At the optimal demand for the plaintiff, there will generally be a positive probability of trial and also of settlement.

The virtues of such asymmetric information models are twofold. First, they include an explicit account of bargaining and thus of the probability of settlement and the magnitude of the settlement offer or demand. (The outcomes of these models depend, however, on essentially

arbitrary modeling choices, such as who makes the settlement proposal, the informed or the uninformed party.) Second, the models explain differences of opinion that give rise to trial in terms of differences in possession of information. (However, the models do not account for why there should be differences in information, given that the parties have incentives to share information and may be forced to do so through legal discovery.)

The private and social incentive to settle generally diverge for several reasons. First, because the litigants do not bear all of the costs of a trial (such as the salaries of judges and the forgone value of juror time), they save less by settling than society does, which tends to make the private incentive to settle socially inadequate. Second, when there is asymmetric information, parties will fail to settle when the plaintiff's demand turns out to have been too high or the defendant's offer too low. But their desire to obtain from each other a greater share of the benefit from settling does not itself translate into any social benefit. Third, the prospect of settlement may reduce deterrence because defendants gain from settlement.

Litigation expenditures. A plaintiff will continue spending on litigation as long as this raises his expected return from settlement or trial (net of litigation costs), and a defendant will make such expenditures as long as this lowers his expected total outlays. The effects of each litigant's expenditures will generally depend on what the other does, and the two will often be spending to rebut one another.

There are several reasons why the private and social incentives to spend on litigation diverge. First, to the extent that their expenditures simply offset each other, without altering trial or settlement outcomes, the expenditures constitute a social waste. Second, the litigants' trial expenditures may mislead the tribunal rather than enhance the accuracy of the outcome, which has negative social value. Third, even if trial expenditures do improve the accuracy of outcomes, they may not be socially optimal in magnitude, for the parties only consider how their

expenditures influence the litigation outcome, without regard to their influence (if any) on deterrence.

Because private and social incentives to spend on litigation may diverge, it may be beneficial for expenditures to be either curtailed or encouraged. In practice, courts often restrict the legal effort that parties can undertake, for example, by limiting the extent of discovery and the number of testifying experts.

Other topics. A number of other topics that relate to litigation and the legal process have been studied, including the selection of suits for litigation (Priest and Klein, 1984); the accuracy of adjudication (Kaplow, 1994, Png, 1986); “discovery,” that is, mandated disclosure of information during litigation (Shavell, 1989); and the appeals process (Daughety and Reinganum, 2000; Shavell, 1995; and Spitzer and Talley, 2000).

5. PUBLIC LAW ENFORCEMENT AND CRIMINAL LAW. Law enforcement often is the result of the efforts of public agents, such as inspectors, tax auditors, and police. We here discuss certain characteristics of optimal public law enforcement. As noted, this subject was first analyzed by Bentham (1789) and Becker (1968) (for a survey, see Polinsky and Shavell, 2000).

Rationale for public enforcement. A basic question is why there is a need for public enforcement of law in the light of the availability of private suits brought by victims (Becker and Stigler, 1974; Landes and Posner, 1975; Polinsky, 1980). The answer depends importantly on the locus of information about the identity of injurers. When victims of harm naturally possess knowledge of the identity of injurers, allowing private suits for damages will motivate victims to sue and thus harness the information they have for purposes of law enforcement. This may help to explain why the enforcement of contractual obligations and of accident law is primarily private. When victims do not know who caused harm, however, or when finding injurers is

difficult, society tends to rely instead on public investigation and prosecution; this is broadly true of crimes and of many violations of environmental and safety regulations.

Basic framework for analyzing public enforcement. Suppose that if an individual commits a harmful act, he obtains a gain and also faces the risk of being caught and sanctioned. The sanction could be a fine or a prison term. Fines will be treated as socially costless because they are mere transfers of money, whereas imprisonment is socially costly because of the expense of operating prisons and the disutility suffered by those imprisoned (which is not offset by gains to others). The higher is the probability of detecting and sanctioning violators, the more resources the state must devote to enforcement.

We assume that social welfare equals the sum of individuals' expected utilities. If individuals are risk neutral, social welfare can be expressed as the gains individuals obtain from committing their harmful acts, less the harms caused, and less the costs of law enforcement. The enforcement authority's problem is to maximize social welfare by choosing enforcement expenditures, or, equivalently, a probability of detection, the form of sanctions, and their level.

