

## LIABILITY FOR ACCIDENTS

STEVEN SHAVELL\*

*School of Law, Harvard University, and National Bureau of Economic Research*

### Contents

|   |     |
|---|-----|
| 1. Introduction   | 142 |
| Part A: Central theory of liability                                       | 142 |
| 2. Incentives   | 143 |
| 2.1. Unilateral accidents and levels of care                              | 143 |
| 2.2. Bilateral accidents and levels of care                               | 144 |
| 2.3. Unilateral accidents, levels of care, and levels of activity         | 146 |
| 2.4. Bilateral accidents, levels of care, and levels of activity          | 147 |
| 3. Risk-bearing and insurance   | 147 |
| 3.1. First-best solution to the accident problem                          | 148 |
| 3.2. The accident problem given liability but in the absence of insurance | 148 |
| 3.3. The accident problem given liability in the presence of insurance    | 149 |
| 4. Administrative costs   | 151 |
| 4.1. Volume of suit   | 151 |
| 4.2. Level of care and volume of suit                                     | 153 |
| 4.3. Comments   | 154 |
| 4.4. Strict liability and negligence                                      | 155 |
| Part B: Extensions  | 156 |
| 5. Liability of firms   | 156 |
| 5.1. Victims are strangers  | 156 |
| 5.2. Victims are consumers  | 157 |
| 6. Aspects of the negligence determination                                | 158 |
| 6.1. Differences among injurers   | 159 |
| 6.2. Imperfect assessment of care by courts                               | 159 |
| 6.3. Imperfect ability to control care by injurers                        | 160 |
| 6.4. Errors in the calculation of due care                                | 160 |
| 7. Causation  | 161 |
| 7.1. Strict liability   | 161 |

\* Research support from the John M. Olin Center for Law, Economics, and Business at Harvard Law School is gratefully acknowledged.

*Handbook of Law and Economics, Volume 1*

*Edited by A. Mitchell Polinsky and Steven Shavell*

© 2007 Elsevier B.V. All rights reserved

DOI: 10.1016/S1574-0730(07)01002-X

|  |     |
|--|-----|
| 7.2. Negligence rule   | 162 |
| 7.3. Uncertainty over causation  | 162 |
| 7.4. Proximate causation   | 163 |
| 8. The magnitude of liability: damages   | 164 |
| 8.1. Basic theory: strict liability  | 164 |
| 8.2. Basic theory: negligence rule   | 165 |
| 8.3. Difficulty in estimating harm   | 166 |
| 8.4. Punitive damages  | 166 |
| 8.5. Accuracy of damages and legal costs   | 167 |
| 9. Multiple injurers   | 167 |
| 9.1. Strict liability  | 168 |
| 9.2. Negligence rule   | 168 |
| 10. The judgment-proof problem   | 168 |
| 10.1. Incentives to engage in harmful activities   | 169 |
| 10.2. Incentives to take care  | 169 |
| 10.3. Incentives to purchase liability insurance   | 170 |
| 10.4. Policies to ameliorate the judgment-proof problem                                    | 170 |
| 11. Vicarious liability  | 171 |
| 12. Nonpecuniary harm  | 173 |
| 12.1. Incentives in the risk-neutral case  | 173 |
| 12.2. Incentives and insurance in the risk-averse case; fines as a supplement to liability | 173 |
| Part C: Methods of controlling risk  | 175 |
| 13. Liability versus other methods of controlling risk                                     | 175 |
| 13.1. Methods of control   | 175 |
| 13.2. Comparison   | 175 |
| 13.3. Resolution of externalities through bargaining by affected parties                   | 178 |
| 14. Conclusion   | 179 |
| References   | 179 |

## Abstract

This is a survey of legal liability for accidents. Three general aspects of accident liability are addressed. The first is the effect of liability on incentives, both whether to engage in activities (for instance, whether to drive) and how much care to exercise (at what speed to travel) to reduce risk when so doing. The second general aspect concerns risk-bearing and insurance, for the liability system acts as an implicit insurer for accident victims and it imposes risk on potential injurers (because they may have to pay judgments to victims). In this regard, victims' accident insurance and injurers' liability insurance are taken into account. The third general aspect of accident liability is its administrative expense, comprising the cost of legal services, the value of litigants' time, and the operating cost of the courts. A range of subtopics is considered, including product liability, causation, punitive damages, the judgment-proof problem, vicarious

liability, and nonpecuniary harm. Liability is also compared to other methods of controlling harmful activities, notably, to corrective taxation and to regulation.

**Keywords**

law and economics, liability, accident(s), tort(s), insurance, product liability, externalities, information, regulation, litigation, suit, trial, court(s)

*JEL classification:* D00, D6, D8, H1, H8, K00, K13, L5

## 1. Introduction

The subject of this chapter is liability for accidents, by which is meant the law determining when the victim of an accident is entitled to recover losses from the injurer. This body of law, included in what is known as tort law, governs, for example, when the victim of an automobile accident can collect from the driver who harmed him, when the victim of pollution can secure compensation from the polluter, or when the victim of an adverse reaction to a drug can obtain a judgment from its manufacturer.

Three aspects of accident liability will be addressed. The first is its effect on incentives, both whether to engage in activities (for instance, whether to drive) and how much care to exercise (at what speed to travel) to reduce risk when so doing. The second aspect concerns risk-bearing and insurance, for the liability system acts as an implicit insurer for victims and it imposes risk on potential injurers (because they may have to pay victims). In this regard, victims' accident insurance and injurers' liability insurance will be taken into account. The third aspect of accident liability is its administrative expense, comprising the cost of legal services, the value of litigants' time, and the operating cost of the courts.<sup>1</sup>

The chapter will begin in [Part A](#) with the central theory of accident liability, where the main points about incentives, risk-bearing, and administrative costs will be presented. Then in [Part B](#) a variety of subsidiary topics and issues will be discussed, including liability of firms, the judgment-proof problem (inability to pay fully for harm done), causation, nonmonetary losses, and vicarious liability. The sections in [Part B](#) can be read more or less independently of one another; they require only an understanding of [Part A](#). Last, [Part C](#) will compare liability to other methods of controlling harmful activities, such as corrective (Pigouvian) taxes and regulation.

Economic analysis of accident liability began with mainly informal contributions of a number of legal scholars, notably, [Calabresi \(1970\)](#) and [Posner \(1972\)](#), and has been developed since then, in large part by economists, using the standard methods of microeconomics.

Before proceeding, it may be remarked that in view of the importance of liability in reality (its virtual omnipresence) and its appeal in theory relative to other means of controlling harmful externalities (addressed in [Part C](#)), one has the sense that it deserves to receive greater attention from economists than it has to this point.

### Part A: Central theory of liability

In this part on the principal theory of liability, the two basic rules of liability will be considered, strict liability and the negligence rule. Under strict liability, an injurer must always pay for harm due to an accident that he causes. Under the negligence rule, an

<sup>1</sup> This chapter is mainly theoretical; for empirical work on liability, see [Kessler and Rubinfeld \(2007\)](#).

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, as will be discussed below.) In fact, the negligence rule is the dominant form of liability; strict liability is reserved mainly for certain especially dangerous activities (such as the use of explosives).<sup>2</sup> The amount paid by a liable party under a liability rule is often referred to as “damages,” and unless otherwise noted, damages will be assumed to equal the harm caused in an accident.

## 2. Incentives

In order to focus on liability and incentives, it is assumed in this section that victims and injurers are risk neutral. Further, it is assumed that they are strangers to one another, or at least are not in a contractual relationship. Additionally, it is assumed for simplicity that victims and injurers are individuals as opposed to firms (although most of what is said carries over to the context of firms; see section 5).

To begin with, accidents are assumed to be unilateral in nature: only injurers can influence risks. Then bilateral accidents are considered; in these accidents victims as well as injurers affect risks. As noted above, two types of decisions of parties are examined: concerning their level of care (or precautions) when engaging in an activity; and concerning their level of activity. First, the choice of care alone will be studied; then both care and activity level will be investigated.

### 2.1. Unilateral accidents and levels of care

Here the assumption is that injurers alone can reduce risk by choosing a level of care. Let  $x$  be expenditures on care (or the value of effort devoted to it) and  $p(x)$  be the probability of an accident that causes harm  $h$ , where  $p$  is declining and convex 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 are required to pay damages equal to  $h$  whenever an accident occurs, and they 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, denoted  $\underline{x}$ , is set equal to  $x^*$ . Hence, under the rule an injurer who causes harm will be found negligent and have to pay  $h$  if  $x < x^*$  but will not be found negligent and will not have to pay anything if  $x \geq x^*$ . It follows that the injurer will choose  $x^*$ : Clearly, he will not choose  $x$  greater than  $x^*$ , for that will cost him more and he will escape liability for negligence by choosing merely  $x^*$ . Moreover, he will not choose  $x < x^*$ , for then he would be

<sup>2</sup> Dobbs (2000).

liable, implying that his expenses would exceed  $x^* + p(x^*)h$ , which is greater than or equal to  $x^*$ .

Thus, under both forms of liability, strict liability and the negligence rule, injurers are induced to take optimal care. These fundamental results were first shown by Brown (1973). Several comments may be made about them.

(a) The informational requirements imposed on the courts are different under the two rules. Under strict liability, the courts need only to observe the harm  $h$ . In contrast, under the negligence rule, the courts need to know more: they must calculate optimal care  $x^*$ , observe actual care  $x$ , and also observe harm  $h$ . (However, the informational advantage of strict liability is attenuated when strict liability with a defense of contributory negligence is considered in the bilateral case below.)

(b) The results about optimal care under the two rules (as well as most others in section 2) hold more generally than in the simple model considered above. They are valid if  $x$  is multidimensional and if  $x$  affects the probability distribution of harm (rather than just a single level of harm), as the reader can easily verify by essentially the argument given. Note that if a component of  $x$  is not observable by the court, the component cannot be included in the due care standard, so that the component would not be selected optimally under the negligence rule but would be under strict liability.

(c) No actual findings of liability occur under the negligence rule when due care is optimal,  $\underline{x} = x^*$ , since injurers are induced to take due care; but findings of liability do occur under strict liability (presuming that  $p(x^*) > 0$ ).

## 2.2. *Bilateral accidents and levels of care*

Assume now that victims also choose a level of care  $y$ , that the probability of an accident is  $p(x, y)$ , which is declining in both variables, that the social goal is to minimize  $x + y + p(x, y)h$ , and that the optimal levels of care  $x^*$  and  $y^*$  are positive and unique.<sup>3</sup>

Under strict liability, it is evident that injurers' incentives are optimal conditional on victims' level of care, but victims choose  $y = 0$ ; victims have no incentive to take care because they are fully compensated for their losses.

However, the natural version of the strict liability rule to consider in bilateral situations is strict liability with a defense of contributory negligence. Under this rule, a due care level  $\underline{y}$  is established by courts for victims, and a victim is said to be contributorily negligent if his care level  $y$  was less than  $\underline{y}$ . An injurer is held liable for harm only if the victim was not contributorily negligent, that is, only if the victim's level of care was at least his due care level  $\underline{y}$ . If  $\underline{y}$  is set by the courts to equal  $y^*$ , then it is a (Nash) equilibrium for both injurers and victims to act optimally. In particular,

<sup>3</sup> In some early, less formal literature on accidents, for example, Calabresi (1970), reference is made to the notion of the "least-cost avoider," the party—injurer or victim—who can avoid an accident at the lower cost. The idea of a least-cost avoider relies on the assumption that either party can undertake a discrete amount of care that is by itself sufficient to prevent an accident. Under this assumption, it is socially best for only the least cost-avoider to take care.

victims will choose  $y^*$  in order to avoid having to bear their losses, assuming that injurers choose  $x^*$ : victims obviously would not choose  $y > y^*$ . Also, victims would not choose  $y < y^*$ , for if they did so, they would bear their own losses (they would be contributorily negligent), so would minimize  $y + p(x^*, y)h$ ; but for any  $y < y^*$ , this must exceed  $y^* + p(x^*, y^*)h \geq y^*$ . Conversely, injurers will choose  $x^*$ , assuming that victims choose  $y^*$ . For then injurers will have to pay for harm, so they will minimize  $x + p(x, y^*)$ .<sup>4</sup>

Under the negligence rule, if due care  $\underline{x}$  is set equal to  $x^*$ , then injurers and victims will also act optimally in equilibrium. Injurers will choose  $x^*$  to avoid being liable, assuming that victims choose  $y^*$ . This is true by essentially the argument showing that injurers choose  $x^*$  under the negligence rule in the unilateral case. Victims will choose  $y^*$ , assuming the injurers choose  $x^*$ . This is so because victims will bear their losses (since injurers behave nonnegligently), meaning that they will select  $y$  to minimize  $y + p(x^*, y)h$ , implying that they will choose  $y^*$ .

A variant of the negligence rule is negligence with the defense of contributory negligence, according to which a negligent injurer is held liable only if the victim was not contributorily negligent. If due care levels are chosen optimally,  $\underline{x} = x^*$  and  $\underline{y} = y^*$ , then injurers and victims will act optimally in equilibrium under this rule. The explanation is similar to that just given for the negligence rule.

Another version of the negligence rule is the comparative negligence rule. By definition of this rule, if both parties are negligent, they each bear a fraction of the harm, the fraction rising the lower their respective levels of care; if only one party is negligent, however, that party pays for the entire harm; and if neither is negligent, the victim bears his losses. Again, if  $\underline{x} = x^*$  and  $\underline{y} = y^*$ , then injurers and victims act optimally in equilibrium under this rule. The explanation is identical to that under the negligence rule with the defense of contributory negligence. That the comparative negligence rule differs from the negligence rule when both parties are negligent is a moot aspect of the rule, since in equilibrium, the circumstance in which both parties are negligent is irrelevant to the calculations of either party.<sup>5</sup>

In summary, strict liability with the defense of contributory negligence and all of the versions of the negligence rules support optimal levels of care  $x^*$  and  $y^*$  in equilibrium, assuming that due care levels are chosen optimally. These conclusions were also established by Brown (1973).<sup>6</sup>

<sup>4</sup> It can also be shown that  $x^*$  and  $y^*$  is the unique equilibrium (when  $\underline{y} = y^*$ ). Under the other rules to be discussed in this section, the equilibria are also unique.

<sup>5</sup> In versions of the model of accidents in which both injurers and victims might be found negligent in equilibrium, the comparative negligence rule obviously leads to different results from other versions of the negligence rule. See Bar-Gill and Ben-Shahar (2003), Edlin (1994), Cooter and Ulen (1986), and Rubinfeld (1987).

<sup>6</sup> Diamond (1974) proved closely related results shortly afterward. See also Green (1976), Emons (1990), and Emons and Sobel (1991), who focus on the case of heterogeneous injurers and victims.

It should be noted that courts need to be able to calculate optimal care levels for at least one party under any of the rules, and in general this requires knowledge of the function  $p(x, y)$ .

### 2.3. Unilateral accidents, levels of care, and levels of activity

Now let us reconsider unilateral accidents, allowing for injurers to choose their level of activity  $z$ , which is interpreted as the (continuously variable) number of times they engage in their activity. Let  $b(z)$  be the injurer's benefit from the activity, where  $b$  is increasing and concave in  $z$ . Assume that  $x + p(x)h$  is the cost of care incurred and the expected harm generated each time that an injurer engages in his activity, so that  $z(x + p(x)h)$  is the total cost of care and expected harm given  $z$ .<sup>7</sup> Suppose that the social object is to maximize  $b(z) - z(x + p(x)h)$ , and let  $x^*$  and  $z^*$  be optimal values of  $x$  and  $z$ . Note that  $x^*$  minimizes  $x + p(x)h$ , so  $x^*$  is as described above in section 2.1. Thus,  $z^*$  is determined by  $b'(z) = x^* + p(x^*)h$ , which is to say, the marginal benefit from the activity equals the marginal social cost, comprising the sum of the cost of optimal care and expected accident losses (given optimal care).

Under strict liability, an injurer will choose both the optimal level of care  $x^*$  and the optimal level of activity  $z^*$ , as his objective is 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 in section 2.1, but his level of activity  $z$  will be socially excessive. In particular, because an injurer will be led to escape liability for negligence by taking care of  $x^*$ , he will choose  $z$  to maximize  $b(z) - zx^*$ , so that  $z$  will satisfy  $b'(z) = x^*$ . Hence, if  $p(x^*)h$  is positive, then  $x^* < x^* + p(x^*)h$ , so that concavity of  $b$  implies that  $z > z^*$ . The explanation for the excessive level of activity is that the injurer's cost of raising his level of activity is only his cost of care  $x^*$ , which is less than the social cost, as that also includes  $p(x^*)h$ . The excessive level of activity under the negligence rule will be more important the larger is expected harm  $p(x^*)h$  from the activity.

The distinction between activity level and care level, and the result that under strict liability, both are chosen optimally, whereas under the negligence rule, the level of activity is excessive, is first developed in Shavell (1980b). Several comments about the conclusions about should be made.

(a) The failure of the negligence rule to control adequately the level of activity arises because negligence is defined here (and for the most part in reality) in terms of care  $x$  alone. A possible justification for this restriction in the definition of appropriate behavior is the difficulty courts would face in determining the optimal  $z^*$  and the actual  $z$ .

(b) The problem that the activity level is not properly controlled under the negligence rule has an analogue in respect to any component of behavior that would be difficult

<sup>7</sup> If the cost of care and expected harm do not rise linearly with  $z$ , the basic nature of the conclusions of this section would not be altered.



to incorporate directly into the negligence due care standard (either because it cannot readily be observed (consider the frequency with which a driver checks his rear view mirror) or because it is not easy to calculate what constitutes the optimal setting of the component). Any such component of behavior will, however, be optimally controlled under strict liability.

#### 2.4. Bilateral accidents, levels of care, and levels of activity

Suppose now that victims as well as injurers choose levels of care and of activity; let victims' level of activity be denoted  $t$  (victims' level of activity might be how many miles a pedestrian walks and exposes himself to risk) and victims' utility from it be  $v(t)$ , where  $v$  is increasing and concave in  $t$ . Suppose that social welfare is  $b(z) + v(t) - zx - ty - ztp(x, y)h$ , where, note, the assumption continues to be that expected accident losses rise linearly with (now victims' as well as injurers') level of activity. In this general situation, none of the liability rules that have been considered leads to full optimality. As just explained in section 2.3, the negligence rule leads injurers to engage excessively in their activity. Similarly, strict liability with a defense of contributory negligence leads victims to engage excessively in their activity (the number of miles pedestrians walk), as they do not bear their losses given that they take due care. Which form of liability, negligence or strict liability (with the defense of contributory negligence), is better will implicitly reflect whether it is more important to control victims' or injurers' level of activity; if injurers' level of activity is more important to control, strict liability will be superior, otherwise the negligence rule will be preferred.

The reason that full optimality cannot be achieved under either of the major types of liability rule is in essence that injurers must bear accident losses to induce them to choose the right level of their activity, but this means that victims will not choose the optimal level of their activity, and conversely. Indeed, for essentially this reason, it can be shown that there does not exist any liability rule that induces optimal behavior,  $x^*$ ,  $y^*$ ,  $s^*$ , and  $t^*$ , assuming that the liability rule is a function only of care levels  $x$  and  $y$ .<sup>8</sup>

### 3. Risk-bearing and insurance

Let us now consider the implications of risk aversion and the role of accident and liability insurance in relation to accident liability. For this purpose, it is convenient to consider the simple unilateral setting with injurers' level of care the only aspect of behavior at issue; how what is said will carry over to the more general context will be clear to the reader. Let  $U$  and  $V$  be the utility functions of injurers and victims respectively, assume that injurers and victims are either risk neutral or risk averse, and let  $u$

<sup>8</sup> This result is shown in Shavell (1980b). However, fully optimal behavior can readily be induced with tools other than liability rules. For example, if injurers have to pay the state for harm caused and victims bear their own losses, both victims and injurers will choose levels of care and of activity optimally.

and  $v$  be their initial levels of wealth. The development below largely follows Shavell (1982a, 1987), which first studied liability, risk-bearing, and insurance.

### 3.1. First-best solution to the accident problem

The socially ideal solution to the accident problem is such that risk-averse parties do not bear risk and that the level of care is the optimal one,  $x^*$ , discussed in section 2.1.

In particular, the socially ideal solution to the accident problem can be identified with the solution that would be obtained by a dictator who could choose in a Pareto optimal way the level of care  $x$  and also levels of wealth contingent on the occurrence of accidents, subject to a resource constraint.<sup>9</sup> Denoting by  $v_n$  the wealth of a victim if he is not involved in an accident,  $v_a$  his wealth if he is, and similarly for  $u_n$  and  $u_a$ , the dictator would maximize

$$EV = (1 - p(x))V(v_n) + p(x)V(v_a)$$

subject to

$$EU = (1 - p(x))U(u_n) + p(x)U(u_a) = k,$$

where  $k$  is a constant, and subject also to

$$((1 - p(x))v_n + p(x)v_a) + ((1 - p(x))u_n + p(x)u_a) + p(x)h + x = u + v.$$

The second constraint is the resource constraint in expected value terms.<sup>10</sup> It is readily shown that when this problem is solved, (a) risk averse parties—be they injurers or victims—are left with the same level of wealth, and (b) the level of care is  $x^*$ , that minimizing  $p(x)h + x$ .

### 3.2. The accident problem given liability but in the absence of insurance

The question addressed here is what the Pareto optimal solution to the accident problem is in the presence of strict liability or the negligence rule. Formally, the problem is to maximize  $EV$  over the parameter(s) of the liability rule subject to the constraint that injurers choose care  $x$  to maximize  $EU$  under the liability rule, and subject also to the constraint that a lump sum ex ante transfer  $r$  be such that  $EU = k$ .

Under strict liability, the following can be shown. First, if injurers are risk neutral, the (Pareto) optimal magnitude of liability  $d$  is the harm  $h$ , and the first-best solution is achieved. The main reason is that, as injurers are risk-neutral, their bearing of risk

<sup>9</sup> In characterizing Pareto optimal solutions to the accident problem, we are of course characterizing social welfare optima. For were we to maximize any social welfare function depending positively on parties' expected utilities, the optimum would be Pareto optimal.

<sup>10</sup> A justification for writing the resource constraint in terms of expected values is that accident risks among the population are independent and that the population is large.

does not reduce their expected utility, and victims are protected against risk because, by definition of strict liability, they are compensated for harm. But second, if injurers are risk averse, the optimal magnitude of liability  $d$  is less than  $h$  and the first-best outcome is not achieved. The explanation is that, since injurers are risk-averse, it is desirable to lower liability a positive amount from  $h$ , since the first-order benefit in terms of reduce risk-bearing is positive, whereas the first-order loss from suboptimal incentives is zero or negative. Note that one interpretation of this result is that when injurers are risk-averse, it is not desirable to fully “internalize” a stochastic externality.

Under the negligence rule, the following is true. First, if victims are risk-neutral, the optimal due care standard  $\underline{x}$  equals  $x^*$  and the first-best solution is achieved. This holds because injurers will be lead to choose  $x^*$  (if  $d = h$ ) and thus will bear no risk (unlike under strict liability), and risk-bearing by victims will not reduce their expected utility as they are risk neutral. Second, if victims are risk-averse, then the optimal  $\underline{x}$  is generally unequal to  $x^*$  and the first-best solution generally is not achieved. In this case, because victims are risk-averse, they bear risk when injurers take due care, implying that the first-best solution is not achieved, and also that it may be desirable for risk to be further lowered by raising  $\underline{x}$  above  $x^*$ .

Note that consideration of risk-bearing alters the comparison of strict liability and negligence. Strict liability becomes more appealing when injurers are less risk averse than victims, since strict liability imposes risk on injurers. The negligence rule becomes more appealing when injurers are more risk-averse than victims, since the negligence rule imposes risk on victims.

### 3.3. The accident problem given liability in the presence of insurance

Now assume that victims can purchase insurance against accident losses that they might bear, and that injurers can purchase liability insurance against liability judgments that might be imposed on them. Assume also that the insurance policies are optimal for insureds in the usual sense that the policies are designed to maximize the expected utility of insureds, given the constraint that the insurance premium is actuarially fair (equal to expected coverage). Two assumptions about the information of insurers will be considered: that insurers can observe care  $x$  and thus make premiums depend on  $x$ ; and that insurers cannot observe care (so that a situation of moral hazard exists).

In this case, then, the formal problem of finding the Pareto optimal liability rule is to maximize  $EV$  over the parameter(s) of the liability rule subject to the two constraints given in section 3.2 and subject also to the constraints that injurers and victims purchase insurance policies that maximize their expected utility given that premiums are fair.

Under strict liability, the following can be shown. The magnitude  $d$  of liability equals the harm  $h$ , and the first-best solution to the accident problem will be achieved unless injurers are risk-averse and liability insurers cannot observe care. Whether or not the first-best solution is achieved, it is not socially desirable for the government to interfere in the insurance market.

The foregoing is clearly true if injurers are risk-neutral, so let us consider the case where injurers are risk-averse. Then if liability insurers can observe care, it is straightforward that the first-best outcome will be achieved. The reason is that injurers will purchase full liability insurance coverage, will pay a premium of  $p(x)h$ , will choose care  $x$  to minimize the cost of care plus the insurance cost, namely, to minimize  $x + p(x)h$ , and hence will choose  $x^*$ . Since injurers do not bear risk (because coverage is full), victims do not bear risk, and care is optimal, the first-best outcome is achieved. The other case is that in which liability insurers cannot observe care, so that moral hazard exists. In this case, as a general matter, injurers will find it optimal (due to moral hazard) to purchase partial liability insurance coverage of  $c < h$ , so injurers will bear risk of  $h - c$ . Hence, they will have a positive incentive to take care, even though they own liability insurance. The first-best outcome will not be achieved because care will generally be different from  $x^*$  and because risk-averse injurers bear risk (of  $h - c$ ). However, it can be shown that the outcome that results when damages  $d = h$  and individuals purchase their privately-optimal liability insurance is second-best (equivalent to what a dictator could achieve if he was not able to directly control care but could control levels of wealth contingent on the occurrence of accidents). Hence, it can be demonstrated that government intervention in the liability insurance market is not desirable. This conclusion, about the undesirability of government intervention, is not transparent.<sup>11</sup>

Under the negligence rule, if due care  $\underline{x}$  is set equal to  $x^*$ , the first-best solution is achieved. The explanation is, essentially, that injurers will be led to take due care and will not want to purchase liability insurance because they will not bear the risk of liability (it can be shown that they would not want to fail to take due care and buy liability insurance).<sup>12</sup> Further, since injurers take due care and are not held liable, victims are induced to purchase accident insurance so will not bear risk themselves.

Three points may be made in summary of the foregoing discussion. First, the presence of liability insurance affects incentives of injurers in a manner that depends on whether insurers can observe their level of care; liability insurance translates and modifies, but does not eliminate, injurers' incentives to take care under the threat of liability. Second, there is no basis for government intervention in liability insurance markets despite the moral hazard that liability insurance creates when insurers cannot observe care. This point is not only of intellectual note. Historically, the sale of liability insurance was viewed with skepticism and delayed in some countries, on the ground that it might interfere with incentives; in the former Soviet Union, liability insurance was proscribed;

<sup>11</sup> However, partial intuition for the conclusion may be helpful to provide. Suppose that it is assumed (rather than proved to be optimal) that the magnitude of liability is  $h$ . Then it must be welfare enhancing to allow injurers to purchase whatever liability insurance policy they please, for doing so must raise their expected utility, and this cannot affect the expected utility of victims since they are fully compensated for harm, given that liability is strict and that  $d = h$ .

<sup>12</sup> The conclusion that injurers do not purchase insurance under the negligence rule does not hold if courts might err in the negligence determination or if other uncertainties lead to the risk that injurers would be found liable. For a model of liability and insurance that considers this possibility, see [Sarath \(1991\)](#).

and today liability insurance is disallowed in this country in some domains.<sup>13</sup> Third, the presence of liability and accident insurance means that the risk-bearing reasons favoring either strict liability or the negligence rule, given above in section 3.2, are essentially mooted; if injurers are risk-averse, they can purchase liability insurance coverage under strict liability, and if victims are risk averse, they can purchase accident insurance coverage under the negligence rule.