Fines. Suppose that the sanction is a fine and that individuals are risk neutral. Then the optimal level of the fine is maximal, f_M , as emphasized in Becker (1968). If the fine were not maximal, society could save enforcement costs by simultaneously raising the fine and lowering the probability without affecting the level of deterrence. Formally, if $f < f_M$, then raise the fine to f_M and lower the probability from p to $(f/f_M)p$; the expected fine is still pf , so that deterrence is maintained, but expenditures on enforcement are reduced, implying that social welfare rises. Moreover, the optimal probability is such that there is some underdeterrence; in other words, at the optimal p , the expected fine pf_M is less than the harm h . The reason for this result is that if pf_M equals h , behavior will be ideal, in which case decreasing p must be socially beneficial because the individuals thereby induced to commit the harmful act cause no net social losses

(because their gains essentially equal the harm), but reducing p saves enforcement costs.

If individuals are risk averse, the optimal fine may well be below the maximal fine, as stressed in Polinsky and Shavell (1979). This is because the use of a very high fine would impose a substantial risk-bearing cost on individuals who commit harmful acts.

Imprisonment. Now suppose that the sanction is imprisonment and that individuals are risk neutral in imprisonment. Then the optimal imprisonment term is maximal. The reasoning is similar to that employed above with respect to fines: if the imprisonment term were not maximal, it could be raised and the probability of detection lowered so as to keep the expected prison term constant; neither individual behavior nor the costs of imposing imprisonment are affected (because the expected prison term is the same), but enforcement expenditures fall.

If, instead, individuals are risk averse in imprisonment (the disutility of each additional year of imprisonment grows with the number of years in prison), there is a stronger argument for setting the imprisonment sanction maximally than when individuals are risk neutral. Now, when the imprisonment term is raised, the probability of detection can be lowered even more than in the risk-neutral case without reducing deterrence. Thus, not only are there greater savings in enforcement expenditures, but also the social costs of imposing imprisonment sanctions decline because the expected prison term falls.

Last, suppose that individuals are risk preferring in imprisonment (the disutility of each additional year of imprisonment declines with the number of years in prison). This possibility seems particularly important: the first years of imprisonment may create unusually high disutility, due to brutalization of the prisoner or due to the stigma of having been imprisoned at all. In addition, individuals generally have positive time discount rates, which are thought to be especially significant for criminals. In the case of risk-preferring individuals, the optimal prison term may well be less than maximal: if the sentence were raised, the probability that maintains

deterrence could not be lowered proportionally, implying that the expected prison term would rise. Thus, although there would be enforcement cost savings, they might not be great enough to offset the increased sanctioning costs.

Fines versus imprisonment. Fines generally are preferable to prison terms as a means of deterrence, since fines are socially cheaper sanctions to impose (Becker, 1968). Hence, fines should be employed to the greatest extent possible — until a party's wealth is exhausted — before imprisonment is imposed. Further, imprisonment should be used as a sanction only if the harm prevented by the added deterrence is sufficiently great.

Fault-based liability. Our discussion so far has presumed that liability is strict, but liability may also be based on fault, an assessment of whether the act that caused harm was socially undesirable (analogous to the negligence rule and due care standard discussed above in the accident context). Fault-based liability, like strict liability, can induce individuals to behave properly, but fault-based liability possesses an advantage when individuals are risk averse: if they act responsibly, they will not be found at fault, so will not bear the risk of being sanctioned. Similarly, fault-based liability is advantageous when the form of the sanction is imprisonment, for then, again, individuals may be led to behave optimally without the actual imposition of sanctions, and thus without social costs being incurred (Shavell, 1987b). To the extent that mistakes are made in determining fault, however, these two advantages are reduced because risk is imposed and sanctioning costs are incurred. Note, too, that fault-based liability is more difficult to implement, because it requires the state to determine optimal behavior.

Incapacitation. Society may reduce harm not only through deterrence but also by imposing sanctions that remove parties from positions in which they are able to cause harm, that is, by incapacitating them. Imprisonment is the primary incapacitative sanction, although there are other examples: individuals can lose their drivers' licenses; businesses can lose their right to

operate in certain domains, and the like.

Suppose that the sole function of imprisonment is to incapacitate. Then it will be desirable to keep someone in jail as long as the reduction in crime from incapacitating him exceeds the costs of imprisonment (Shavell, 1987c). Although this condition could hold for a long period, it is unlikely to, unless the harm prevented is very high, because the proclivity to commit crimes apparently declines sharply with age.