#### 4. Administrative costs

The administrative costs of the liability system are the legal expenses and the time and effort of litigants and of the state that are generated by the bringing and the resolution of suits. These costs are substantial; a number of estimates suggest that on average, administrative costs of a dollar or more are incurred for every dollar that a victim receives through the liability system.<sup>14</sup>

The factor of administrative costs raises the issue of whether the use of the liability system is worthwhile, that is, about the socially desirable volume of suits. The presence of administrative costs also leads us to reexamine optimal risk-reducing behavior and to inquire about the comparison between strict liability and negligence. In addressing these questions, it will be convenient to assume that parties are risk neutral and to consider the unilateral model of accidents with injurers' care being the only variable. The social goal will be taken to be minimization of social costs, comprised of the costs of care, expected harm, and now also expected litigation costs.

##### 4.1. Volume of suit

In order to relate the private incentive to sue to what is socially desirable—and to show that they fundamentally diverge—it will be useful to examine in this section a discrete version of the unilateral model, where there is just a single positive level of care. (This allows us to isolate the issue of the volume of suit from the issue of the level of care, since care is not continuously variable.) Let  $x$  be the single level of care that injurers can exercise,  $p$  the probability of harm  $h$  if care is not taken, and  $p' < p$  the probability of harm if care is taken. If  $x < (p - p')h$ , let us say that care is “efficient” since it lowers social costs, other things being equal. Further, let  $c_V$  be the cost of suit for a victim,  $c_I$  the cost of suit for an injurer,  $c_p$  the cost born by the public, and suppose that liability is strict.

When is suit socially desirable? In a general sense, the answer is: when the benefit of suit, in terms of inducing the exercise of care and reducing expected harm, exceeds

<sup>13</sup> On the historical resistance to the sale of liability insurance, see Tunc (1974, pp. 50–52). In this country today, liability insurance coverage is barred against punitive damages in some jurisdictions and also against certain types of fines; see Jerry (1996, pp. 471–477).

<sup>14</sup> See Danzon (1985, p. 187), Kakalik and Pace (1986, p. vii), and Tillinghast-Towers Perrin (2000, p. 12).

expected litigation costs. To amplify, if suit does not induce care to be taken, then plainly suit cannot be socially desirable: for if suit is not brought, social costs will be  $ph$ , whereas if suit is brought, social costs will be  $p(h + c_V + c_I + c_P)$ , which is higher due to litigation costs. Suppose now that suit does induce care to be taken. Then if suit is brought, social costs will be  $p'(h + c_V + c_I + c_P) + x$ . Hence, suit is socially worthwhile if and only if  $p'(h + c_V + c_I + c_P) + x < ph$ , or equivalently, if and only if the following key social condition holds:

$$p'(c_V + c_I + c_P) < (p - p')h - x.$$

This is exactly the condition that expected litigation costs are less than the net deterrence benefits of suit.

It is readily seen that suit may be brought by private parties even though that is socially undesirable. Suit will be brought by a victim of harm whenever  $c_V < h$ . This can be true even though suit is not desirable. For example, suit might not affect care at all—might not induce  $x$ —yet still be brought by victims since  $c_V < h$ ; in this case, suit would be a pure waste, as the expected litigation costs of  $p(c_V + c_I + c_P)$  would accomplish nothing. Even if suit induces care, the key social condition of the last paragraph for suit to be desirable might not hold when  $c_V < h$ . An aspect of this problem, note, is that in contemplating suit, the victim takes no account of the possibility that suit may fail to induce care or that the exercise of care may have little social value.

It is also evident that suit may not be brought even though it would be socially desirable that it be brought. This would be so if the bringing of suit would induce care and thereby produce a valuable deterrent; in other words the key social condition may hold even though  $c_V > h$ . An example is this:  $h$  is 100,  $c_V$  is 200, other litigation costs are 0,  $x$  is 1,  $p$  is 1, and  $p'$  is .01. Here suit is not brought, since 200 exceeds 100, so social costs are 100 (for  $p$  is 1). But were suit brought care of 1 would be exercised, the probability of harm would drop to .01, so expected social costs would be  $.01(100 + 200) + 1 = 4$ , which is far lower. The problem here is in part that when the victim considers suit, the fact that his willingness to sue would create very beneficial deterrence is of no moment to him.

Several remarks about the foregoing are worth making, helping to explain why the actual incentive to sue may be different from what is socially best. First, from the purely formal perspective, the private condition determining suit, whether  $c_V < h$  holds, is facially quite different from the key social condition,  $p'(c_V + c_I + c_P) < (p - p')h - x$ . Second, there are two intuitively understandable social/private divergences at work. One is that the victim does not take into account the point that bringing suit causes a negative externality in the form of litigation costs on the injurer and society, namely  $c_I + c_P$ . The other is that, as mentioned, the victim does not take into account a positive externality due to suit, the creation of incentives to take care and to lower risks (for the victim sues only after an accident occurs). These two private/social divergences, working in opposite directions, make understandable how it is that there can be either too much or too little suit.

The private/social divergence may be of substantial importance in fact. For example, the financial reasons to sue for losses sustained in an automobile accident are often high and result in a tremendous amount of litigation and attendant expense—about half of all tort litigation in the United States concerns automobile accidents. Yet the incentive to drive safely may be relatively little affected by the prospect of suit (because incentives to drive safely, such as they are, have mainly to do with fear of injury from an accident or fear of criminal liability). If so, the large volume of automobile-accident litigation may largely constitute a social waste (my point is not so much that it *is* a social waste as it is that it could be; that the litigation is observed does not signal its social utility).<sup>15</sup>

The private/social divergence could be remedied by, for example, the state barring suit where it is undesirable but would be brought, or by the state subsidizing suit where it would be desirable but not be brought. For the state to do this, however, requires that it determine, among other things, the incentive that suit generates, in other words, that the key social condition must be ascertained. This imposes a high informational burden on the state. There is no easy fix, no simple remedy (such as making the victim pay the full litigation costs of suit) that will result in the socially desirable level of suit or necessarily to an improvement.

The points made here about the private versus the social incentive to use the legal system are first made in [Shavell \(1982b\)](#).

#### 4.2. Level of care and volume of suit

Now let us return to the setting with care continuously variable, so that we can consider not only the issue of the volume of suit but also that of variation in the level of care. As will be seen, the level of care should be higher than the otherwise socially best level  $x^*$ , due to litigation costs. The reason is essentially that the occurrence of an accident creates greater social costs than just the harm  $h$  itself, as the social costs equal the harm plus litigation costs.<sup>16</sup>

To amplify, consider the following second-best problem: a dictator with the goal of minimizing total social costs can order victims when to sue and how much liable injurers should pay, but the dictator cannot directly control the level of care  $x$ . The solution to this second-best problem is a natural standard of comparison for the functioning of the liability system. It can be shown that the second-best solution has the following character. If suit is brought, injurers pay  $h + c$ , where  $c = c_V + c_I + c_P$ , so injurers should bear the harm plus total litigation costs. The explanation for this result is that, if suit is brought, then the occurrence of harm creates social costs of  $h + c$ , so that the injurer should be given an incentive to take care reflecting this amount. Additionally, under the second-best solution, suits are brought if and only if the following key social

<sup>15</sup> For empirical study of the effect of liability on automobile accidents, see [Cummins et al. \(2001\)](#), [Deweese et al. \(1996\)](#), [E. Landes \(1982\)](#), and [Sloan \(1998\)](#).

<sup>16</sup> This section follows [Shavell \(1999\)](#).

condition is satisfied:

$$p(x^*(h+c))c \leq [p(0) - p(x^*(h+c))]h - x^*(h+c),$$

where  $x^*(h+c)$  is the  $x$  that minimizes  $x + p(x)(h+c)$ , that is, the optimal level of care when accident losses are  $h+c$ . The left side of this condition is expected litigation costs and the right side is the reduction in expected harm net of the cost of care induced by suit.

Under strict liability, with damages  $d = h$ , suit is brought when  $c_V < h$ . For reasons that are essentially those given in section 4.1, it is possible that suit is brought when the key social condition does not hold and suit is undesirable; and it is also possible that suit is not brought when the key social condition does hold and suit is desirable. If suit is brought, the level of care that will be taken is  $x^*(h+c_I)$ , since the injurer bears his own litigation costs; this level of care is less than the second-best optimum level of care of  $x^*(h+c)$ . In summary, suit might be excessive or inadequate; and when suit is brought, the level of care is too low.

The second-best outcome can be achieved, however, assuming that the state can determine whether the key social condition holds. For suit can be barred if that is optimal and subsidized if need be. Further, if suit is optimal and is brought, the correct level of care can be induced by raising damages by making the injurer pay  $h + c_V + c_P$ ; since he naturally bears his own litigation costs, this level of damages means that the injurer's total expenses are  $h+c$  so that his level of care will be second-best optimal.

### 4.3. Comments

(a) The basic points made above about the social versus private incentive to make use of the liability system, given that it is costly to employ, are robust. Whatever the specific nature of the model, it will tend to be true that a party contemplating making an expenditure will not take into personal account the negative externality that his expenditure will engender, in the form of expenditures by the other side of litigation and by the state; also, the party will not take into account the incentive effects of his expenditure on the behavior of others. As suggested in section 4.1, the possibility of a divergence seems to be of substantial policy interest because of the magnitude of administrative costs.

(b) Also, the point that the level of care (as distinguished from the volume of suit) should reflect the fact that when harm occurs and suit is brought, the social consequences are not limited to the harm but include litigation costs is general, and it yields a relatively easily applied policy prescription, that injurers' total payments should reflect total litigation costs. Such payments need not be in the form of damages; they could be in the form of fines on top of damages.<sup>17</sup>

(c) The difference between the private and the social motive to litigate is, as noted, initially developed in Shavell (1982b), and is extended in various ways in Menell (1983),

<sup>17</sup> An advantage of having litigation costs paid as fines is that then victims would not have an incentive to spend too much on litigation.



Kaplow (1986), Rose-Ackerman and Geistfeld (1987), and Spier (1997). Also, Polinsky and Rubinfeld (1988) consider the incentive to bring suit and injurers' level of care, presuming that the only policy instrument is the magnitude of damages. Under this assumption, the state is not able to induce both the optimal volume of suit and the optimal level of care, and as a consequence, the optimal level of damages does not equal harm plus others' litigation costs and may even be less than harm (if it is desirable to discourage suit and suit cannot be barred or taxed, damages need to be lowered). Polinsky and Che (1991), Hylton (1990), and Shavell (1999), however, allow for the level of suit to be controlled separately from injurers' level of care. For further discussion of litigation and private versus social incentives, see Shavell (1997) and Spier (2007).

#### 4.4. *Strict liability and negligence*

Administrative costs enter into the comparison of strict liability and negligence as forms of liability. However, there does not seem to be a clear a priori difference in litigation costs under the two. Strict liability would be expected to result in a higher volume of cases than the negligence rule, for cases will be brought under strict liability whenever  $c_V < h$ , whereas cases will often not be brought under the negligence rule when  $c_V < h$  since the injurer will often be known not to have acted negligently.<sup>18</sup> Although this factor of the volume of cases implies that strict liability is more expensive than the negligence rule, the litigation cost per case disputed is likely to be higher under the negligence rule, since negligence has to be determined, and since settlement of cases is less likely than under strict liability. Hence, either strict liability or the negligence rule could turn out to be the more costly on grounds of administrative costs, depending on the relative importance of the volume of cases (making strict liability more expensive) and the litigation cost per case (making the negligence rule more expensive).

Another issue bearing on strict liability versus the negligence rule concerns private and social incentives to bring suit. There is some basis for thinking that the problem of socially excessive suit discussed above is less likely under the negligence rule than under strict liability. To understand why, observe that if the negligence rule functions perfectly, it will be socially advantageous for suit to be subsidized so that it would be free to bring: for injurers will then decide to act nonnegligently, no suits will in fact be brought, and no litigation costs will be incurred. However, if as is realistic one assumes that courts may err in the negligence determination and/or that victims may not be able to tell whether injurers are nonnegligent, suits will sometimes be brought under the negligence rule. In consequence, the general qualitative results reached under strict liability will apply under the negligence rule as well, although the likelihood of excessive litigation would seem to be lower.

<sup>18</sup> Farber and White (1991) provide evidence that many medical malpractice cases are dropped when plaintiffs learn that the defendant probably was not negligent. Relatedly, Ordover (1978) examines a model in which victims' decision whether to sue under the negligence rule depends on their beliefs about the negligence of defendants.

## Part B: Extensions

### 5. Liability of firms

Let us reconsider the theory of liability under the assumption that injurers are firms. This will require us to take into account the relationship between liability and market price, and also the level of production. Two cases will be distinguished: where accident victims are strangers to firms, such as where a person's car is damaged by a firm's truck; and where accident victims are the consumers of firms and are injured because of their purchase of a product or service, such as where a person's water heater ruptures and damages his property. Where accident victims are consumers, firms may be concerned about the risks generated by their products because consumers will pay less for risky products to the degree that they perceive risk; hence the need for liability as an incentive tends to be reduced.

Firms will be presumed to be identical, to maximize profits, and to operate in a perfectly competitive environment, so that product price will equal unit cost of production, including expected liability costs. Firms and victims will be assumed to be risk neutral, accidents to be unilateral (only firms' behavior affects risk), and the liability system to operate without administrative cost. The measure of social welfare is analogous to that considered in section 2 with activity levels: the utility consumers derive from products (such as from water heaters) and, where relevant, the utility that strangers obtain from their activities (such as from driving), minus expected harm, the costs of care, and direct costs of production. The development below follows the lines of Shavell (1987).

#### 5.1. Victims are strangers

Let  $c$  be the direct production cost per unit of a firm's product,  $x$  the cost of care per unit of the product,  $s$  the quantity of the product produced and consumed, and  $u(s)$  the utility consumers obtain from the product. Then social welfare is  $u(s) - s[c + x + p(x)h]$ , where the term in brackets is the production cost, cost of care, and expected harm suffered by strangers per unit. It is clear that optimal care  $x^*$  is, as earlier, the  $x$  minimizing  $x + p(x)h$ . Hence, the optimal quantity of the product  $s^*$  is determined by  $u'(s) = c + x^* + p(x^*)h$ , which has the familiar interpretation that marginal utility must equal the production cost of a unit, here including cost of care and expected harm.

Under strict liability the level of care and also the quantity produced will be socially optimal. Firms will choose  $x^*$ , in order to minimize unit production cost. Hence, unit production cost and price will equal  $c + x^* + p(x^*)h$ , implying that  $u'(s) = c + x^* + p(x^*)h$ , so that  $s$  will equal  $s^*$ . The quantity produced will be optimal because the price will reflect the expected harm caused by the product.