Note that, as a matter of economic logic, the incapacitation rationale might imply that a person should be imprisoned even if he has not committed a crime — because the danger he poses to society makes incapacitating him worthwhile. In practice, however, the fact that a person has committed a harmful act may be the best basis for predicting his future behavior, in which case the incapacitation rationale would suggest imprisoning an individual only if he has committed such an act.

Two observations are worth noting about the relationship between optimal enforcement when incapacitation is the goal versus when deterrence is the goal. First, when enforcement is based on incapacitation, the optimal magnitude of the sanction is independent of the probability of apprehension, which contrasts with the case when enforcement is based on deterrence. Second, when enforcement is deterrence-oriented, the probability and magnitude of sanctions depend on the ability to deter, and if this ability is limited (as, for instance, with the insane), a low expected sanction may be optimal, whereas a high sanction still might be called for to incapacitate.

Other issues. A number of other topics have been studied in the economic analysis of public law enforcement, including mistake, marginal deterrence (the effect of sanctions in reducing the severity of harm a party causes), self-reporting of violations (Kaplow and Shavell, 1994a; Innes, 1999), repeat offenses, plea bargaining (Reinganum, 1988), general enforcement

(when detection resources simultaneously influence the deterrence of a range of harmful acts) (Mookherjee and Png, 1992; and Shavell, 1991), and corruption of law enforcement agents (Shleifer and Vishny, 1993; Rose-Ackerman, 1999; and Polinsky and Shavell, 2001).

Criminal law. The subject of criminal law may be viewed in the light of the theory of public law enforcement (Posner, 1985; Shavell, 1985). First, the fact that the acts in the core area of crime (robbery, murder, rape, and so forth) are punished by the sanction of imprisonment makes basic sense. Were society to rely on fines alone, deterrence of the acts in question would be grossly inadequate. Notably, the probability of detecting many of these acts is low, making the money sanction necessary for deterrence high, but the assets of individuals who commit these acts often are insubstantial. Hence, the threat of prison is needed for deterrence. Moreover, the incapacitative aspect of imprisonment is valuable because of the difficulty of deterring individuals who are prone to commit criminal acts.

Second, many of the doctrines of criminal law appear to enhance social welfare. This seems true of the basic feature of criminal law that punishment is not imposed on all harmful acts, but instead is usually confined to those that are undesirable. (For example, murder is subject to criminal sanctions, but not all accidental killing.) As we have stressed, when the socially costly sanction of imprisonment is employed, the fault system is desirable because it results in less frequent imposition of punishment than strict liability. Also, the focus on intent in criminal law as a precondition for imposing sanctions may be sensible with regard to deterrence because those who intend to do harm are more likely to conceal their acts, and may be harder to discourage because of the benefits they anticipate. That unsuccessful attempts to do harm are punished in criminal law is an implicit way of raising the likelihood of sanctions for undesirable acts. Study of specific doctrines of criminal law seems to afford a rich opportunity for economic analysis.

6. CRITICISM OF ECONOMIC ANALYSIS OF LAW. Many observers, and particularly noneconomists, view economic analysis of law with skepticism. We consider several such criticisms here.

Description of behavior. It is sometimes claimed that individuals and firms do not respond to legal rules as rational maximizers of their well-being. For example, it is often asserted that decisions to commit crimes are not governed by economists' usual assumptions. Some skeptics also suggest that, in predicting individuals' behavior, certain standard assumptions are inapplicable. For example, in predicting compliance with a law, the assumption that preferences be taken as given would be inappropriate if a legal rule would change people's preferences, as some say was the case with civil rights laws and environmental laws. In addition, laws may frame individuals' understanding of problems, which could affect their probability assessments or willingness to pay. The emerging field of behavioral economics, as well as work in various disciplines that address social norms, is beginning to examine these sorts of issues (Jolls et al., 1998).

Distribution of income. A frequent criticism of economic analysis of law concerns its focus on efficiency, to the exclusion of the distribution of income. The claim of critics is that legal rules should be selected in a manner that takes into account their effects on the rich and the poor. But achieving sought-after redistribution through income tax and transfer programs tends to be superior to redistribution through the choice of legal rules. This is because redistribution through legal rules and the tax-transfer system both will distort individuals' labor-leisure decisions in the same manner, but redistribution through legal rules often will require choosing an inefficient rule, which imposes an additional cost (Shavell, 1981; Kaplow and Shavell, 1994b).