Under the negligence rule, if due care  $\underline{x}$  equals  $x^*$ , then by the logic of section 2.1, firms will choose  $x^*$ , so that unit costs and product price will be  $c + x^*$ . Hence, the quantity sold will be determined by  $u'(s) = c + x^*$ , implying that the quantity under the negligence rule is socially excessive, since  $c + x^* < c + x^* + p(x^*)h$ . The reason

is that the price does not reflect expected harm under the negligence rule. The problem with the negligence rule is analogous to that discussed in section 2.3; here the analog to the activity level is the quantity produced of the good.

Thus, the way that the liability rules function is similar to how they function when injurers are individuals; the effect of liability on firms' care is the same as on individual injurers' care, and the effect of liability on purchases via prices is analogous to its effect on individuals' choice of activity levels. The point that strict liability, but not the negligence rule, results in prices that induce optimal production is made in Polinsky (1980b) and Shavell (1980b).

## 5.2. Victims are consumers

Define the full price of the product as the market price plus the perceived expected accident losses that would be borne by a consumer. Consumers will choose their purchases  $s$  to maximize their utility  $u(s)$  given the full price.

Suppose first that consumers can observe the risk  $p(x)$  associated with a firm's product. Then the outcome will always be optimal. In the absence of liability, firms that choose  $x$  will charge  $c+x$ , and consumers will view the full price as  $c+x+p(x)h$ . Since consumers will buy from firms with the lowest full price, firms will be led to choose  $x^*$  to minimize the full price. Since the full price will be  $c+x^*+p(x^*)h$ , consumers will choose  $s$  to maximize  $u(s) - s[c+x^*+p(x^*)h]$ , so they will choose  $s^*$ . Under the negligence rule with due care of  $x^*$ , firms will choose  $x^*$  and charge  $c+x^*$ , the full price will again be  $c+x^*+p(x^*)h$ , so consumers will choose  $s^*$ . Under strict liability, firms will choose  $x^*$  to minimize the market price, which will be  $c+x^*+p(x^*)h$ ; this will also be the full price (since consumers do not bear any losses), and consumers will choose  $s^*$ . That the presence of liability does not matter is due to the assumption that consumer information is perfect, meaning that market incentives result in proper care and consumer knowledge of the full price leads to correct purchases.

Suppose next that consumers cannot observe  $p(x)$  for an individual firm but do know  $p(x)$  on average. In the absence of liability, a firm would not choose positive  $x$ , for it could not charge a higher price for so doing. Hence,  $x = 0$  in equilibrium, and since consumers know the risk to be  $p(0)$  (as they know the average risk), the full price is  $c+p(0)h$ , and  $s$  satisfies  $u'(s) = c+p(0)h$ . Thus, care is too low and  $s$  is unequal to  $s^*$ , but  $s$  is optimal given  $p(0)$ . Under the negligence rule and under strict liability, firms act as they do in the previous case, so that  $x^*$  and  $s^*$  are chosen. In this case, then, the presence of liability matters, since care will not be taken without liability. Even the negligence rule results in the correct level of purchases  $s^*$  since consumers know the risk  $p(x^*)$  in equilibrium.

Suppose now that consumers misperceive average risk, say they believe the risk is  $\alpha p(x)$  where  $\alpha$  is unequal to 1. To illustrate the difference this makes from the previous situation, suppose  $\alpha < 1$ , so consumers underestimate the average risk (the case of  $\alpha > 1$  is analogous). Then in the absence of liability, firms again choose  $x = 0$ , and individuals choose  $s$  to minimize the full price of  $c + \alpha p(0)h$ , so they choose  $s$

that is too high given the risk of  $p(0)$ . Under the negligence rule with due care of  $x^*$ , firms choose  $x^*$  and the full price is  $c + x^* + \alpha p(x^*)h$ , so that  $s$  is too high. Under strict liability, however, since firms choose  $x^*$  and the market price of  $c + x^* + p(x^*)h$  is the full price,  $s^*$  is chosen. In other words, when consumers misperceive average risks, liability is needed to induce firms to take optimal care, and strict liability rather than the negligence rule is needed to induce consumers to purchase the correct quantity.

To summarize, when consumer information about product risk is perfect at the level of the individual firm, liability is not needed to induce optimal care nor to provide consumers with the right price on which to base purchases, since they take expected harm into account in deciding on purchases. When, however, consumer information is not perfect, liability is generally needed to achieve optimality, and the nature of the imperfection of information influences how, exactly, and the degree to which, liability improves outcomes.

Last, let us comment on warranties, which is to say, the effective choice of liability rules by consumers. Consumers will choose the form of warranty that results in the lowest full price. It follows that if consumers correctly perceive average risk, they will purchase warranties imposing either negligence or strict liability on firms, and the outcome will be optimal. If consumers misperceive risks, however, warranty choice may not lead to the optimal outcome. To illustrate, suppose that consumers underestimate average risk. Then consumers will tend not to want a warranty giving strict liability, for this would lead to a market price that appears expensive to them (in the extreme case where consumers believe the risk to be zero, the warranty would raise the price but appear to have no value to consumers). As a consequence, consumers might elect a negligence rule form of warranty, or none at all; thus, the outcome, would not be optimal; notably, purchases would be excessive. Therefore, misperception of risk may lead to problems with choice of warranties.

Literature on firms' liability to consumers, so-called product liability, includes early papers by Oi (1973) and Hamada (1976), assuming that consumers have perfect information about risk. Articles that study imperfect information about product risk and/or liability include Goldberg (1974), Schwartz and Wilde (1979), and Shavell (1980b). Key articles on the theory of warranties are Grossman (1981) and Spence (1977); see also Priest (1981) for an important informal treatment of warranties emphasizing consumer behavior.

## 6. Aspects of the negligence determination

In this section, I extend the simple model of the negligence determination in several ways. In doing this, I consider the unilateral model of accidents, but what is said about injurers will carry over to victims in regard to the determination of contributory negligence.

### 6.1. Differences among injurers

One issue that bears on the negligence determination concerns differences among injurers that lead to differences in the optimal level of care. In particular, suppose that injurers vary in their cost of exercising care (were the difference to concern some other factor, such as the likelihood of causing harm, the conclusions would be similar). Let  $kx$  be the cost of exercising care of  $x$  for an injurer of type  $k$ , where  $k$  is distributed on an interval  $[0, K]$ . The optimal level of care for an injurer of type  $k$  is the  $x$  that minimizes  $kx + p(x)h$ ; denote this by  $x^*(k)$ , which is increasing in  $k$ .

If courts can determine an injurer's type, they can set the due care standard for each type optimally and induce all to act optimally; that is if  $\underline{x}(k) = x^*(k)$ , then each type of injurer will take due care and act optimally. However, if courts cannot observe  $k$ , and must set a single due care standard  $\underline{x}$  in  $[x^*(0), x^*(K)]$ , then it can be shown that the optimal level of this uniform due care standard is greater than or equal to  $x^*(E(k))$ , the optimal level of care for the person with the mean cost parameter. (The optimal level of care for this mean person is sometimes referred to as the standard of due care for the "reasonable" man.)

Now make the additional assumption that injurers have a choice whether or not to engage in their activity, and if they so do they obtain a benefit of  $u$  (there is only one level of activity, for simplicity). Then it is socially beneficial for an injurer of type  $k$  to engage in the activity if and only if  $u > kx^*(k) + p(x^*(k))h$ . However, we know from above (see section 2.3) that, under the negligence rule, injurers may decide to engage in the activity even though doing so is socially undesirable; the problem is that injurers can avoid having to pay for the activity by taking due care. This problem is most serious in regard to injurers with high  $k$ —for whom it is very expensive to take care or who are awkward or inept; for them  $kx^*(k) + p(x^*(k))h$  will be high—perhaps greatly exceeding  $u$ —since  $x^*(k)$  will be low. This has an implication for the setting of due care. Even if the courts can observe injurers' type  $k$ , the optimal level of due care may not be  $x^*(k)$  for all  $k$ ; rather optimal due care may be  $\underline{x} > x^*(k)$  for all  $k$  above some threshold  $k'$ . By setting such a due care standard, engaging in the activity may become too expensive to be worthwhile for high  $k$  types, thus implicitly combating the problem of excessive engagement in the activity for these most dangerous types. For similar reasons, in the case where  $k$  cannot be observed, the optimal single level of due care will tend to be higher than in the last paragraph.<sup>19</sup>

### 6.2. Imperfect assessment of care by courts

Another issue that bears on the determination of negligence is that courts may make mistakes in assessing the care exercised by injurers. Let  $e$  be the error in the level of

<sup>19</sup> On differences among individuals and the negligence determination, see Diamond (1974) and Shavell (1987, pp. 86–91).

care, so that  $x + e$  is the level of care observed by a court. Thus, if  $\underline{x}$  is the due care standard, negligence will be found when  $x + e < \underline{x}$ , which is to say, when  $e < \underline{x} - x$ . Accordingly, even if  $x = \underline{x}$ , negligence might be found in the event of an accident, for if  $e < 0$  negligence would be found.

An injurer subject to the negligence rule will choose  $x$  to minimize  $x + \text{Prob}[x + e < \underline{x}]p(x)h$ , where Prob denotes probability. It can be shown that if the due care level  $\underline{x}$  would be met in the absence of error, then in the presence of error injurers will choose  $x > \underline{x}$ , provided that the distribution of  $e$  is sufficiently concentrated. In particular, care will be excessive under this condition if due care  $\underline{x}$  is set at the optimal level  $x^*$ . Also, negligence will sometimes be found in general. The intuition behind this result that care tends to be excessive due to error in its determination is that the injurer can guard against being found negligent by taking excessive care; for taking more than due care means that even if actual care is underestimated, he might still escape liability, and taking more than due care may well cost less than the reduction in expected liability that it brings about. The analysis of error in observing care is first examined in Diamond (1974) and the point about excessive care is emphasized in Craswell and Calfee (1986).<sup>20</sup>

### 6.3. Imperfect ability to control care by injurers

An additional issue that affects the finding of negligence is that individuals may not be able to control perfectly what might be called their momentary level of care, such as a driver's behavior at a particular instant. A driver might be unable to control his or her behavior at a particular instant due to a sneeze (causing the driver to swerve) or to an inadvertent lapse in attention. In contrast, what an injurer is able to control is his general habit of attention, his mental attitude (which might be reflected, for instance, by the frequency with which a driver looks at his rear view mirror, how frequently he switches lanes, and the like). But a court would usually find it difficult to observe this mental attitude, as opposed to the momentary level of care. We might formalize these observations by assuming that an individual chooses  $x$ , his general level of attention, whereas his momentary level of care is  $x + e$  where  $e$  is a random term, and where it is  $x + e$  that the court observes. Hence, the injurer will be found negligent in the event of an accident if  $x + e < \underline{x}$ , and the analysis of his behavior will be similar to that in the last section, with the conclusion being that care will tend to be excessive if  $\underline{x} = x^*$ , and that negligence will sometimes be found.<sup>21</sup>

### 6.4. Errors in the calculation of due care

A further issue concerns whether courts may make mistakes in ascertaining the optimal due care standard. If courts' information about the function  $p(x)$  or about the cost of care is imperfect, then their calculation of  $x^*$  will tend to be incorrect.

<sup>20</sup> The analysis is further developed in Shavell (1987, pp. 93–99).

<sup>21</sup> On the momentary level of care, see Diamond (1974) and Shavell (1987, pp. 96–97). See also Grady (1983).

One possibility is that injurers do not know in advance how the courts will err in determining  $x^*$  and thus in determining due care  $\underline{x}$ . In this case, injurers will tend to take excessive care in order to escape liability due to mistake, and negligence will sometimes be found, for reasons closely associated with those discussed in the last two sections.

The other possibility is that injurers anticipate what the (mistaken) due care standard will be. Under this assumption, if due care is too low,  $\underline{x} < x^*$ , injurers will take care of  $\underline{x}$  and will not be found negligent. If due care is too high,  $\underline{x} > x^*$ , injurers will take care of  $\underline{x}$  and will not be found negligent, provided that  $\underline{x}$  does not exceed  $x^*$  by too much. If  $\underline{x}$  is sufficiently high, however, injurers will decide to be negligent and will choose  $x^*$ .<sup>22</sup>

## 7. Causation

A fundamental principle of liability law is that a party cannot be held liable unless he was the cause of losses. For example, if a surgeon negligently performed a procedure and his patient died on the operating table, but the patient would have died even if the surgeon had properly carried out the procedure, the surgeon would not be held liable. Or if a firm pollutes a stream with a carcinogenic agent and a person living nearby develops cancer, but the cancer is shown to be due to another carcinogenic factor, the firm will not be held liable.

The idea of causation of harm can be expressed as follows. Let  $s$  denote a state of the world in the universe of possible states. Let  $h(s, x)$  be the harm that occurs in state  $s$  if action  $x$  (which could be a level of care) is taken. Then, given a state  $s$ , taking action  $x$  may be said to be a cause of losses of  $h$  relative to some other reference or comparison action  $x_0$  if  $h(s, x) - h(s, x_0) = h$ . Often, when it is said that an injurer is a cause of losses, the comparison action is not mentioned but it will be implicit from the context.

Let us now consider the effect of the causation requirement—that liability is imposed only if the injurer was the cause of harm—on incentives under liability rules. On this issue, see originally Calabresi (1975) and Shavell (1980a).<sup>23</sup>

### 7.1. Strict liability

Under strict liability, it is readily shown that if injurers are held liable for the harm that their activity causes, but not for other harm, then their incentives to take care and to engage in harmful activities will be optimal. Here, by harm that their activity causes is meant harm that is caused by their engaging in their activity relative to their not engaging in their activity.

If injurers are held liable for less than the harm they cause, or for more than that harm, their incentives will generally be suboptimal. Notably, if they were held liable for harm that they do not cause—if firms are held liable for cancer that their pollution does not

<sup>22</sup> See Shavell (1987, pp. 97–98).

<sup>23</sup> See also W.M. Landes and Posner (1987, pp. 228–255) and Shavell (1987, pp. 118–123).

cause—then they might be led to take excessive care<sup>24</sup> or be undesirably discouraged from engaging in their activity.

### 7.2. Negligence rule

Under the negligence rule, it can be shown that injurers will be led to take due care, assuming that due care equals optimal care, if they are held liable only if they caused harm and pay damages equal to harm. Here two possible definitions of harm are natural to consider. One is harm relative to the harm that would have occurred had the injurer not engaged in his activity (for example, not driven a car). The other possible definition is harm relative to the harm that would have occurred had the injurer taken due care; in other words, harm would be measured by  $h(s, x) - h(s, \underline{x})$ . I have employed the first definition in the present chapter and will continue to do so below.<sup>25</sup>

If injurers were held liable more often than when they caused harm, this would not lead to their taking excessive care; it would only augment their already sufficient incentive to take due care. Thus, unlike under strict liability, there is not an affirmative reason to restrict liability to accidents caused by injurers under the negligence rule in the simple unilateral model of accidents.

Nevertheless, there exist advantages of restricting liability on the basis of causation under variations of the basic unilateral model. Specifically, in the presence of uncertainty about the negligence determination, relaxation of the causation requirement might exacerbate the tendency to take excessive care (this tendency exists under the assumption that damages are measured according to the first definition of harm). Further, relaxation of the causation requirement would tend to increase the volume of litigation and thus administrative costs.

### 7.3. Uncertainty over causation<sup>26</sup>

In many situations there is uncertainty about causation. For example, it may not be known which manufacturer out of many sold the product that resulted in injury, or whether harm was due to the defendant firm's pollution or to background factors. The traditional approach of the law is to employ a 50% probability threshold: to hold a defendant liable if and only if the probability that the defendant was the cause of losses

<sup>24</sup> To illustrate why an injurer might take excessive care, imagine that a firm would be made to pay for cancers that its pollution did not cause if it discharges some pollution, but not otherwise. Then the firm would have an excessive incentive to reduce pollution to zero (perhaps by installing a socially undesirably expensive pollution control device).

<sup>25</sup> The main reason that I adopt the first definition of harm is that courts frequently have insufficient information about  $h(s, \underline{x})$  to calculate  $h(s, x) - h(s, \underline{x})$ . Were the second definition of harm employed, it can be shown that injurers would still be induced to take care of  $x^*$  if  $\underline{x} = x^*$ . The main differences that the second definition of harm would make to the analysis flow from the implication that expected liability would be continuous in  $x$  at  $\underline{x}$  rather than discontinuous. Kahan (1989) and Grady (1983) focus on the implications of the second definition of harm.

<sup>26</sup> On the subject of this section, see Rosenberg (1984) and Shavell (1985), and Shavell (1987, pp. 123–126).



exceeds 50%. More generally, the courts could employ a probability threshold  $t$  different from 50%.

The probability threshold approach may lead to suboptimal behavior. On one hand, if an injurer knows that the probability that he was the cause of losses will be less than 50% (or whatever is  $t$ ), he will never be held liable so will have no motive to reduce risk. On the other hand, if he anticipates that the probability that he was the cause of losses will exceed 50%, he will always be held liable, and his incentives will also be suboptimal, at least under strict liability. For under strict liability, his excessive liability implies that he will have too low an incentive to engage in the activity. Under the negligence rule, however, his being held liable even when he was not the cause of losses will only increase his incentive to take due care, so should not result in excessive care, assuming that the negligence determination occurs without error.

As an alternative to the probability threshold criterion, the legal system has sometimes adopted the proportional liability approach, that is, of imposing liability in proportion to the likelihood of causation. Under this approach, an injurer whose likelihood of being the cause of harm was  $q$  would pay damages of  $qh$  in a case for which the harm was  $h$ . (The proportional approach is often called market-share liability, since in a context in which different firms cause accidents and it is hard to identify which firm caused a particular harm, the likelihood of causation may be approximated by a firm's market share.) The proportional liability approach may lead to the same incentives to reduce risk that exist in the absence of uncertainty over causation. In particular, suppose that  $p(x)$  is the probability that an injurer would cause an accident, and that  $c$  is the probability of an accident due to other causes (assume these are mutually exclusive events for simplicity). Then, the probability of harm is  $c + p(x)$  and the probability  $q$  that harm was due to the injurer, as calculated by the court, will be  $q = p(x)/(c + p(x))$ . (Note that the court must know  $p(x)$  in order to compute  $q$ .) Under strict liability using the proportional approach, the probability of having to pay damages is  $c + p(x)$  and the amount paid would be  $qh$ , so that expected liability would be  $(c + p(x))qh = (c + p(x))[p(x)/(c + p(x))]h = p(x)h$ , implying that care,  $x$ , will be chosen optimally and the decision to engage in the activity will also be optimal.<sup>27</sup> Likewise, incentives to take due care under the negligence rule will be correct if due care is set optimally.

#### 7.4. Proximate causation<sup>28</sup>

Even if a party is a cause of losses, he may still escape liability under the law because he was not the "proximate cause" of losses. There are two major categories of losses that

<sup>27</sup> Incentives will not be optimal if, however, the courts cannot properly calculate  $q = p(x)/(c + p(x))$ . A context in which problems might arise is where many firms market a product that causes harm and market share is employed to determine  $q$ . It might be that a particular firm's product is safer than another's (one firm's drug might cause fewer side effects than another's). If so, its  $p(x)$  is lower than another's; but if only market share, and not  $p(x)$ , is considered by courts, incentives would not be optimal.

<sup>28</sup> For analysis of proximate causation, see Shavell (1980a) and Shavell (1987, pp. 110–115, 121–123).

are said not to be proximately caused. One is accidents that came about in an atypical, freak manner, for example, where a dog imbibes nitroglycerin left at a mining site and then explodes, injuring nearby persons. Allowing parties to escape liability for unusual accidents is sometimes thought not to undermine incentives, on the ground that no one could have foreseen such accidents. This argument, however, is subject to the criticism that courts may find it difficult to discriminate between accidents that can and cannot be foreseen and that it reduces the incentives of injurers to become informed. Moreover, the argument leads to the *reductio ad absurdum* that there should never be liability: after all, any accident may be viewed as extraordinarily unlikely (of essentially zero probability) if it is described in sufficient detail.

The other major category of accidents said not to be proximately caused, and for which parties are not held liable, are those coming about on account of what might be described as coincidence. For example, in a well known case, a speeding bus happened to be at just the “right” point on its route to be struck by a falling tree. Here, the bus company escaped liability for the injuries to passengers even though the injuries would not have occurred but for the excessive speed of the bus. Allowing parties to escape liability for such coincidental accidents might not affect incentives to take precautions, however. If the probability of a bus being struck by a falling tree is independent of its speed, failing to impose liability when trees fall on buses would not affect incentives to speed.

## 8. The magnitude of liability: damages

### 8.1. Basic theory: strict liability

Under strict liability, if the magnitude of liability, which as noted is called damages  $d$ , equals harm, then incentives to take care will be optimal in the unilateral model of accidents. For if so, the injurer’s problem will be to minimize  $x + p(x)h$ , which is the social problem, so he will choose  $x^*$ ; whereas if  $d < h$  the level of care will be inadequate and if  $d > h$  the level of care will be excessive. Likewise, for levels of activity to be optimal, damages must equal harm.<sup>29</sup>

The point that damages equal to harm is optimal essentially carries over to the situation, not yet considered, where the magnitude of harm is stochastic. Suppose that, if an accident occurs, the harm  $h$  is governed by a probability distribution with density  $f(h; x)$ ; thus, note, the level of care may affect the probability distribution of damages. The social problem in this situation is to minimize  $x + p(x) \int hf(h; x)dh = x + p(x)E(h; x)$ , where  $E(h, x)$  is expected harm conditional on an accident occurring

<sup>29</sup> In the bilateral model of accidents, damages equal to harm is optimal under the rule of strict liability with a defense of contributory negligence, if victims’ activity level is taken to be fixed. If, however, victims’ activity level is variable, then optimal damages may well be less than harm, for then some part of losses will be borne by victims and will induce them desirably to moderate their level of activity.

when care is  $x$ . It is clear from the expression of the social objective that if  $d = h$ , the injurer will minimize the social objective so will choose  $x^*$ .

Several observations are worth making about this more general situation where the magnitude of harm is stochastic.

(a) It is not necessary that  $d = h$  for incentives to be correct. If damages equal  $E(h; x)$  regardless of the actual  $h$ , then incentives will be correct.

(b) Legal doctrines that impose a ceiling on damages (on the ground that no one could have expected such high damages) or that exclude damages when the probability density  $f$  is sufficiently low (on the ground that no one could have expected such an unlikely event) lead to inadequate care and excessive activity levels; for such doctrines truncate expected liability conditional on an accident occurring to a level below  $E(h; x)$ .

(c) Legal doctrines that impose damages exceeding harm for certain types of outcome (see section 8.4 below on punitive damages) lead to excessive care and inadequate activity levels.

## 8.2. Basic theory: negligence rule

Under the negligence rule, analysis of the optimal magnitude of damages is different: on one hand, damages can be somewhat less than harm and optimal care will still be induced; on the other hand, damages can exceed harm and optimal care will be induced. To begin with, consider the simple model with one level of harm  $h$ , due care  $x$  set optimally at  $x^*$ , and no errors in the negligence determination. It was shown in section 2.1 that if  $d = h$ , the injurer will take due care of  $x^*$ . In fact, the incentive to take care of  $x^*$  rather than less care is sharp, in the sense that expected liability rises discontinuously from 0 to a positive level if  $x$  is reduced from  $x^*$ , since then expected liability jumps from 0 to  $p(x)h$ . This discontinuity in expected liability can be shown to imply there is an interval  $[h, h - k]$  for some positive  $k$  such that if  $d$  is in the interval, the injurer will still be induced to choose  $x^*$ . If  $d > h$ , then it is obvious that  $x^*$  will be taken and that greater care will not be taken, since if care is just  $x^*$  then the injurer will be relieved of liability.

These points require some qualification, however.

(a) If there is uncertainty in the negligence determination, then damages exceeding harm may exacerbate the problem of excessive care (see section 6.2).<sup>30</sup> In this situation, damages equal to harm will not induce optimal care, but some level of damages (perhaps less than  $h$ ) will induce optimal care.

(b) If the magnitude of harm is stochastic, the points made in the preceding paragraph largely carry through.

<sup>30</sup> An additional issue is that erroneous findings of liability tend to remedy the problem of excessive levels of activity under the negligence rule, raising the possibility that setting damages above  $h$  would be desirable.

### 8.3. *Difficulty in estimating harm*

Some components of harm are hard to estimate, for example, the decline in profits that will be caused by a fire at a store. The courts sometimes exclude such difficult-to-measure elements of harm from damages (on the ground that they are too speculative). This legal practice might be justified if the cost of ascertaining a component of harm outweighs the value of the improvement in incentives that its inclusion would accomplish. However, the cost of estimating a component of loss would be low if rough estimates are used, suggesting that that approach may be superior to omitting speculative components of harm from damages.

### 8.4. *Punitive damages*

When an injurer's behavior departs substantially from what is appropriate, damages in excess of harm, so-called punitive damages, may be imposed. If imposition of such damages causes expected liability to exceed expected harm, injurers will be induced to take excessive precautions, at least under strict liability, and they will also reduce their levels of activity undesirably. Under the negligence rule, excessive damages may also lead to undesirable levels of care when there is uncertainty about the negligence determination, as discussed in section 6.2. Hence, the use of punitive damages seems socially detrimental as a general matter.

There is, however, a fairly general circumstance in which damages exceeding liability are desirable: when injurers sometimes escape liability. This possibility may arise because injurers are hard to identify as the sources of harm (the origin of pollution may be difficult to trace) or because victims do not choose to bring suit (litigation costs may discourage legal action). If injurers who ought to be found liable for harm  $h$  are in fact only found liable and made to pay damages with probability  $q$ , then if damages are raised to  $(1/q)h$ , injurers' expected liability will be  $h$ . Thus, the more likely a party is to escape liability, the higher should be damages when the party is found liable. Accordingly, a firm that dumps toxic wastes at night, or an individual who tries to conceal a bad act, should have to pay punitive damages, but not an injurer who causes harm in a noticeable way. On these points and others, see, for example, Cooter (1989), Diamond (2002), and Polinsky and Shavell (1998).<sup>31</sup>

Another assumption that would justify damages exceeding harm is that injurers obtain "illicit" utility—utility not credited in social welfare—from certain harmful acts, such as spitefully destroying a person's property. Under this assumption, it is readily shown that damages exceeding harm may be desirable, for they are required to induce injurers not to commit the harmful acts, in order to offset their illicit utility. The assumption of illicit utility might be considered problematic, however, as it implies that social welfare

<sup>31</sup> See also Daughety and Reinganum (1997). For empirical study of punitive damages, see Eisenberg et al. (1997), Karpoff and Lott (1999), and Polinsky (1997).

is not a function of individuals' utilities (and thus leads to conflicts with the Pareto principle<sup>32</sup>).

### 8.5. Accuracy of damages and legal costs

Much expense is incurred in litigation about the magnitude of a victim's harm, which raises the question of what the social value of greater accuracy is and whether the private value of accuracy is different from the social value. It turns out that the private value of information about harm tends to exceed the social value. To explain, there is social value in establishing harm accurately primarily when injurers know, at the time that they choose their level of care, how much harm they might cause, for then they can take appropriate steps to prevent harm. For example, if an injurer anticipates that the atypically large harm he might cause will be accurately measured, he will exercise a properly high degree of care, as is socially desirable. However, injurers often lack considerable information about the harm they might cause when they decide on their precautions. Drivers, for example, know relatively little about how much harm a potential victim would suffer in an accident (the seriousness of injuries, the magnitude of lost earnings). In such cases, there is limited social value in assessing damages accurately because injurers behavior would not be much affected by doing so; drivers' incentives to avoid accidents would be largely the same if, instead of striving for fairly precise measurements of harm, courts employed rough averages (based, perhaps, upon abbreviated litigation over damages or upon figures from a table).<sup>33</sup> Regardless of the possibly attenuated social value of accuracy in the measure of damages, victims and injurers have very strong private incentives to spend to establish damages accurately in court. A victim will always be willing to spend up to a dollar to prove that harm is a dollar higher, and an injurer will always be willing to spend up to a dollar to prove that harm is a dollar lower. This suggests why, as shown in [Kaplow and Shavell \(1996b\)](#) the private value of information about damages exceeds the social, so that too much is spent by litigants proving damages.

## 9. Multiple injurers

In some contexts more than one injurer jointly contributes to harmful events, such as when two drivers race each other and cause another car to drive off the road, or where two firms operate a refinery which explodes. For simplicity, let us consider the unilateral model of accidents and that there are just two injurers (the case of  $n$  injurers is essentially identical). Let  $p(x_1, x_2)$  be the probability of harm, so that the social problem is

<sup>32</sup> See [Kaplow and Shavell \(2001\)](#).