Moreover, it is difficult to redistribute income systematically through the choice of legal rules: many individuals are never involved in litigation; and for those who are, there is substantial income heterogeneity among plaintiffs as well as among defendants. Additionally, in contractual contexts, the choice of a legal rule often will not have any distributional effect because contract terms, notably, the price, will adjust so that any agreement into which parties enter will continue to reflect the initial distribution of bargaining power between them.

Concerns for fairness. An additional criticism is that the conventional economic approach slights important concerns about fairness, justice, and rights. Some of these notions refer implicitly to the appropriateness of the distribution of income and, accordingly, are encompassed by our preceding remarks. Also, to some degree, the notions are motivated by instrumental concerns. For example, the attraction of paying fair compensation to victims must derive in part from the beneficial risk reduction effected by such payments, and the appeal of obeying contractual promises must rest in part on the desirable consequences contract performance has on production and exchange. To some extent, therefore, critics' concerns are already taken into account in standard economic analysis.

However, many who promote fairness, justice, and rights do not regard these notions merely as some sort of proxy for attaining instrumental objectives. Instead, they believe that satisfying these notions is intrinsically valuable. This view also can be partially reconciled with the economic conception of social welfare: if individuals have a preference for a legal rule or institution because they regard it as fair, that should be credited in the determination of social welfare, just as any preference should.

But many commentators take the position that conceptions of fairness are important as ethical principles in themselves, without regard to any possible relationship the principles may have to individuals' welfare. This opinion is the subject of longstanding debate among moral

philosophers. Some readers may be skeptical of normative views that are not grounded in individuals' well-being because embracing such views entails a willingness to sacrifice individuals' well-being. Indeed, consistently pursuing any non-welfarist principle must sometimes result in everyone being made worse off; see Kaplow and Shavell (2001, 2002).

Efficiency of judge-made law. Also criticized is the contention of some economically-oriented legal academics, notably Posner (1972), that judge-made law tends to be efficient (in contrast to legislation, which is said to reflect the influence of special interest groups). Some critics believe that judge-made law is guided by notions of fairness, or is influenced by legal culture or judges' biases, and thus will not necessarily be efficient. Whatever is the merit of the critics' claims, they are descriptive assertions about the law, and their validity does not bear on the power of economics to predict behavior in response to legal rules or on the value of normative economic analysis of law.

References

- Ayres, I., and Gertner, R. 1989. Filling gaps in incomplete contracts: an economic theory of default rules. *Yale Law Journal* 99, 87-130.
- Bebchuk, L.A. 1984. Litigation and settlement under imperfect information. *RAND Journal of Economics* 15, 404-415.
- Bebchuk, L.A., and Ben-Shahar, O. 2001. Pre-contractual reliance. *Journal of Legal Studies* 30, 423-457.
- Becker, G.S. 1968. Crime and punishment: an economic approach. *Journal of Political Economy* 76, 169-217.
- Becker, G.S., and Stigler, G.J. 1974. Law enforcement, malfeasance, and compensation of enforcers. *Journal of Legal Studies* 3, 1-18.
- Bentham, J. 1789. *An Introduction to the Principles of Morals and Legislation*, in: *The Utilitarians*, 1973 ed. Garden City, NY: Anchor Books, 5-398.
- Blume, L., Rubinfeld, D.L., and Shapiro, P. 1984. The taking of land: when should compensation be paid? *Quarterly Journal of Economics* 99, 71-92.
- Bolton, P., and Dewatripont, M. 2005. *Contract Theory*. Cambridge, MA: MIT Press.
- Brown, J.P. 1973. Toward an economic theory of liability. *Journal of Legal Studies* 2, 323-349.
- Calabresi, G. 1970. *The Costs of Accidents: A Legal and Economic Analysis*. New Haven: Yale University Press.
- Calabresi, G., and Melamed, A.D. 1972. Property rules, liability rules, and inalienability: one view of the cathedral. *Harvard Law Review* 85, 1089-1128.
- Coase, R.H. 1960. The problem of social cost. *Journal of Law and Economics* 3, 1-44.
- Cooter, R.D., and Ulen, T. 2003. *Law and Economics*, 4th ed. Reading, MA: Addison-Wesley.
- Craswell, R. 1996. Offer, acceptance, and efficient reliance. *Stanford Law Review* 48, 481-553.