<sup>33</sup> A qualification to this point, emphasized by [Spier \(1994\)](#), arises where the probability distribution of harm is affected by an injurer's degree of care. If this is so, then accuracy in assessing harm will influence an injurer's incentives to reduce risk even when, at the time he chooses his level of care, he does not have information about the harm that would occur in an accident.

to minimize  $x_1 + x_2 + p(x_1, x_2)h$ , and let  $x_i^*$  denote the optimal  $x_i$ , which are assumed to be positive.

### 9.1. *Strict liability*

By strict liability is meant imposition of damages on injurers in amounts that are independent of their levels of care and that the sum of damages equals  $h$  (the usual rule).

Suppose that the injurers act independently of one another. Then the optimal outcome is not achievable, since to induce optimal care of  $x_i^*$  injurer  $i$  must bear damages of  $h$  under strict liability.

If injurers act in concert, then it is obvious that they will behave optimally since their problem will be the social problem.

### 9.2. *Negligence rule*

By the negligence rule is meant the rule under which each injurer escapes liability if his care is at least his due care level  $\underline{x}_i$ ; if he alone is negligent, he bears damages of  $h$ , and if he and the other injurer are both negligent, they will jointly pay  $h$ , in amounts that may depend on their  $x_i$ .

Suppose that injurers act independently of each other. Then, if due care levels are optimal,  $\underline{x}_i = x_i^*$ , it is clear (from the argument in section 2.1) that each injurer taking due care is an equilibrium, and it can be shown that it is the unique equilibrium. Thus the negligence rule is superior to strict liability. The superiority of the negligence rule has to do with the fact that, by the nature of the negligence rule, each party is threatened with damages equal to the *entire* harm  $h$  when other parties act optimally; whereas under strict liability, each party is threatened with a lesser amount.

If injurers act in concert, then they will also act optimally if due care levels are optimal.

These points about behavior of multiple injurers under the negligence rule were first made by W.M. Landes and Posner (1980).

## 10. The judgment-proof problem

The possibility that injurers may not be able to pay in full for the harm they cause is known as the judgment-proof problem and is of substantial importance, for individuals and firms often pose risks significantly exceeding their assets (a person of modest means could cause a devastating fire; a small firm's product could cause many deaths). The judgment-proof problem reduces the effectiveness of accident liability in combating risk and also lowers the incentive to purchase liability insurance. On the judgment-proof problem, see originally Summers (1983) and Shavell (1986).

### 10.1. Incentives to engage in harmful activities

When injurers are unable to pay fully for the harm they may cause, their incentives to engage in risky activities will be greater than otherwise. Under strict liability, their incentives to engage in an activity are optimal if their assets are enough to cover the harm they might cause, but their incentives will be inadequate if they are unable to pay for the harm. Under the negligence rule, the situation is somewhat different. If the rule functions without error and injurers are induced to take care (as they might despite being judgment proof—see below), they bear no liability and their incentives to engage in the activity (which are too great, by the argument of section 2.3) are not altered by the existence of the judgment-proof problem. If, though, injurers are not induced to take optimal care, or there are errors in the negligence determination that sometimes result in findings of negligence, then the existence of the judgment-proof problem does lead injurers to engage more often in the activity than they otherwise would.

### 10.2. Incentives to take care

Likewise, when injurers are unable to pay for all the harm they might cause, their incentives to take care tend to be diluted. Consider the injurer's problem of choosing care  $x$  under strict liability, when his assets are  $y < h$ . The injurer's problem is often formulated as minimizing  $x + p(x)y$ ; hence, the injurer chooses  $x(y)$  determined by  $-p'(x)y = 1$  instead of  $-p'(x)h = 1$ , so that  $x(y) < x^*$  (and the lower is  $y$ , the lower is  $x(y)$ ). Under this formulation of the injurer's problem, note that it is implicitly assumed that care is nonmonetary, for if care is monetary, the injurer's wealth after spending on care would be  $y - x$ , and only this amount would be left to be paid in a judgment. If care is monetary, in other words, the injurer's problem is to minimize  $x + p(x)(y - x)$ . The solution  $x(y)$  to this problem is determined by the condition  $-p'(x)(y - x) = 1 - p(x)$ . The interpretation is that the marginal benefit of taking care, saving the person's remaining wealth of  $y - x$  with a higher probability of  $-p'(x)$ , equals the marginal cost of the expenditure, which is not 1 but  $1 - p(x)$ . In effect, spending on care has a private cost of less than the social cost of 1, since there is a chance  $p(x)$  that an expenditure will not be a private cost as it would be paid in a judgment. This gives rise to the theoretical possibility that care could be excessive,  $x(y) > x^*$ ; but that would not be the case when  $p$  is low. For simplicity, I will assume below that care is nonmonetary.

Under the negligence rule, with due care  $\underline{x}$  set equal to  $x^*$ , the injurer might choose  $x^*$  rather than risk liability with a lower  $x$  and incur costs of  $x + p(x)y$ ; indeed, since  $x^* + p(x^*)h < x^*$ , there exists an interval  $[h - k, h]$  with positive  $k$  such that if  $y$  is in this interval,  $x^*$  will be induced. So, unlike with strict liability, the judgment-proof problem does not lead to inadequate care under the negligence rule unless assets  $y$  are sufficiently less than harm  $h$ .

### *10.3. Incentives to purchase liability insurance*

Risk-averse injurers who may not be able to pay for the entire harm they cause will tend not to purchase full liability insurance or any at all. This is because purchase of full coverage would require them to pay premiums for damage payments that they would not make in the absence of coverage: if a person with assets of \$10,000 buys coverage against liability of \$100,000, he is purchasing coverage against \$90,000 of losses that he would not suffer if he did not have coverage.<sup>34</sup> The nature and consequences of this effect of the judgment-proof problem depend, among other things, on whether liability insurers have information about risk and thus link premiums to it. If liability insurers have information about risk and base premiums on it, then reduction in the purchase of liability insurance tends to reduce incentives to take care (as in the last section). If liability insurers do not have information about risk, then reduced purchase of liability insurance tends to increase incentives to take care (even though the incentives generally remain suboptimal). In both cases, reduction in the purchase of liability insurance tends to undesirably increase incentives to engage in the harmful activity.

### *10.4. Policies to ameliorate the judgment-proof problem*

Several types of policy responses to the dilution of incentives caused by the judgment-proof problem are of interest. If there is a second party who has some control over the behavior of the party whose assets are limited and who directly affects risk, then the second party can be held vicariously liable for the losses caused by the first, inducing the vicariously liable party to cause risks to be lowered. Thus, holding a large contractor liable for the accidents caused by a small subcontractor or an employer for accidents caused by its employees will lead the former to control the risks posed by the latter. See section 11.

Additionally, parties with assets less than a specified amount could in some contexts be prevented from engaging in an activity. However, such minimum asset requirements may undesirably prevent some individuals who ought to engage in the activity from doing so; although their assets are low and their care would be inadequate, their benefits might still exceed the expected harm that they create. Minimum asset requirements are a somewhat blunt instrument for alleviating the incentive problems under consideration.<sup>35</sup>

A third response to inadequate incentives caused by the judgment-proof problem is regulation of liability insurance. One form of insurance regulation would mandate purchase of (perhaps full) liability insurance coverage. This approach would be especially appealing when insurers can observe the precautions taken by injurers and link premiums to it. In that case, if injurers purchase full liability insurance coverage, their incentives to reduce risk would be optimal (see section 3.3). However, if insurers cannot

<sup>34</sup> See Huberman et al. (1983), Keeton and Kwerel (1984), and Shavell (1986).

<sup>35</sup> On minimum asset requirements, see Shavell (2005).



observe risk and moral hazard exists, then requiring the purchase of liability insurance further reduces the incentive to take care. Therefore, mandating the purchase of liability coverage may not be desirable. This point suggests that an opposite form of insurance regulation may be advantageous: barring purchase of liability insurance (a policy that is sometimes employed, as noted in section 3.3). Forbidding purchase of liability insurance can improve incentives to take care if, without a prohibition, injurers would have purchased positive coverage and insurers cannot observe injurers' level of care.<sup>36</sup>

Fourth, the use of corrective taxes equal to expected harm may help to alleviate the judgment-proof problem. When harm is caused with a low probability, the expected harm is much less than the actual harm. Hence, parties with limited assets may be able to pay the appropriate tax on risk-creating behavior even though they could not pay for the harm itself. Suppose that the release of pollution will cause harm of \$1 million with a 1% probability. Thus, a firm with \$100,000 of assets would be able to pay \$10,000 for the expected harm it would cause were it to cause the pollution, even though it would only be able to pay one tenth of the actual \$1 million harm it might generate, and so its incentives to reduce risk would be much too low under the liability system.

A fifth way of correcting for dilution of incentives is for the state to regulate parties' behavior directly, such as with traffic laws or safety requirements. Regulation, however, may involve inefficiency because of regulators' limited knowledge of risk and of the cost and ability to reduce it.

Another way of mitigating dilution of incentives is resort to criminal liability. A party who would not take care if only his assets were at stake might be induced to do so for fear of imprisonment.

## 11. Vicarious liability

Vicarious liability, as mentioned in the last section, is the imposition of liability on a party related to the actual author of harm, where the vicariously liable party usually has some control over the party who directly causes the harm. The forms of vicarious liability vary widely and include holding parents liable for harms caused by their children, contractors for harms caused by subcontractors, firms for the harms caused by employees (a particularly important example), and even, occasionally, lenders for the harms caused by borrowers.

There are two major reasons that vicarious liability may be socially desirable. One is that the injurer may not have proper information about reduction of harm, whereas the vicariously liable party may have good, or at least superior, information and be able to influence the risk-reducing behavior of the injurer. This feature of vicarious liability might apply, for example, in regard to a parent in relation to a child (suppose a teenage child does not realize how dangerous using a motorboat might be to swimmers and

<sup>36</sup> On regulation of liability insurance, see Jost (1996), Polborn (1998), and Shavell (1987, 2000, 2005).

that the parent can monitor the child's use of the boat), or in regard to a firm and its employees (suppose a worker does not understand how dangerous a toxic waste product is, but the firm does and thus controls how it is transported by the worker to a dump site).

The other basic reason that vicarious liability may be desirable is that it helps to ameliorate the judgment-proof problem as it applies to the injurer, as noted in the previous section.<sup>37</sup> Under vicarious liability, the vicariously liable party's assets are at risk as well as the injurer's, giving the vicariously liable party a motive to reduce risk or to moderate the injurer's level of activity. There are various ways in which the vicariously liable party can affect risk or the level of activity. (a) Vicariously liable parties may be able to affect the behavior of injurers in some direct manner, such as where a parent prevents a child from using a speedboat, where an employer does not allow an employee to transport hazardous wastes until the employee has been through a training course in handling the wastes, or where a contractor requires a subcontractor to use a particular safety device. Here, note, the control by the vicariously liable party may be over the level of care (as in the case of the employee who must take a training course) or over the level of activity (as in the case of the child who is prevented from using the speedboat). (b) Another possibility is that vicariously liable parties can themselves take precautions that alter the risk that injurers present. For example, the employer can purchase a tanker truck to transport hazardous wastes that is safer than some other kind of tanker truck (perhaps it has a double-walled container). (c) Sometimes vicariously liable parties may control participation in activities because they act as "gatekeepers"; they are able to prevent injurers from engaging in their activity by withholding financing or a required service. Thus, a vicariously liable lender or parent corporation may withhold funding to an injurer; or a vicariously liable architect may not agree to sign a statement that a hotel is safe from collapse in an earthquake and thus effectively deny the hotel from doing business if such a statement is legally required.

The main disadvantage of vicarious liability is that it increases litigation costs, since it means that vicariously liable parties can be sued as well as the injurer. The increase in litigation cost must be weighed against the improvement in incentives. Hence, imposing vicarious liability on a homeowner for an accident caused by a delivery truck that is bringing something to his home would probably not be worthwhile, since the extra incentive gained would be modest. Another disadvantage is that it leads potential contracting parties, who might become vicariously liable were they to make a contract, to expend effort determining the degree of their liability exposure.<sup>38</sup>

<sup>37</sup> Both reasons often apply jointly; for instance, children and employees are often judgment proof as well as in possession of less information than vicariously liable parties. Other reasons for vicarious liability are sometimes emphasized. For example, where a firm is held vicariously liable for its employees, a victim need not have to prove which employee actually caused the harm suffered.

<sup>38</sup> Another disadvantage, of a subtle nature, is that a vicariously liable contracting party will charge a party with low assets an amount reflecting the vicariously liable party's liability exposure; this charge will itself reduce the assets of the low asset party, exacerbating the judgment proof problem in regard to him; see Pitchford (1995).

Literature on vicarious liability originates with Kornhauser (1982) and Sykes (1981, 1984).<sup>39</sup>

## 12. Nonpecuniary harm

It is often the case that the victim of an accident suffers a nonpecuniary harm, such as when an irreplaceable photo album is destroyed in a fire, a person suffers a disabling injury, or a scenic beach is ruined by an oil spill. In these cases, the harm is called nonpecuniary because the loss is not replaceable with a good or service that can be purchased on a market. Let us sketch the implications of such nonpecuniary harms for the analysis of the liability system.

### 12.1. Incentives in the risk-neutral case

Suppose initially that individuals are risk neutral. In this case, it is natural to assume that the victim's loss is the sum of a pecuniary component  $h$  and a nonpecuniary component  $t$ . Thus, the optimal level of care  $x^*$  is the  $x$  minimizing  $x + p(x)(h+t)$ ; the optimal level  $z^*$  at which to engage in the activity is the  $z$  maximizing  $b(z) - z[x^* + p(x^*)(h+t)]$ ; and so forth. Hence, it is evident that the nonpecuniary harm should influence damages and the due care standard: the optimal level of damages is  $d = h + t$ ; and due care should be  $x^*$  and not the lower level of care minimizing only  $x + p(x)h$ .

However, courts often face an informational problem in determining nonpecuniary harm  $t$ , and one that is more difficult than that experienced in determining pecuniary harm. Pecuniary harm is measurable since, by definition, it is equal to the cost of purchasing a replacement good. If a car is damaged beyond repair, the same car (or essentially the same one) can generally be purchased on the market, so the court needs only to estimate the cost of buying the replacement vehicle. In contrast, nonpecuniary harm is not usually associated with a market observable and, moreover, varies, perhaps greatly, across individuals. The disutility of losing an irreplaceable family photo album might be very different from one person to the next. Thus, courts are likely to err in computing damages for nonpecuniary harm (and in ascertaining the proper due care level when such harm is involved).

### 12.2. Incentives and insurance in the risk-averse case; fines as a supplement to liability

Now suppose that victims are risk averse and for simplicity that a victim's utility is of the form  $u(y)$  plus a nonpecuniary utility component, where  $u(y)$  is utility from wealth

<sup>39</sup> See also Boyd and Ingberman (1997), Boyer and Laffont (1997), Feess (1999), Hansmann and Kraakman (1991), Kraakman (1986, 1998), Pitchford (1995), and Shavell (1987, pp. 170–177, 182–185).

$y$  and  $u$  is concave. If there is an accident causing a pecuniary harm of  $h$  and a utility loss of  $t$ , the victim's utility would be  $u(y - h) - t$ ; whereas if there is no accident, his utility is  $u(y)$ . Before considering liability and incentives, observe that the optimal insurance policy for a victim facing a loss of  $h$  and  $t$  is a policy that covers only the pecuniary loss  $h$ . For, as is well known, optimal insurance coverage is such that the marginal utility of wealth is the same if an accident occurs as if it does not; and since the utility loss  $t$  does not affect the utility of wealth, purchasing insurance against it would only lower expected utility.<sup>40</sup> Another way to put this point is that although nonpecuniary losses may involve substantial disutility (consider the death of a child or a parent), they may not create a need for money (do not raise the marginal utility of money), so it makes no sense to insure against them.

This point about the optimal insurance purchases against nonpecuniary losses suggests that the first-best solution to the accident problem when losses involve nonpecuniary components and victims are risk averse has the following character. Victims receive an amount equal to their pecuniary losses alone; but injurer risk-reducing behavior reflects nonpecuniary as well as pecuniary harm.

The first-best solution cannot be achieved under strict liability for the simple reason that if victims receive optimal insurance coverage equal to pecuniary harm, incentives of injurers will be too low; and if injurers pay damages sufficient to give them proper incentives, victims' insurance coverage will be excessive. The second-best level of damages strikes an implicit compromise between the goal of providing proper incentives to injurers and optimal insurance for victims.

However, the first-best solution can be achieved under strict liability if damages are supplemented by fines collected by the state. For then damages can be set equal to the pecuniary harm  $h$ , optimally insuring victims, and the fine can be set high enough so as to provide, together with damages, optimal incentives to reduce risk for injurers. (One way to appreciate the superiority of the system of fines is that it can be shown to be Pareto superior to any liability system. In a fully-specified model, the fine revenue can benefit individuals, such as by reducing their taxes or by allowing a lump sum payment to be made to individuals *ex ante*.)

Under the negligence rule, if it is perfectly functioning, the first-best solution can be achieved without the use of fines. For if damages are set high enough to induce due care of  $x^*$ , victims will bear their losses and will purchase accident insurance covering only their pecuniary harm  $h$ . However, if there are errors in the negligence determination, then, as under strict liability, a system with fines as a supplement to liability will tend to be superior to the pure liability system.

The use of supplemental fines would tend to resolve certain tensions arising from application of standard principles of tort law in calculating damages. It is frequently observed, for example, that because damages for death tend to be based on the present value of foregone earnings, payments for the death of children (whose earnings are

<sup>40</sup> See Arrow (1974) and Cook and Graham (1977).

many years in the future) or for the elderly are low. Such damages lead to inadequate deterrence (and also may seem jarring to the public), even though the damages are not inappropriate from the standpoint of insurance (the death of a child or of an elderly person usually does not create the need for money). The use of fines could help to increase deterrence without altering what victims receive.<sup>41</sup>

## **Part C: Methods of controlling risk**

### **13. Liability versus other methods of controlling risk**

#### *13.1. Methods of control*

Liability for accidents is one among a number of methods that society may employ to control potentially harmful behavior. Another is direct regulation, under which the state restricts permissible behavior, for example, by requiring that factories install smoke arrestors. Closely related to state regulation is the legal 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; the injunction is effectively privately-initiated regulation. Society can also make use of the corrective tax, under which a party pays the state an amount equal to the expected harm he causes, a standard example being the expected harm due to the discharge of a pollutant into a lake.<sup>42</sup> In fact, liability and regulation are the preeminent tools that society utilizes to control risky activities; use of corrective taxes is unusual in a relative sense.

#### *13.2. Comparison*

I will now sketch several factors bearing on the relative desirability of these methods of controlling externalities. The review of factors will show that any of the methods (or a combination) could be the best, depending on the context.<sup>43</sup>

One factor of relevance is the quality of the state's information. If the state has complete information, that is, knows injurers' benefits from acts, costs of precautions and their effectiveness, and the harm that injurers would cause, then all of the methods of

<sup>41</sup> The subject of liability, risk-bearing, and non-pecuniary harms was first studied in Spence (1977), who pointed out the advantage of fines as a supplement to liability. See also Calfee and Rubin (1992), Shavell (1987, pp. 231–235, 247–254), and Viscusi and Evans (1990).

<sup>42</sup> Other tools include subsidies for taking precautions and fines for causing harm, but for simplicity only the four methods mentioned in the text are considered here.

<sup>43</sup> This section follows Shavell (1993). There are a number of articles that focus on comparisons between pairs of methods of controlling externalities. See Calabresi and Melamed (1972), Ellickson (1973), Kaplow and Shavell (1996a, 2002), Kolstad et al. (1990), Polinsky (1980a), Shavell (1984a, 1984b), Weitzman (1974), and Wittman (1977).

control allow achievement of optimality. If the state's information is imperfect, the state will not be able to calculate which actions (such as installing a smoke arrestor) are desirable and thus sometimes will err. Hence, methods of control that require the state to determine optimal actions—namely, regulation, the injunction, and liability based on the negligence rule—will not perform well. In contrast, as long as the state can ascertain expected harm, we know that it can induce injurers to act optimally under the corrective tax; and as long as the state can determine the actual harm, we know that it can induce optimal behavior under strict liability; for under either the corrective tax or strict liability, injurers will balance their costs of precautions and benefits from engaging in activities against the reduction in expected harm that would be brought about by taking precautions and modifying their levels of activity.

This basic informational argument favoring corrective taxes or strict liability over regulation, the injunction, and the negligence rule applies despite possible uncertainty about the magnitude of harm. The reason, essentially, is that under corrective taxes, the state only needs to estimate expected harm and set the tax equal to it, and under strict liability, the state only needs to allow harm to occur and set damages equal to actual harm, whereas under the other methods the state must also estimate precaution costs and/or injurers' benefits from engaging in their activities. The present point about the superiority of corrective taxes (and of strict liability) holds notwithstanding Weitzman (1974), which suggests that regulation may be superior to corrective taxation.<sup>44</sup> The state's information also bears on the relative appeal of the corrective taxes versus strict liability. Under the corrective tax, the expected harm must be known, such as the expected harm that would be caused by pollution or by leaving ice on sidewalks. Under strict liability, only the actual harm need be known. When, as is often the case, actual harm is more readily measured than expected harm (arguably the case for icy sidewalks), strict liability enjoys an informational advantage over the corrective tax. In some situations, though, expected harm may be easier to determine than actual harm (especially when harm is difficult to trace to its origin).

A second factor of relevance to the comparison of methods of control is the information available to victims. For many externalities, victims have better information than the state about who is causing harm or about its extent—because they are the parties who actually suffer the harm. Hence, one would suppose that victims are the most appropriate enforcement agents, suggesting the desirability of the liability tool or the injunction. For some externalities, however, victims may have poor information about harm, worse than the state's. Consider, for instance, that victims may be unaware that harm is even occurring when they are exposed to odorless and invisible carcinogenic agents. In such circumstances, the state would often be a better enforcer than individuals, and use of regulation or corrective taxes might be advantageous.

<sup>44</sup> Weitzman's (1974) conclusion that regulation could be superior to taxation rests on the assumption that the tax rate does not depend on the quantity of pollution. If, as would usually be easy to implement, the tax rate can vary with the quantity of pollution, taxes are superior to regulation; see Roberts and Spence (1976) and Kaplow and Shavell (2002).

A third factor concerns the level of activity of an injurer (how much a firm produces, how many miles a person drives), as opposed to the precautions an injurer takes given the level of activity (whether a firm uses a smoke scrubber while producing, whether a person exercises care when driving). Regulation and the negligence rule are most often concerned with precautions taken but not with the level of activity (a firm may be required by regulation to install smoke scrubbers or to take certain safety precautions but not to reduce its output). Thus injurers may not have incentives to moderate their level of activity although that would be desirable (their activity may result in harm despite the exercise of optimal precautions—even with smoke scrubbers, some pollution will result). In contrast, under the corrective tax and strict liability, injurers pay for harm done, so that they will optimally moderate their level of activity (as well as efficiently choose their level of precautions).

A fourth pertinent factor, noted above, is the ameliorative behavior of victims. Under regulation, the injunction, and corrective taxation, victims are not compensated for their harm, so that victims have a natural incentive to take optimal precautions or to alter their activity. Similarly, under the negligence rule, victims have an incentive to take precautions, since injurers will tend to be led to behave nonnegligently so will not compensate victims. Under strict liability, victims incentives are not as good; although the defense of contributory negligence gives victims an incentive to take precautions, they are not led to moderate their level of activity (or any aspect of behavior not included in the contributory negligence determination).

Still another factor is administrative costs, the costs borne by the state in applying a legal rule and the legal and related costs borne by the affected parties. Liability rules possess a general administrative cost advantage over regulation in that under liability rules, administrative costs are incurred only if harm is done. This advantage may be significant when the likelihood of harm is small. Nevertheless, administrative costs will sometimes be lower under other approaches. For example, compliance with a regulation may readily be detected in some circumstances (determining whether factory smokestacks are sufficiently high would be easy) and also may be accomplished through random monitoring, saving enforcement resources. Also, imposing corrective taxes might be inexpensive. Notably, suppose that they are levied at the time of the purchase of a product. In contrast, liability rules might be expensive to employ. For example, demonstrating the source of a particular harm and its extent may be difficult. Also, when industrial pollution affects millions of individuals on an ongoing basis, the cost of a continuous flow of individual suits (or even class actions) that measure damages victim-by-victim is likely to be in excess of the cost of alternatives.

Last, the ability of injurers to pay for harm is of relevance. For liability rules to induce potential injurers to behave appropriately, injurers must have assets sufficient to make the required payments; otherwise they will have inadequate incentives to reduce harm, as discussed in section 10. Where inability to pay is a problem, bonding requirements may be helpful, and regulation may become more appealing (although it may need to be enforced through the threat of nonmonetary, criminal sanctions). In addition, corrective taxes have an advantage over liability rules when harm is probabilistic because, under

the corrective tax, an injurer would pay only the expected harm (with certainty) rather than the higher actual harm (with a probability); see section 10.4.

### *13.3. Resolution of externalities through bargaining by affected parties*

Parties affected by unregulated externalities will sometimes have the opportunity to make mutually beneficial agreements with those who generate the externalities. In the classic example, if a firm's pollution causes harm of \$1,000 that can be prevented by installing a smoke scrubber that costs \$100, then, in the absence of any legal obligation on the firm, one might expect a potential victim of pollution to pay the firm an amount between \$100 and \$1,000 to install the scrubber.<sup>45</sup>

If it is posited that there are no obstacles to reaching a mutually beneficial agreement concerning externalities, then that will occur. This tautology is one version of the Coase Theorem; Coase (1960) stressed the point that externality problems could be remedied through private bargains. A closely related version of the Coase Theorem asserts that the outcome regarding the externality—whether a smoke scrubber is installed or instead pollution is generated—does not depend on the legal rule that applies. For example, if the scrubber costs \$100 and there is no law that controls pollution, a bargain as we have described it will come about and the scrubber will be installed; and likewise if there is a law that leads to installation of the scrubber, the same will happen. The outcome, however, might be affected by the legal rule because of the level of wealth of parties. Most obviously, the potential victims might not have assets sufficient to pay for the scrubber, in which case the scrubber would not be installed unless a legal rule leads to this; moreover, legal rules may affect the distribution of wealth and thus the demand for goods, including that of being free from pollution.<sup>46</sup>

In any case, there are many obstacles to bargaining. Bargaining may fail to occur when victims are numerous and face collective action problems in coming together. This is often the situation with respect to victims of industrial pollution, for example. Similarly, in important contexts, bargaining will be impractical because victims will not know in advance who will injure them; this is the case for automobile accidents and most other accidents between strangers. Another reason that bargaining may not occur is that victims might not know that they are exposed to a risk. Also, the cost of bargaining between just one potential victim and one potential injurer who know of each other can discourage them from engaging in the process. If these reasons do not apply and victims and injurers do engage in bargaining, asymmetry of information may lead to bargaining impasses. Altogether, these problems that reduce the likelihood of bargaining occurring, and also its success if it does take place, make the importance of legal rules to remedy externalities substantial.

<sup>45</sup> For experimental studies on bargaining and entitlements, see, for example, Hoffman and Spitzer (1982) and Croson and Johnston (2000).

<sup>46</sup> The outcome following from a legal rule might also be affected by an "endowment effect," under which individuals' valuations depend on whether or not they originally enjoy legal protection. See, for example, Kahneman et al. (1990).



## 14. Conclusion

The subject of liability for accidents as a device for controlling participation in risky activities and for inducing the taking of precautions when parties engage in them has been reviewed here. Economic analysis of this subject has been undertaken for only a relatively short time, and many avenues for theoretical research are apparent, concerning, for example, the relationship between liability insurance and the liability system, the ramifications of the judgment-proof problem, and the effect of liability on investigation of risks. More generally, as the discussion in section 13 should have indicated, analysis of the comparative strengths of (and of complementarities between) liability and other devices for controlling accidents seems to be of particular importance.