- Daughety, A.F. 2000. Settlement. *Encyclopedia of Law and Economics*, Bouckaert, B., and De Geest, G., eds. Cheltenham, UK: Edward Elgar. 5, 95-158.
- Daughety, A.F., and Reinganum, J.F. 2000. Appealing judgments. *RAND Journal of Economics* 31, 502-525.
- Demsetz, H. 1967. Toward a theory of property rights. *American Economic Review: Papers and Proceedings* 57, 347-359.
- Diamond, P.A. 1974. Single activity accidents. *Journal of Legal Studies* 3, 107-164.
- Diamond, P.A., and Maskin, E. 1979. An equilibrium analysis of search and breach of contract, I: steady states. *Bell Journal of Economics* 10, 282-316.
- Hadfield, G.K. 1994. Judicial competence and the interpretation of incomplete contracts. *Journal of Legal Studies* 23, 159-184.
- Hart, O.D. 1987. Incomplete contracts. *The New Palgrave Dictionary of Economics*, Eatwell, J., Milgate, M., and Newman, P., eds. New York: Macmillan Press. 2, 752-759.
- Hart, O.D., and Moore, J. 1988. Incomplete contracts and renegotiation. *Econometrica* 56, 755-758.
- Hay, B.L., and Spier, K.E. 1998. Settlement of litigation. *The New Palgrave Dictionary of Economics and the Law*, Newman, P., ed. London: Macmillan. 3, 442-451.
- Innes, R. 1999. Remediation and self-reporting in optimal law enforcement. *Journal of Public Economics* 72, 379-393.
- Jolls, C., Sunstein, C.R. and Thaler, R. 1998. A behavioral approach to law and economics. *Stanford Law Review* 50, 1471-1550.
- Kaplow, L. 1986. An economic analysis of legal transitions. *Harvard Law Review* 99, 509-617.
- Kaplow, L. 1994. The value of accuracy in adjudication: an economic analysis. *Journal of Legal Studies* 23, 307-401.

- Kaplow, L. and Shavell, S. 1994a. Optimal law enforcement with self-reporting of behavior. *Journal of Political Economy* 102, 583-606.
- Kaplow, L. and Shavell, S. 1994b. Why the legal system is less efficient than the income tax in redistributing income. *Journal of Legal Studies* 23, 667-681.
- Kaplow, L., and Shavell, S. 2001. Any non-welfarist method of policy assessment violates the Pareto principle. *Journal of Political Economy* 109, 281-286.
- Kaplow, L., and Shavell, S. 2002. *Fairness versus Welfare*. Cambridge, MA: Harvard University Press.
- Katz, A. 1990. The strategic structure of offer and acceptance: game theory and the law of contract formation. *Michigan Law Review* 89, 215-295.
- Katz, A. 1996. When should an offer stick? The economics of promissory estoppel in preliminary negotiations. *Yale Law Journal* 105, 1249-1309.
- Kronman, A.T. 1978. Mistake, disclosure, information, and the law of contracts. *Journal of Legal Studies* 7, 1-34.
- Landes, W.M., and Posner, R.A. 1975. The private enforcement of law. *Journal of Legal Studies* 4, 1-46.
- Landes, W.M., and Posner, R.A. 1987a. *The Economic Structure of Tort Law*. Cambridge, MA: Harvard University Press.
- Landes, W.M., and Posner, R.A. 1987b. Trademark law: an economic perspective. *Journal of Law and Economics* 30, 265-309.
- Landes, W.M., and Posner, R.A. 2003. *The Economic Structure of Intellectual Property Law*. Cambridge, MA: Harvard University Press.
- Libecap, G.D. 1986. Property rights in economic history: implications for research. *Explorations in Economic History* 23, 227-252.

- Miceli, T.J. 1997. *Economics of the Law: Torts, Contracts, Property, Litigation*. New York: Oxford University Press.
- Mookherjee, D., and Png, I.P.L. 1992. Monitoring vis-à-vis investigation in enforcement of law. *American Economic Review* 82, 556-565.
- Png, I.P.L. 1986. Optimal subsidies and damages in the presence of judicial error. *International Review of Law and Economics* 6, 101-105.
- Polinsky, A.M. 1980. Private versus public enforcement of fines. *Journal of Legal Studies* 9, 105-127.
- Polinsky, A.M. 1983. Risk sharing through breach of contract remedies. *Journal of Legal Studies* 12, 427-444.
- Polinsky, A.M., and Shavell, S. 1979. The optimal tradeoff between the probability and magnitude of fines. *American Economic Review* 69, 880-891.
- Polinsky, A.M., and Shavell, S. 2000. The economic theory of public enforcement of law. *Journal of Economic Literature* 38, 45-76.
- Polinsky, A.M., and Shavell, S. 2001. Corruption and optimal law enforcement. *Journal of Public Economics* 81, 1-24.
- Polinsky, A.M., and Shavell, S., eds. 2006. *Handbook of Law and Economics*, vol. 1. Amsterdam: Elsevier.
- Posner, R.A. 1972. *Economic Analysis of Law*. Boston: Little, Brown and Company.
- Posner, R.A. 1985. An economic theory of the criminal law. *Columbia Law Review* 85, 1193-1231.
- Posner, R.A. 2003. *Economic Analysis of Law*, 6th ed. New York: Aspen Publishers.
- Priest, G.L. and B. Klein. 1984. The selection of disputes for litigation. *Journal of Legal Studies* 13, 1-55.