## References

- Arrow, K.J. (1974). "Optimal insurance and generalized deductibles". *Scandinavian Actuarial Journal* 1974, 1–42.
- Bar-Gill, O., Ben-Shahar, O. (2003). "The uneasy case for comparative negligence". *American Law and Economics Review* 5, 433–469.
- Boyd, J., Ingberman, D.E. (1997). "The search for deep pockets: is "extended liability expensive liability?" *Journal of Law, Economics, and Organization* 13, 232–258.
- Boyer, M., Laffont, J.-J. (1997). "Environmental risks and bank liability". *European Economic Review* 41, 1427–1459.
- Brown, J.P. (1973). "Toward an economic theory of liability". *Journal of Legal Studies* 2, 323–349.
- Calabresi, G. (1970). *The Costs of Accidents*. Yale University Press, New Haven.
- Calabresi, G. (1975). "Concerning cause and the law of torts". *University of Chicago Law Review* 43, 69–108.
- Calabresi, G., Melamed, A.D. (1972). "Property rules, liability rules, and inalienability: one view of the cathedral". *Harvard Law Review* 85, 1089–1128.
- Calfee, J.E., Rubin, P.H. (1992). "Some implications of damage payments for nonpecuniary losses". *Journal of Legal Studies* 21, 371–411.
- Coase, R.H. (1960). "The problem of social cost". *Journal of Law and Economics* 3, 1–44.
- Cook, P., Graham, D. (1977). "The demand for insurance and protection: the case of irreplaceable commodities". *Quarterly Journal of Economics* 91, 143–156.
- Cooter, R.D. (1989). "Punitive damages for deterrence: when and how much?" *Alabama Law Review* 40, 1143–1196.
- Cooter, R.D., Ulen, T.S. (1986). "An economic case for comparative negligence". *New York University Law Review* 61, 1067–1110.
- Craswell, R., Calfee, J.E. (1986). "Deterrence and uncertain legal standards". *Journal of Law, Economics, & Organization* 2, 279–303.
- Crosan, R., Johnston, J.S. (2000). "Experimental results on bargaining under alternative property rights regimes". *Journal of Law, Economics, & Organization* 16, 50–73.
- Cummins, J.D., Phillips, R.D., Weiss, M.D. (2001). "The incentive effects of no-fault automobile insurance". *Journal of Law and Economics* 44, 427–464.
- Danzon, P.M. (1985). *Medical Malpractice: Theory, Evidence, and Public Policy*. Harvard University Press, Cambridge, MA.
- Daughety, A.F., Reinganum, J.F. (1997). "Everybody out of the pool: products liability, punitive damages, and competition". *Journal of Law, Economics, & Organization* 13, 410–432.
- Deweese, D., Duff, D., Trebilcock, M. (1996). *Exploring the Domain of Accident Law: Taking the Facts Seriously*. Oxford University Press, New York.

- Diamond, P.A. (1974). "Single activity accidents". *Journal of Legal Studies* 3, 107–164.
- Diamond, P.A. (2002). "Integrating punishment and efficiency concerns in punitive damages for reckless disregard of risks to others". *Journal of Law, Economics, & Organization* 18, 117–139.
- Dobbs, Dan, B. (2000). *The Law of Torts*. West, St. Paul.
- Edlin, A. (1994). "Efficient standards of due care: should courts find more parties negligent under comparative negligence?". *International Review of Law and Economics* 14, 21–34.
- Eisenberg, T., Goerd, J., Ostrom, B.J., Rottman, D., Wells, M.T. (1997). "The predictability of punitive damages". *Journal of Legal Studies* 26, 623–661.
- Ellickson, R.C. (1973). "Alternatives to zoning: covenants, nuisance rules, and fines as land use controls". *University of Chicago Law Review* 40, 681–781.
- Emons, W. (1990). "Efficient liability rules for an economy with non-identical individuals". *Journal of Public Economics* 42, 89–104.
- Emons, W., Sobel, J. (1991). "On the effectiveness of liability rules when agents are not identical". *Review of Economic Studies* 58, 375–390.
- Farber, H.S., White, M.J. (1991). "Medical malpractice: an empirical examination of the litigation process". *Rand Journal of Economics* 22, 199–217.
- Feess, E. (1999). "Lender liability for environmental harm: an argument against negligence based rules". *European Journal of Law and Economics* 8, 231–250.
- Goldberg, V.P. (1974). "The economics of product safety and imperfect information". *Bell Journal of Economics* 5, 683–688.
- Grady, M.F. (1983). "A new positive economic theory of negligence". *Yale Law Journal* 92, 799–829.
- Green, J. (1976). "On the optimal structure of liability laws". *Bell Journal of Economics* 7, 553–574.
- Grossman, S.J. (1981). "The informational role of warranties and private disclosure about product quality". *Journal of Law and Economics* 24, 461–483.
- Hamada, K. (1976). "Liability rules and income distribution in product liability". *American Economic Review* 66, 228–234.
- Hansmann, H., Kraakman (1991), R. (1991). "The uneasy case for limiting shareholder liability in tort". *Yale Law Journal* 100, 1879–1934.
- Hoffman, E., Spitzer, M. (1982). "The Coase theorem: some experimental tests". *Journal of Law and Economics* 25, 73–98.
- Huberman, G., Mayers, D., Smith, C. (1983). "Optimal insurance policy indemnity schedules". *Bell Journal of Economics* 14, 415–426.
- Hylton, K.N. (1990). "The influence of litigation costs on deterrence under strict liability and under negligence". *International Review of Law and Economics* 10, 161–171.
- Jerry, R.H. (1996). *Understanding Insurance Law*. Matthew Bender, New York.
- Jost, P.-J. (1996). "Limited liability and the requirement to purchase insurance". *International Review of Law and Economics* 16, 259–276.
- Kahan, M. (1989). "Causation and incentives to take care under the negligence rule". *Journal of Legal Studies* 18, 427–447.
- Kahneman, D., Knetsch, J.L., Thaler, R.H. (1990). "Experimental tests of the endowment effect and the Coase theorem". *Journal of Political Economy* 98, 1325–1348.
- Kakalik, J.S., Pace, N.M. (1986). *Costs and Compensation Paid in Tort Litigation*, Report R-3391-ICJ. Rand Corporation, Santa Monica, CA.
- Kaplow, L. (1986). "Private versus social costs in bringing suit". *Journal of Legal Studies* 15, 371–385.
- Kaplow, L., Shavell, S. (1996a). "Property rules versus liability rules: an economic analysis". *Harvard Law Review* 109, 713–790.
- Kaplow, L., Shavell, S. (1996b). "Accuracy in the assessment of damages". *Journal of Law and Economics* 39, 191–210.
- Kaplow, L., Shavell, S. (2001). "Any non-welfarist method of policy assessment violates the Pareto principle". *Journal of Political Economy* 109, 281–286.
- Kaplow, L., Shavell, S. (2002). "On the superiority of corrective taxes to quantity regulation". *American Law and Economics Review* 4, 1–17.

- Karpoff, J.M., Lott, J.R. (1999). "On the determinants and importance of punitive damage awards". *Journal of Law and Economics* 42, 527–573.
- Keeton, W.R., Kwerel, E. (1984). "Externalities in automobile insurance and the underinsured driver problem". *Journal of Law and Economics* 27, 149–179.
- Kessler, D., Rubinfeld, D.L. (2007). "Empirical Study of the Common Law and Legal Process". In: Polinsky, A.M., Shavell, S. (Eds.), *Handbook of Law and Economics*, vol. 1. North-Holland, Amsterdam.
- Kolstad, C.D., Ulen, T.S., Johnson, G.V. (1990). "Ex post liability for harm vs. ex ante safety regulation: substitutes or complements?" *American Economic Review* 80, 888–901.
- Kornhauser, L.A. (1982). "An economic analysis of the choice between enterprise and personal liability for accidents". *California Law Review* 70, 1345–1392.
- Kraakman, R. (1986). "Gatekeepers: the anatomy of a third-party enforcement strategy". *Journal of Law, Economics, & Organization* 2, 53–104.
- Kraakman, R. (1998). "Limited shareholder liability". In: Newman, P. (Ed.), *The New Palgrave Dictionary of Economics and the Law*, vol. 3. Macmillan, London, pp. 648–654.
- Landes, E. (1982). "Insurance, liability, and accidents: a theoretical and empirical investigation of the effect of no-fault accidents". *Journal of Law and Economics* 25, 49–65.
- Landes, W.M., Posner, R.A. (1980). "Joint and multiple torts: an economic analysis". *Journal of Legal Studies* 9, 517–556.
- Landes, W.M., Posner, R.A. (1987). *The Economic Structure of Tort Law*. Harvard University Press, Cambridge, MA.
- Menell, P.S. (1983). "A note on private versus social incentives to sue in a costly legal system". *Journal of Legal Studies* 12, 41–52.
- Oi, W. (1973). "The economics of product safety". *Bell Journal of Economics* 4, 3–28.
- Ordover, J.A. (1978). "Costly litigation in the model of single activity accidents". *Journal of Legal Studies* 7, 243–261.
- Pitchford, R. (1995). "How liable should a lender be? The case of judgment-proof firms and environmental risk". *American Economic Review* 85, 1171–1186.
- Polborn, M.K. (1998). "Mandatory insurance and the judgment-proof problem". *International Review of Law and Economics* 18, 141–146.
- Polinsky, A.M. (1980a). "Resolving nuisance disputes: the simple economics of injunctive and damage remedies". *Stanford Law Review* 32, 1075–1112.
- Polinsky, A.M. (1980b). "Strict liability vs. negligence in a market setting". *American Economic Review (Papers and Proceedings)* 70, 363–367.
- Polinsky, A.M. (1997). "Are punitive damages really insignificant, predictable, and rational?" *Journal of Legal Studies* 26, 663–677.
- Polinsky, A.M., Che, Y.-K. (1991). "Decoupling liability: optimal incentives for care and litigation". *Rand Journal of Economics* 22, 562–570.
- Polinsky, A.M., Rubinfeld, D.L. (1988). "The welfare implications of costly litigation for the level of liability". *Journal of Legal Studies* 17, 151–164.
- Polinsky, A.M., Shavell, S. (1998). "Punitive damages: an economic analysis". *Harvard Law Review* 111, 869–962.
- Posner, R.A. (1972). *Economic Analysis of Law*, 1st edn. Little, Brown and Company, Boston.
- Priest, G.L. (1981). "A theory of the consumer product warranty". *Yale Law Journal* 90, 1297–1352.
- Roberts, M.J., Spence, M. (1976). "Effluent charges and licenses under uncertainty". *Journal of Public Economics* 5, 193–208.
- Rose-Ackerman, S., Geistfeld, M. (1987). "The divergence between the social and private incentives to sue: a comment on Shavell, Menell, and Kaplow". *Journal of Legal Studies* 16, 483–491.
- Rosenberg, D. (1984). "The causal connection in mass exposure cases: a 'public law' vision of the tort system". *Harvard Law Review* 97, 849–929.
- Rubinfeld, D.L. (1987). "The efficiency of comparative negligence". *Journal of Legal Studies* 16, 375–394.
- Sarath, B. (1991). "Uncertain litigation and liability insurance". *Rand Journal of Economics* 22, 218–231.

- Schwartz, A., Wilde, L.L. (1979). "Intervening in markets on the basis of imperfect information: a legal and economic analysis". *University of Pennsylvania Law Review* 127, 630–682.
- Shavell, S. (1980a). "An analysis of causation and the scope of liability in the law of torts". *Journal of Legal Studies* 9, 463–516.
- Shavell, S. (1980b). "Strict liability versus negligence". *Journal of Legal Studies* 9, 1–25.
- 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. (1984a). "A model of the optimal use of liability and safety regulation". *Rand Journal of Economics* 15, 271–280.
- Shavell, S. (1984b). "Liability for harm versus regulation of safety". *Journal of Legal Studies* 13, 357–374.
- Shavell, S. (1985). "Uncertainty over causation and the determination of civil liability". *Journal of Law and Economics* 28, 587–609.
- Shavell, S. (1986). "The judgment proof problem". *International Review of Law and Economics* 6, 45–58.
- Shavell, S. (1987). *Economic Analysis of Accident Law*. Harvard University Press, Cambridge, MA.
- Shavell, S. (1993). "The optimal structure of law enforcement". *Journal of Law and Economics* 36, 255–287.
- 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. (1999). "The level of litigation: private versus social optimality". *International Review of Law and Economics* 19, 99–115.
- Shavell, S. (2000). "On the social function and the regulation of liability insurance". *Geneva Papers on Risk and Insurance, Issues and Practice* 25, 166–179.
- Shavell, S. (2005). "Minimum asset requirements and compulsory liability insurance as solutions to the judgment-proof problem". *Rand Journal of Economics* 36, 63–77.
- Sloan, F.A. (1998). "Automobile accidents, insurance, and tort liability". In: Newman, P. (Ed.), *The New Palgrave Dictionary of Economics and the Law*, vol. 1. Macmillan, London, pp. 140–144.
- Spence, M. (1977). "Consumer misperceptions, product failure, and product liability". *Review of Economic Studies* 44, 561–572.
- Spier, K.E. (1994). "Settlement bargaining and the design of damage awards". *Journal of Law, Economics, & Organization* 10, 84–95.
- Spier, K.E. (1997). "A note on the divergence between the private and the social motive to settle under a negligence rule". *Journal of Legal Studies* 26, 613–621.
- Spier, K.E. (2007). "Litigation". In: Polinsky, A.M., Shavell, S. (Eds.), *Handbook of Law and Economics*, vol. 1. North-Holland, Amsterdam.
- Summers, J. (1983). "The case of the disappearing defendant: an economic analysis". *University of Pennsylvania Law Review* 132, 145–185.
- Sykes, A.O. (1981). "An efficiency analysis of vicarious liability under the law of agency". *Yale Law Journal* 91, 168–206.
- Sykes, A.O. (1984). "The economics of vicarious liability". *Yale Law Journal* 93, 1231–1282.
- Tillinghast-Towers Perrin (2000). *U.S. Tort Cases, 2000: Trends and Findings on the Costs of the U.S. Tort System*. Tillinghast-Towers Perrin, Chicago.
- Tunc, A. (1974). "Introduction". In: *Torts. International Encyclopedia of Comparative Law*, vol. II. J.C.B. Mohr, Tübingen, Germany, Chapter I.
- Viscusi, W.K., Evans, W. (1990). "Utility functions that depend on health status: estimates and economic implications". *American Economic Review* 80, 353–374.
- Weitzman, M.L. (1974). "Prices vs. quantities". *Review of Economic Studies* 41, 477–491.
- Wittman, D. (1977). "Prior regulation versus post liability: The choice between input and output monitoring". *Journal of Legal Studies* 6, 193–212.