- Reinganum, J.F. 1988. Plea bargaining and prosecutorial discretion, *American Economic Review* 78, 713-728.
- Reinganum, J.F., and Wilde, L.L. 1986. Settlement, litigation, and the allocation of litigation costs. *RAND Journal of Economics* 17, 557-566.
- Rogerson, W.P. 1984. Efficient reliance and damage measures for breach of contract. *RAND Journal of Economics* 15, 39-53.
- Rose-Ackerman, S. 1999. *Corruption and Government: Causes, Consequences and Reform*. New York: Cambridge University Press.
- Schwartz, A. 1992. Relational contracts in the courts: an analysis of incomplete agreements and judicial strategies. *Journal of Legal Studies* 21, 271-318.
- Schweizer, U. 1989. Litigation and settlement under two-sided incomplete information. *Review of Economic Studies* 56, 163-178.
- Shavell, S. 1980a. Damage measures for breach of contract. *Bell Journal of Economics* 11, 466-490.
- Shavell, S. 1980b. Strict liability versus negligence. *Journal of Legal Studies* 9, 1-25.
- Shavell, S. 1981. A note on efficiency vs. distributional equity in legal rulemaking: should distributional equity matter given optimal income taxation? *American Economic Review: Papers and Proceedings* 71, 414-418.
- Shavell, S. 1982a. On liability and insurance. *Bell Journal of Economics* 13, 120-132.
- Shavell, S. 1982b. The social versus the private incentive to bring suit in a costly legal system. *Journal of Legal Studies* 11, 333-339.
- Shavell, S. 1985. Criminal law and the optimal use of nonmonetary sanctions as a deterrent. *Columbia Law Review* 85, 1232-1262.
- Shavell, S. 1987a. *Economic Analysis of Accident Law*. Cambridge, MA: Harvard University

- Press.
- Shavell, S. 1987b. The optimal use of nonmonetary sanctions as a deterrent. *American Economic Review* 77, 584-592.
- Shavell, S. 1987c. A model of optimal incapacitation. *American Economic Review: Papers and Proceedings* 77, 107-110.
- Shavell, S. 1989. Sharing of information prior to settlement or litigation. *RAND Journal of Economics* 20, 183-195.
- Shavell, S. 1991. Specific versus general enforcement of law. *Journal of Political Economy* 99, 1088-1108.
- Shavell, S. 1993. The optimal structure of law enforcement. *Journal of Law and Economics* 36, 255-287.
- Shavell, S. 1994. Acquisition and disclosure of information prior to sale. *RAND Journal of Economics* 25, 20-36.
- Shavell, S. 1995. The appeals process as a means of error correction. *Journal of Legal Studies* 24, 379-426.
- Shavell, S. 1997. The fundamental divergence between the private and the social motive to use the legal system. *Journal of Legal Studies* 26, 575-612.
- Shavell, S. 2004. *Foundations of Economic Analysis of Law*. Cambridge, MA: Harvard University Press.
- Shavell, S. 2006. On the writing and interpretation of contracts. *Journal of Law, Economics, & Organization*, forthcoming.
- Shleifer, A., and Vishny, R.W. 1993. Corruption. *Quarterly Journal of Economics* 108, 599-617.
- Spier, K.E. 1992. The dynamics of pretrial negotiation. *Review of Economic Studies* 59, 93-108.
- Spitzer, M., and Talley, E. 2000. Judicial auditing. *Journal of Legal Studies* 29, 649-683.

Umbeck, J.R. 1981. *A Theory of Property Rights with Application to the California Gold Rush*.

Ames, Iowa: Iowa State University Press.

Wils, W.P.J. 1993. Who should bear the costs of failed negotiations? A functional inquiry into precontractual liability. *Journal des Economistes et des Etudes Humaines* 4, 93-134.